UNIVERSIDAD NACIONAL DEL CALLAO

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

ESCUELA ACADÉMICA PROFESIONAL DE MATEMÁTICA

Un Método Proximal Continuo para Minimizar Funciones Cuasi-convexas

Tesis para optar el Título profesional de

Licenciado en Matemática

ROSA QUISPE LLAMOCA

CALLAO - PERÚ

Febrero - 2009

HOJA DE PRESENTACIÓN

Un Método Proximal Continuo para Minimizar Funciones Cuasi-convexas

Rosa Quispe Llamoca

Tesis presentada a consideración del Cuerpo de Docentes de la Facultad de Ciencias Naturales y Matemática de la Universidad Nacional del Callao, como parte de los requisitos para obtener el Título Profesional de Licenciado en Matemática.

Aprobado por:

Lic. Sofía Irena Duran Quiñones.

Lic. Gabriel Rodríguez Varillas.

Mg. Edinson Montoro Alegre.

Dr. Erik Alex Papa Quiroz.

CALLAO - PERÚ

Febrero - 2009

FICHA CATALOGRÁFICA

QUISPE LLAMOCA, ROSA

Un Método Proximal Continuo para Minimizar Funciones Cuasi-convexas, Callao [2009].
XI, 137 p. 29,7 cm (UNAC, Licenciado en Matemática, 2009)

Tesis, Universidad Nacional del Callao, Facultad de Ciencias Naturales y Matemática.
Matemática.
I. UNAC/FCNM II. Título (Serie)

Dedicatoria

A Dios y a mi mamá Belia.

AGRADECIMIENTOS

- ⋆ Gracias a Dios, a mi mamá por su cariño, guía y apoyo, agradeciendole por haberme brindado una carrera y con la promesa de seguir adelante.
- ⋆ A mi asesor de tesis, el Doctor Erik Alex Papa Quiroz, por su valiosa colaboración y dirección en la realización de la tesis.
- ⋆ En general, a mis profesores y amigos que me apoyaron en la realización de esta tesis.

RESUMEN

Un Método Proximal Continuo para Minimizar Funciones Cuasi-convexas

ROSA QUISPE LLAMOCA

Febrero - 2009

Asesor: Dr. Erik Alex Papa Quiroz
Título Obtenido: Licenciado en Matemática

En esta tesis consideramos el siguiente problema de minimización convexa

$$(P) \quad \begin{cases} \min f(x) \\ x \in R^n_+ \end{cases}$$

donde $f : R^n \to (-\infty, \infty]$ es una función convexa propia, cerrada y $R^n_+ = \{x \in R^n : x_j \geq 0,\ j = 1, ..., n\}$. Para resolver el problema (P), estudiamos el siguiente Sistema Dinámico de Lotka-Volterra:

$$(S) \quad \begin{cases} x'_j(t) + \left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right] \frac{\partial f(x(t))}{\partial x_j} = 0, \quad j = 1, ..., n \\ x_j(0) = x_{0_j} \in R^n_{++}, \quad j = 1, .., n. \end{cases}$$

Probamos la existencia global y unicidad de la trayectoria $x : [0, \infty) \to R^n$ para el sistema (S), cuando $f : R^n \to R$ es una función limitada inferiormente en R^n_+ y de clase $C^2(R^n)$. Estudiamos la viabilidad de la trayectoria en el octante positivo y analizamos su comportamiento asintótico donde obtenemos la convergencia global de la trayectoria para un punto óptimo cuando f es convexa. También, extendemos los resultados de convergencia del caso convexo para demostrar que en el caso cuasi-convexo la trayectoria converge a un punto candidato a solución. Estos resultados son útiles en las aplicaciones de la matemática a la economía, por ejemplo las funciones de costo, producción y utilidad, que caracterizan al problema de decisión del consumidor, suelen ser convexas o cuasi-convexas y el conjunto de decisión del consumidor se encuentra generalmente en el octante no negativo R^n_+. Así, nuestro trabajo resolverá en particular problemas en la teoría económica de la decisión.

Palabras Claves:

Sistema Dinámico Lotka-Volterra, método proximal continuo, convergencia asintótica, economía matemática.

ABSTRACT

A Continuous Proximal Method for Minimizing Quasi-convex Functions

ROSA QUISPE LLAMOCA
February - 2009

Adviser: Dr. Erik Alex Papa Quiroz

Obtained Degree: Mathematician

In this thesis, we consider the convex minimization problem

$$(P) \quad \begin{cases} \min f(x) \\ x \in R^n_+ \end{cases}$$

where $f : R^n \to (-\infty, \infty]$ is a closed proper convex function and $R^n_+ = \{x \in R^n : x_j \geq 0,\ j = 1, ..., n\}$. To solve the problem (P), we study the following Lotka-Volterra Dynamical System:

$$(S) \quad \begin{cases} x'_j(t) + \left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right] \frac{\partial f(x(t))}{\partial x_j} = 0, \quad j = 1, ..., n \\ x_j(0) = x_{0_j} \in R^n_{++}, \quad j = 1, .., n \end{cases}$$

We prove the global existence and unicity of the trajectory $x : [0, \infty) \to R^n$ for the system (S), where $f : R^n \to R$ be a $C^2(R^n)$ function which is bounded from below on R^n_+. We study the viability in the positive octant and analyze the asymptotic behavior of the trajectory where we obtain the global convergence of this path to an optimal solution when f is convex. Also, we extend the global converge to a KKT point from the convex case to the quasi-convex one. These results are usefull in the application of the mathematics to the economy. For example cost functions, production and consumer decision problem are often convex or quasi-convex functions on the nonnegative octant R^n_+. Thus, our work will solve, in particular, problems in decision economic theory.

Keywords:

Lotka-Volterra Dynamical System, proximal method continuous asymptotic convergence, Mathematical Economics.

Contenido

Lista de Figuras

Notaciones

En el desarrollo del Trabajo de Tesis utilizamos las siguientes terminologías:

$\bar{R} = R \cup \{\pm\infty\}$.

$R^n = \{x = (x_1, x_2, ..., x_n) : x_i \in R, \forall i = 1, ..., n\}$: espacio euclidiano n-dimensional.

$R^n_+ = \{x = (x_1, x_2, ..., x_n) \in R^n : \ x_i \geq 0 \ \forall i = 1, ..., n\}$: ortante no negativo.

$R^n_{++} = \{x = (x_1, x_2, ..., x_n) \in R^n : x_i > 0, \forall i = 1, ..., n\}$: ortante positivo.

Dado $x \in R^n$, $x \geq 0$ significa $x \in R^n_+$.

Dado $x \in R^n$, $x > 0$ significa $x \in R^n_{++}$.

$B(\bar{x}, \epsilon)$: denota la bola abierta de centro $\bar{x}$ y radio ϵ.

$\frac{\partial f}{\partial v}$: denota la derivada direccional de f.

$\nabla f\,(x)$: denota el gradiente de f en x.

$L\,(f, \alpha)$: denota el conjunto de nivel inferior de f.

$U\,(f, \alpha)$: denota el conjunto de nivel superior de f.

$Y(f, \alpha)$: denota la superficie de nivel de f.

$\precsim$: denota una relación de preferencia.

$\sim$: denota una relación de indiferencia.

Introducción

En diversas áreas de ciencias e ingeniería existen modelos matemáticos que pueden expresarse como un problema de optimización, esto es,

$$\begin{cases} \textit{optimizar } f(x) \\ \quad s.a \\ \quad\ g_i(x) \leq 0 \\ \quad\ h_i(x) = 0 \end{cases}$$

donde *optimizar* significa, minimizar $f(x)$ o maximizar $f(x)$, g_i y h_i son funciones dadas. Un problema de optimización modela así un problema de tomar una decisión óptima para maximizar (ganancias, velocidad, eficiencia, etc.) o minimizar (costos, tiempo, riesgo, error, etc.) un criterio determinado. Las restricciones, g_i y h_i, significan que no cualquier decisión es posible.
La Optimización puede considerarse como la búsqueda de la mejor solución (solución óptima) de un problema, el término mejor depende del contexto en el que se trabaje, podría significar solución que minimiza los costos, o maximiza los beneficios, o que hace que la distancia recorrida sea mínima, etc. Esta introducción sobre lo que se entiende por Optimización refleja claramente las importantísimas e indudables aplicaciones de esta área de las matemáticas a un amplio espectro de problemas.
Una clase bien amplia de problemas de optimización es la optimización convexa donde f es una función convexa, es decir, se cumple

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y), \ \forall \ \lambda \in [0,1],$$

y el conjunto

$$X = \{x : g(x) \leq 0\} \cap \{x : h(x) = 0\}$$

es un conjunto convexo y diferente de vacío. El interés del concepto de funciones convexas es importante porque tiene numerosas aplicaciones en Economía.
Cabe destacar que la genera-lización de funciones convexas son las funciones cuasi-convexas, las cuales también tienen importantes aplicaciones en teoría económica. Las funciones cuasi-convexas tienen un comportamiento similar al de las funciones cuasi-cóncavas, desde que f es cuasi-cóncava si y solo $-f$ es cuasi-convexa. Un ejemplo importante de funciones cuasi-cóncavas se observa en las funciones de utilidad la cual puede representar una relación de preferencia $\precsim$, dada por

$$y \precsim x \Leftrightarrow u(y) \leq u(x).$$

Elegir la mejor alternativa se convierte así en encontrar una alternativa que maximice la función f . Por lo tanto, concluimos que resolver el problema de decisión del consumidor es equivalente a resolver el problema de maximización:

$$\begin{cases} \max f(x) \\ x \in X \end{cases}$$

el cual es también equivalente a

$$\begin{cases} \min -f(x) \\ x \in X \end{cases}$$

donde $-f$ es cuasi-convexa. Este modelo es en particular un problema de optimización cuasi-convexa.
En este trabajo estamos interesados en resolver el problema

$$\begin{cases} \min f(x) \\ x \in R^n_+ \end{cases}$$

donde, $f : R^n \to (-\infty, +\infty]$ es una función convexa, propia, cerrada y $R^n_+ = \{x \in R^n : x_j \geq 0,\ j = 1, ..., n\}$. Para ello presentamos el siguiente sistema

$$(S) \quad \begin{cases} x'_j(t) + \left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right] \frac{\partial f(x(t))}{\partial x_j} = 0 & j = 1, ..., n \\ x_j(0) = x_{0_j} \in R^n_{++} & j = 1, ..., n \end{cases}$$

donde $\mu > 0$, $\nu > 0$ son parámetros dados. Estudiamos este sistema como un método para resolver el problema de minimización convexa y cuasi-convexa, el cual fue introducido por Attouch, H., y Teboulle, M. [3] en el artículo titulado "Regularized Lotka-Volterra Dynamical System as Continuous proximal - Like Method in Optimization".

La tesis está dividida en dos partes. En la primera parte unificamos definiciones, proposiciones, caracterizaciones, teoremas y corolarios de cuasi-convexidad, cuasi-convexidad estricta, cuasi-convexidad semi-estricta, pseudo-convexidad donde presentamos una aplicación de cuasi-convexidad en teoría económica de decisión.
En la segunda parte desarrollamos el artículo de Attouch, H y Teboulle, M. [3], donde se demuestra la existencia y unicidad de la trayectoria cuando f está acotada inferiormente y es de clase $C^2(R^n)$. Además, estudiamos la viabilidad de la trayectoria y el comportamiento asintótico en el caso convexo y cuasi-convexo. El resultado novedoso de esta tesis es la convergencia de la trayectoria del Sistema Dinámico (S) a un punto candidato a solución del problema de minimización cuasi-convexa.

El trabajo está organizado de la siguiente manera:

En el Capítulo 1, presentamos algunos conceptos básicos de convexidad, análisis real y optimización en R^n_+, los cuales serán usados a lo largo de la tesis.

En el Capítulo 2, se desarrolla la teoría económica de decisión, definimos la relación de preferencia, $\precsim$, la cual puede ser representada mediante una función de utilidad.

De esta manera, la relación de preferencia da un criterio de valorización entre dos bienes o mercancías por medio de la función de utilidad. Esta función es cuasi-cóncava y monótona si, la relación de prefencia es convexa y monótona. Este capítulo, justifica el estudio y la importancia de las funciones cuasi-cóncavas.

En el Capítulo 3, definimos las funciones cuasi-convexas por la convexidad de sus conjuntos de nivel inferior, presentamos sus caracterizaciones y propiedades. Además, introducimos definiciones las cuales permitirán obtener proposiciones, teoremas y corolarios de funciones cuasi-convexas. También, definimos cuasi-convexidad estricta, cuasi-convexidad semi-estricta, pseudo-convexidad, cada una de estas con sus respectivas caracterizaciones y propiedades las cuales son de vital importancia para el estudio de optimización cuasi-convexa.

En el Capítulo 4, estudiamos las condiciones de optimalidad para un problema de minimización cuasi-convexa en el octante no negativo. Daremos las condiciones necesarias y suficientes para un mínimo. Además, presentamos modelos matemáticos que pueden ser vistos como problemas de maximización cuasi-cóncava.

En el Capítulo 5, partiremos considerando la función $f : R^n \to R$ como una función suave y el sistema dinámico (S) en el espacio euclidiano R^n, con trayectoria $x : [0, +\infty) \to R^n$ definida por

$$(S) \quad \left\{ \begin{array}{c} x_j'(t) + \left[\frac{x_j(t)}{\mu+\nu x_j(t)}\right] \frac{\partial f(x(t))}{\partial x_j} = 0, \quad j = 1, ..., n \\ x_j(0) = x_{0_j} \in R^p_{++}, \quad j = 1, .., n \end{array} \right.$$

donde $\mu > 0$, $\nu > 0$ son parámetros dados.
Este sistema es derivado de la versión continua de un esquema de discretización implícita basado en un método proximal.
La motivación y la naturaleza del sistema dinámico propuesto es explicado en la Sección 1 del Capítulo 5. En la Sección 2, estudiamos la existencia global y resultados de viabilidad para el sistema (S) con una función suave generalizada f (no necesariamente convexa); partimos con un punto inicial $x_0 \in R^n_{++}$, probamos que la trayectoria $t \to x(t)$ del sistema (S) es definida en R^n_+ y satisface $x(t) \in R^n_{++}$ para todo $t \in R_+$. El comportamiento asintótico de la trayectoria para una función suave es estudiada en la sección 3. En la Sección 4, consideramos el caso convexo y probamos la convergencia de la trayectoria para una solución óptima del problema de minimización convexa sobre el octante no negativo. Mostramos la localización del resultado para el punto límite resultante. En la Sección 5, consideramos el esquema de Euler implícito el cual fue la causa del sistema continuo. En la Sección 6, fijamos $\mu > 0$ y consideramos la familia del sistema regularizado Lotka-Volterra parametrizado por ν y probamos que la resultante genera trayectorias $x^\nu(t)$ que son uniformemente convergente para la solución del sistema clásico regularizado Lotka-Volterra. En la Sección 7, daremos nuevos resultados obtenidos cuando la función es cuasi-convexa, donde se concluye que la trayectoria converge a un punto candidato a solución.

Finalmente, en nuestro estudio concluimos y presentamos la convergencia a un punto óptimo para el caso convexo y como un nuevo resultado tenemos la convergencia a un punto KKT cuando la función es cuasi-convexa. Parte de este material, orientado y asesorado por el Doctor Erik Alex Papa Quiroz, fue presentado en el concurso de Iniciación Científica del IV Congreso Internacional de Matemática Aplicada y Computacional - IV CIMAC ocupando el primer puesto en su categoría.

Revisión de la Bibliografía

Para realizar el siguiente trabajo se ha revisado las siguientes referencias.

- Attouch, H., y Teboulle, M [3] presentan el Sistema Dinámico Lotka Volterra como un método proximal continuo para minimizar funciones convexas, el cual se extenderá para funciones cuasi-convexas.

- Aliprantes, C. D., Broen, D. J., y Birkinshaw, O., [1] hacen un análisis de la función de utilidad como una función cuasi-cóncava, cuyo estudio servirá de justificación para el estudio del método proximal continuo.

- Papa Quiroz, E. A y Oliveira, P.R [38] presentan el método de punto proximal con distancia de Bregman como un método proximal iterativo para minimizar funciones cuasi-convexas, el cual mediante la discretización implícita de Euler se presenta como un método proximal continuo.

Capítulo 1

Preliminares

En este Capítulo, recopilamos las definiciones y resultados en el espacio euclidiano R^n y en espacios más generales que serán usados en el desarrollo de la tesis. En la Sección 1.1 se desarrollan las nociones de convexidad presentando definiciones y teoremas. En la Sección 1.2 presentamos definiciones y teoremas los cuales son resultados de análisis real y definiciones de espacios vectoriales y topológicos. En la Sección 1.3 mencionamos algunas definiciones y teoremas de la teoría de optimización.

1.1 Nociones de Convexidad

Definición 1.1 *Un conjunto $C \subset R^n$ es convexo, si para todo par de puntos x e y en C se tiene*

$$\lambda x + (1-\lambda)\, y \in C, \ \forall \lambda \in [0,1] .$$

Definición 1.2 *El segmento cerrado que une los puntos x e y es denotado por $[x,y]$ y está definido por*

$$[x,y] = \{z = \lambda x + (1-\lambda)\, y \ / \ \lambda \in [0,1]\} .$$

La expresión $\lambda x + (1-\lambda)\, y$ se llama combinación convexa de x e y.

Teorema 1.1 *Sea $C \subset R^n$ un conjunto convexo, entonces se satisfacen las siguientes condiciones:*

i. *rC, es un conjunto convexo para todo $r \in R$.*

ii. *$v + C = \{x = v + c \ / \ c \in C\}$, es un conjunto convexo para $v \in R^n$ un vector fijo dado.*

iii. *Si C_1, C_2 son conjuntos convexos entonces*

$$C_1 + C_2 = \{z = x_1 + x_2 \ / \ x_1 \in C_1 \wedge x_2 \in C_2\}$$

es un conjunto convexo.

Demostración. Ver [10], pág 11. ■

Lema 1.1 *Sea $\{C_\alpha\}$ una colección de conjuntos convexos tal que*

$$C = \bigcap_{\alpha \in I} C_\alpha$$

es no vacío, entonces C es convexo.

Demostración. Ver [9], pág 36. ■

Definición 1.3 *Un hiperplano H_a^α en R^n es definido por el conjunto de puntos que satisfacen la ecuación $\langle a, x\rangle = \alpha$, donde $\alpha \neq 0$ y $\alpha \in R$. De esta manera,*

$$H_a^\alpha = \{x : \langle a, x\rangle = \alpha\}.$$

El vector a se llama el vector normal del hiperplano.

Definición 1.4 *Dos conjuntos X e Y son separados por un hiperplano H_a^α si,*

$$\langle a, x\rangle \geq \alpha, \forall x \in X \ \wedge \ \langle a, y\rangle \leq \alpha, \forall y \in Y.$$

Teorema 1.2 *Sea $C \subset R^n$ un conjunto convexo e $y \notin C$. Entonces existe un hiperplano H_a^α tal que para todo $x \in C$*

$$\langle a, x\rangle \leq \alpha \ \wedge \ \langle a, y\rangle = \alpha.$$

Si $int(C) \neq \emptyset$, entonces para todo $x \in int(C)$

$$\langle a, x\rangle < \alpha.$$

Demostración. Ver [9], pág 51. ■

Definición 1.5 *Sea $C \subset R^n$ un conjunto convexo. Un punto $x \in C$ se llama un punto extremo de C, si x no puede ser expresado como una combinación convexa estricta de dos puntos distintos de C.*

Definición 1.6 *Sea $f : C \subset R^n \to R$ una función definida sobre un conjunto convexo C, sus conjuntos de nivel inferior, están determinados por*

$$L(f, \alpha) = \{x \in C : f(x) \leq \alpha\}$$

para todo $\alpha \in R$.

Definición 1.7 *Sea $f : C \subset R^n \to R$ una función definida sobre un conjunto convexo C, sus conjuntos de nivel superior, están determinados por*

$$U(f, \alpha) = \{x \in C : \alpha \leq f(x)\}$$

para todo $\alpha \in R$.

1.2 Resultados de Análisis

Definición 1.8 *Dado un punto $x_0 \in R^n$ y un número real $\delta > 0$, definimos:*

1. *La bola abierta de centro x_0 y radio δ, $B_\delta(x_0)$, por:*

$$B_\delta(x_0) = \{x \in R^n : \|x - x_0\| < \delta\}.$$

2. *La bola cerrada de centro x_0 y radio δ, $\bar{B}_\delta(x_0)$, por:*

$$\bar{B}_\delta(x_0) = \{x \in R^n : \|x - x_0\| \leq \delta\}.$$

Definición 1.9 *Un conjunto $X \subset R^n$ es abierto si para cada $x \in X$ existe $\delta > 0$ tal que $B_\delta(x_0) \subset X$.*

Definición 1.10 *Un conjunto $X \subset R^n$ es llamado cerrado si su complemento es abierto.*

Definición 1.11 *Sea $X \subset R^n$, una vecindad de un conjunto $A \subset X$ es un conjunto abierto $U \subset X$ tal que $A \subset U$.*

Definición 1.12 *Una función $f : X \subset R^n \to R$ definida en el conjunto X es semicontinua inferior en X si su conjunto de nivel inferior $L(f, \alpha)$ es cerrado relativo en X para todo α.*
Analogamente, f es semicontinua superior en X si su conjunto de nivel superior $U(f, \alpha)$ es cerrado relativo en X para todo α real. Una función es continua si es semicontinua inferior y semicontinua superior.
Equivalentemente, se puede considerar que f es semicontinua inferior en X si y solo si para todo $x_0 \in X$ y toda sucesión $x_k \to x_0$ tenemos

$$f(x_0) = f(\lim x_k) \leq \liminf f(x_k)$$

Definición 1.13 *(Caminos Diferenciales) Un camino en R^n es una aplicación continua $f : I \to R^n$ cuyo dominio es un intervalo $I \subset R$ tal que para cada $t \in I$ tenemos*

$$f(t) = (f_1(t), .., f_n(t)),$$

donde $f_1, .., f_n : I \to R$, las funciones coordenadas de f, son continuas.

Definición 1.14 *Sea $f : U \subset R^n \to R$ una función real definida en un subconjunto abierto U. Dado un punto $a \in U$, la i-ésima derivada parcial de f en el punto a es el límite*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t}$$

cuando tal límite existe.

Definición 1.15 *Una función $f : X \to R$, definida en un conjunto $X \subset R^n$, se dice que es continua en el punto $a \in X$ cuando, para cada $\epsilon > 0$ dado, se puede obtener un $\delta > 0$ tal que si $x \in X$ y $\| x - a\| < \delta$ implica que $|f(x) - f(a)| < \epsilon$. En lenguaje simbólico esto es equivalente a:*

$$\forall \epsilon > 0, \ \exists \delta = \delta(\epsilon) > 0 \ / \ x \in X, \| x - a\| < \delta \Rightarrow |f(x) - f(a)| < \epsilon.$$

Se dice que $f : X \to R$ es una función continua en X si, y sólo si es continua en a, para todo $a \in X$.

Sea $x = (x_1, x_2, \ldots, x_n) \in R^n$, se definen las siguientes normas

$$\| x\| = \sqrt{\sum_{k=1}^{n} x_k^2}\,, \qquad \| x\|_S = \sum_{i=1}^{n} |x_i|\,, \qquad \| x\|_\infty = \max_{i=1,\ldots,n} |x_i|,$$

llamadas: norma euclidiana, norma de la suma y norma del máximo, respectivamente.

Definición 1.16 *Dado un conjunto $X \neq \emptyset$, una familia de subconjuntos de X, τ, definida por*

$$\tau = \{G_i \subset X \ / \ i \in I\}$$

es una topología en X si verifica los siguientes axiomas:

1. $\emptyset \in \tau$, $X \in \tau$.

2. *Para cualquier $J \subset I$, $J \neq \emptyset$ se tiene*

$$\bigcup_{i \in I} G_i \in \tau.$$

3. *Para cualquier $i, j \in I$ se tiene*

$$G_i \cap G_j \in \tau$$

El par (X, τ) es un espacio topológico y los elementos de τ se les llama abiertos del espacio topológico (X, τ).

1.3 Resultados de Optimización

Definición 1.17 *Diremos que un punto $x_0 \in D$ es*

a. *minimizador local en D si*

$$f(x_0) \leq f(x), \ \forall x \in B_\delta(x_0) \cap D.$$

b. *minimizador global en D si*

$$f(x_0) \leq f(x), \ \forall x \in D.$$

Teorema 1.3 *(Condición necesaria de primer orden) Supongamos que $f : R^n \to R$ es una función diferenciable en el punto $x_0 \in R^n$. Si x_0 es un mínimo local del problema de minimización irrestricto entonces,*

$$\nabla f(x_0) = 0.$$

Demostración. Ver [16], pág 15. ∎

Teorema 1.4 *(Condición necesaria de segundo orden) Supongamos que la función $f : R^n \to R$ es dos veces diferenciable en el punto $x_0 \in R^n$. Si x_0 es un mínimo local del problema de minimización sin restricciones, entonces $\nabla f(x_0) = 0$ y la Hesiana de f en el punto x_0 es semidefinida positiva, es decir,*

$$v^T \nabla^2 f(x_0) v \geq 0, \ \forall v \in R^n.$$

Demostración. Ver [16], pág 16. ∎

Teorema 1.5 *(Condición suficiente de segundo orden) Sea $f : R^n \to R$ una función dos veces diferenciable en el punto $x_0 \in R^n$. Si $\nabla f(x_0) = 0$ y si la matriz Hesiana de f en x_0 es definida positiva, es decir,*

$$v^T \nabla^2 f(x_0) v > 0, \ \forall v \in R^n \setminus \{0\}$$

entonces x_0 es un minimizador local estricto de f del problema de minimización sin restricciones.

Demostración. Ver [16], pág 17. ∎

Teorema 1.6 *Sea $f : C \subset R^n \to R$ una función diferenciable en un conjunto abierto y convexo C. Entonces f es convexa si y solo si, para todo x_0 en C se satisface*

$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle, \ \forall x \in C.$$

Demostración. Ver [9], pág 109. ∎

Teorema 1.7 *Sea $f : C \subset R^n \to R$ una función convexa definida en el conjunto convexo C*

i. *Si f alcanza un mínimo local en x_0, entonces f alcanza un mínimo global en x_0.*

ii. *Si f es estrictamente convexa y f alcanza un mínimo en x_0, entonces x_0 es único.*

Demostración. Ver [9], pág 137. ∎

Definición 1.18 *Un punto $\bar{x} \in R^n$ es un punto estacionario del problema*

$$\left\{ \begin{array}{c} \min f(x) \\ x \in D \end{array} \right. \tag{1.1}$$

donde $D = \{x \in R^n : h(x) = 0, g(x) \leq 0\}$, si $\bar{x} \in D$ y existen $\bar{\lambda} \in R^n$ y $\bar{\mu} \in R_+^m$ tales que satisfacen

$$\left\{ \begin{array}{c} \nabla f(\bar{x}) + \bar{\lambda}^T \nabla h(\bar{x}) + \bar{\mu}^T \nabla g(\bar{x}) = 0 \\ \bar{\mu}_i g_i(\bar{x}) = 0, \ i = 1, ..., m \end{array} \right.$$

Los elementos $\bar{\lambda}$ y $\bar{\mu}$ son llamados multiplicadores de Lagrange asociados al punto estacionario $\bar{x}$.

Teorema 1.8 *Sea el problema de minimización*

$$(PL) \left\{ \begin{array}{c} \min f(x) \\ g(x) \leq 0 \end{array} \right. \tag{1.2}$$

donde f, g son funciones convexas. Sea $\bar{x} \in R^n$ y $\bar{\mu} \in R_+^m$ tal que

i. $g(\bar{x}) \leq 0$,

ii. $\bar{\mu} \geq 0$

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

iv. $\nabla f(\bar{x}) + \bar{\mu}^t \nabla g(\bar{x}) = 0$.

Entonces $\bar{x}$ es una solución del problema de minimización (PL).

Demostración. Ver [9], pág 159. ■

Capítulo 2

Minimización Cuasi-convexa en la Teoría Económica de Decisión

Las aplicaciones de la optimización matemática son muy importantes y de gran interés en la comunidad científica motivo por el cual su estudio se ha intensificado en los últimos años. Uno de estos modelos es el problema de decisión del consumidor (o llamado también agente económico) que consiste en maximizar la utilidad entre todos los bienes que se encuentran a su disposición, es decir, se trata de escoger la mejor opción entre aquellas que sean posibles. Hay tres elementos que caracterizan el problema de decisión:

- El conjunto de elección, que nos dice cual es el universo de alternativas del consumidor.
- El criterio de valorización, que estudia la manera en la que el consumidor evalúa las diferentes alternativas que se le ofrecen. Este criterio se define mediante una cierta relación binaria $\precsim$, interpretada económicamente como "es al menos tan preferible", que refleja sus preferencias sobre el conjunto de elección. En economía esta relación es llamada preferencia.
- Las restricciones, que delimitan el conjunto de oportunidades donde el consumidor puede efectivamente elegir.

A partir de este problema, presentamos las siguientes definiciones y sus respectivas notaciones utilizadas en el desarrollo de este capítulo.

Definición 2.1 *Dado un conjunto $X \neq \emptyset$, definimos una relación binaria $\precsim$ de X en X como un subconjunto $\precsim \subset X \times X$.*

Notación 2.1 *Si $(x, y) \in \precsim$ escribiremos $x \precsim y$.*

Definición 2.2 *Dado un conjunto $X \neq \emptyset$, una relación $\precsim$ definida en X es:*

1. *Reflexiva : Si $x \precsim x$ $\forall x \in X$.*
2. *Transitiva : Si $x \precsim y \ \wedge \ y \precsim z$ implica $x \precsim z$ $\forall x, y, z \in X$.*

3. *Completa : Si* $x \precsim y$ *o* $y \precsim x$ $\forall x, y \in X$.

A continuación, daremos la definición de relación de preferencia.

Definición 2.3 *Una relación de preferencia en un conjunto no vacío* X *es una relación reflexiva, completa y transitiva en* X.

Sea $\precsim$ una relación de preferencia en X, la notación $y \precsim x$ se interpreta como "x es al menos tan preferido como y" o "x no es peor que y". Además, la notación $y \prec x$ interpretada como "x es preferido a y" o "x es mejor que y" significa

$$(y, x) \in \precsim \ \wedge \ (x, y) \notin \precsim .$$

Sin embargo, si tenemos $y \precsim x$ y $x \precsim y$ diremos que "x es indiferente a y", lo cual denotaremos como $x \sim y$.
Si x es un elemento del conjunto X, entonces el conjunto

$$\{y \in X : x \prec y\}$$

es llamado conjunto que es mas preferido de x y el conjunto

$$\{y \in X : y \prec x\}$$

es llamado el conjunto que es menos preferido de x. Analogamente, para la relación $\precsim$ se define el conjunto $\{y \in X : x \precsim y\}$ denominado el conjunto que es al menos tan preferido que x y el conjunto $\{y \in X : y \precsim x\}$ es denominado el conjunto que no es al menos tan preferible como x.
Hasta ahora el conjunto X es un conjunto arbitrario pero, si X tiene una estructura topológica (X es un espacio topológico) la continuidad y semicontinuidad de las preferencias $\precsim$ se definen de la siguiente manera.

Definición 2.4 *Una relación de preferencia* $\precsim$ *en un espacio topológico* X *se dice*

1. *Semicontinua superior: si para cada* $x \in X$ *el conjunto*

$$\{y \in X : x \precsim y\}$$

es cerrado.

2. *Semicontinua inferior: si para cada* $x \in X$ *el conjunto*

$$\{y \in X : y \precsim x\}$$

es cerrado.

3. *Continua: si* $\precsim$ *es semicontinua superior y semicontinua inferior, es decir, para cada* $x \in X$ *los conjuntos*

$$\{y \in X : x \precsim y\} \ y \ \{y \in X : y \precsim x\}$$

son cerrados.

Afirmamos que, el complemento del conjunto

$$\{y \in X : x \precsim y\}$$

es el conjunto

$$\{y \in X : y \prec x\}.$$

En efecto, consideremos

$$z \notin \{y \in X : x \precsim y\}$$

desde que $\precsim$ es completo

$$(z, x) \in \precsim \ \wedge \ (x, z) \notin \precsim$$

entonces,

$$(z, x) \in \prec$$

luego,

$$z \in \{y \in X : y \prec x\}.$$

Analogamente, se observa que el complemento del conjunto

$$\{y \in X : y \precsim x\}$$

está dado por el conjunto

$$\{y \in X : x \prec y\}.$$

En efecto, sea

$$z \notin \{x \in X : y \precsim x\}$$

desde que $\precsim$ es completo

$$(x, z) \in \precsim \ \wedge \ (z, x) \notin \precsim$$

entonces,

$$(x, z) \in \prec$$

luego,

$$z \in \{y \in X : x \prec y\}.$$

Inmediatamente, se deduce que una relación de preferencia en un espacio topológico X es continuo si y solo si para cada $x \in X$, los conjuntos

$$\{y \in X : x \prec y\} \ \wedge \ \{y \in X : y \prec x\}$$

son abiertos. Las preferencias continuas se caracterizan de la siguiente manera.

Teorema 2.1 *Para una relación de preferencia $\precsim$ en un espacio topológico X los siguientes enunciados son equivalentes:*

1. *La preferencia $\precsim$ es continua.*

2. *La preferencia $\precsim$ (considerada como un subconjunto de $X \times X$) es cerrada en $X \times X$.*

3. Si $y \prec x$ en X, entonces existen vecindades disjuntas U_x y U_y de x e y respectivamente, tal que $a \in U_x$ y $b \in U_y$ implica $b \prec a$.

Demostración. Primero probaremos que el item 1 implica 3.
Supongamos que existe algún $z \in X$ talque $y \prec z \prec x$. Consideremos las siguientes vecindades

$$U_x = \{a \in X : z \prec a\} \ \wedge \ U_y = \{b \in X : b \prec z\}$$

de ahí $a \in U_x$ y $b \in U_y$ entonces $b \prec z \prec a$ por transitividad de $\precsim$, tenemos $b \prec a$.
Ahora, supongamos que no existe $z \in X$ que satisface $y \prec z \prec x$.
En este caso, consideremos

$$U_x = \{a \in X : y \prec a\} \ \wedge U_y = \{b \in X : b \prec x\}$$

como $a \in U_x$ y $b \in U_y$, tenemos

$$y \prec a \ \wedge \ b \prec x$$

desde que $y \prec x$ (por hipótesis) afirmamos que $b \prec a$.
En efecto, supongamos que no se cumple $b \prec a$ lo cual implica

$$(b, a) \notin \precsim \ \vee \ (a, b) \in \precsim .$$

Si $(b, a) \notin \precsim$, por completitud de $\precsim$ tenemos $(a, b) \in \precsim$. De ahí,

$$y \prec a \precsim b \prec x$$

la transitividad de $\prec$ implica que $y \prec a \prec x$ lo cual es una contradicción porque en este caso se trabajó bajo el supuesto que no existe ningún elemento entre y y x.
Por otro lado, afirmamos que los conjuntos

$$U_x = \{a \in X : z \prec a\} \ \wedge \ U_y = \{b \in X : b \prec z\}$$

son disjuntos,

$$U_x \cap U_y = \emptyset.$$

En efecto, supongamos que $U_x \cap U_y \neq \emptyset$, es decir,

$$\exists \ m \in X \setminus m \in U_x \ \wedge \ m \in U_y$$

entonces,

$$z \prec m \ \wedge \ \ m \prec z$$

por transividad de $\prec$,

$$m \prec m,$$

lo cual es una contradicción pues $\prec$ es irreflexiva.
Acontinuación, demostraremos que el item 3 implica 2.
Tomemos (x_α, y_α) una sucesión de $\precsim$ por definición tenemos,

$$x_\alpha \precsim y_\alpha$$

talque $(x_\alpha, y_\alpha) \to (x, y)$. Nos proponemos demostrar que $(x, y) \in \precsim$.
Supongamos que $(x, y) \notin \precsim$, por completación de $\precsim$ tenemos $(y, x) \in \precsim$ entonces $y \prec x$.
Sabemos que $x_\alpha \to x$, desde que $x \in U_x$, por definición existe $n_0 \in N$ talque $\alpha > n_0$ entonces $x_\alpha \in U_x$. Por otro lado, como $y_\alpha \to y$ y desde que $y \in U_y$, por definición existe $n_1 \in N$ talque $\alpha > n_1$ entonces $y_\alpha \in U_y$.
Consideremos $n_2 = \max\{n_0, n_1\}$ talque $\alpha > n_2$ entonces

$$x_\alpha \in U_x \ \wedge \ y_\alpha \in U_y$$

de la hipótesis del item 3 tenemos $y_\alpha \prec x_\alpha$ lo cual implica

$$(y_\alpha, x_\alpha) \in \precsim \ \wedge \ (x_\alpha, y_\alpha) \notin \precsim$$

es una contradicción desde que $(x_\alpha, y_\alpha) \in \precsim$. Por lo tanto, $(x, y) \in \precsim$.
Finalmente, probaremos que el item 2 implica 1.
Sea $\{y_\alpha\}$ una sucesión del conjunto $\{y \in X : y \precsim x\}$ satisfaciendo $y_\alpha \to z$ en X, desde que

$$\{(y_\alpha, x)\} \in \precsim$$

luego,

$$y_\alpha \precsim x$$

se satisface $(y_\alpha, x) \to (z, x)$ en $X \times X$. Por la cerradura de $\precsim$ en $X \times X$, tenemos $(z, x) \in \precsim$, es decir $z \precsim x$. Por lo que queda demostrado que el conjunto

$$\{y \in X : y \precsim x\}$$

es cerrado. Por tanto, $\precsim$ es semicontinua inferior.
Igualmente, sea $\{y_\alpha\}$ una sucesión de $\{y \in X : x \precsim y\}$ satisfaciendo $y_\alpha \to z$ en X, desde que,

$$\{(x, y_\alpha)\} \in \precsim$$

luego,

$$x \precsim y_\alpha$$

se satisface $(x, y_\alpha) \to (x, z)$ en $X \times X$.
Por la cerradura de $\precsim$ en $X \times X$ tenemos $(x, z) \in \precsim$, es decir $x \precsim z$. Por lo que queda demostrado que el conjunto

$$\{y \in X : x \precsim y\}$$

es cerrado. Por tanto, $\precsim$ es semicontinua superior.
Por la semicontinuidad superior e inferior, concluimos que la preferencia $\precsim$ es continua. ∎

La siguiente definición, considera como una relación de preferencia puede ser representada por una función de utilidad.

Definición 2.5 *La función $u : X \to R$ se le llama función de utilidad representando una relación de preferencia $\precsim$ en un conjunto X, si para todo $x, y \in X$*

$$y \precsim x \Leftrightarrow u(y) \leq u(x).$$

En el caso de relación estricta $\prec$ se tiene

$$y \prec x \Leftrightarrow u(y) < u(x).$$

Como la convexidad es una herramienta interesante en los problemas económicos definamos entonces la convexidad de $\precsim$ en un conjunto convexo X de un espacio vectorial.

Definición 2.6 *Una relación de preferencia definida en un subconjunto convexo X de un espacio vectorial es:*

1. *Convexa; si $x \precsim y$, $x \precsim z$ en X y $0 \leq \alpha \leq 1$ implica que $x \precsim \alpha y + (1-\alpha) z$.*

2. *Estrictamente convexa; si $x \precsim y$, $x \precsim z$ y $y \neq z$ implica que $x \prec \alpha y + (1-\alpha) z$ para $0 < \alpha < 1$.*

Es claro que la relación de preferencia definida en un conjunto convexo X es convexa si y solo si el conjunto $\{y \in X : x \precsim y\}$ es convexo para cada $x \in X$.

Definición 2.7 *Una función definida en un conjunto convexo no vacío C de un espacio vectorial es*

1. *Cuasi-cóncava; si para todo $x, y \in C$ y $0 \leq \alpha \leq 1$ se tiene*

$$u(\alpha x + (1-\alpha) y) \geq \min\{u(x), u(y)\}$$

2. *Estrictamente cuasi-cóncava; si para todo $x, y \in C$ con $x \neq y$ y $0 < \alpha < 1$ se tiene*

$$u(\alpha x + (1-\alpha) y) > \min\{u(x), u(y)\}$$

3. *Cóncava; si para todo $x, y \in C$ y $0 \leq \alpha \leq 1$ se tiene*

$$u(\alpha x + (1-\alpha) y) \geq \alpha u(x) + (1-\alpha) u(y)$$

4. *Estrictamente cóncava; si para todo $x, y \in C$ con $x \neq y$ y $0 < \alpha < 1$ se tiene*

$$u(\alpha x + (1-\alpha) y) > \alpha u(x) + (1-\alpha) u(y)$$

Una función $u : C \to R$ definida en un subconjunto convexo C de un espacio vectorial se dice convexa si $-u$ es cóncava, es decir, para todo $x, y \in C$ y $0 \leq \alpha \leq 1$

$$u(\alpha x + (1-\alpha) y) \leq \alpha u(x) + (1-\alpha) u(y)$$

Similarmente, una función u se dice estrictamente convexa si $-u$ es estrictamente cóncava.
Podemos afirmar, que toda función cóncava es cuasi-cóncava.
En efecto, consideremos que $u : C \to R$ es una función cóncava y sean $x, y \in C$ con $0 \leq \alpha \leq 1$ tomemos $m = \min\{u(x), u(y)\}$

$$u(\alpha x + (1-\alpha) y) \geq \alpha u(x) + (1-\alpha) u(y) \geq \alpha m + (1-\alpha) m = m.$$

Para probar que la recíproca es falsa, basta considerar el siguiente ejemplo.

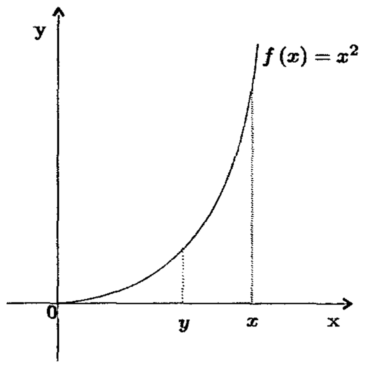

Figura 2.1: Función Cuasi-cóncava, no cóncava

Ejemplo 2.1 *Sea la función $u : [0, \infty) \to R$ definida por $u(x) = x^2$ la cual es una función cuasi-cóncava pero no cóncava.*

En efecto, sin pérdida de generalidad consideremos $y \leq x$ entonces $x - y \geq 0$

$$0 \leq \lambda (x - y) \leq x - y$$

sumando x,

$$y \leq y + \lambda (x - y) \leq x$$

desde que $y \geq 0$ elevamos al cuadrado

$$y^2 \leq (\lambda x + (1 - \lambda) y)^2 \leq x^2$$

luego,

$$f(y) \leq f(\lambda x + (1 - \lambda) y) \leq f(x), \ \forall \lambda \in [0, 1]$$

de aquí,

$$\min \{f(x), f(y)\} \leq f(\lambda x + (1 - \lambda) y), \ \forall \lambda \in [0, 1].$$

Por tanto, f es una función cuasi-cóncava. Sin embargo, la función no es cóncava, es decir, no satisface

$$f(\lambda x + (1 - \lambda) y) \geq \lambda f(x) + (1 - \lambda) f(y), \ \forall \lambda \in [0, 1].$$

Basta considerar,

$$x = m, \ y = m + 1 \ \wedge \ \lambda = \frac{1}{2}$$

con $m \geq 0$, entonces

$$f\left(\frac{1}{2} m + \frac{1}{2}(m + 1)\right) = f\left(\frac{2m + 1}{2}\right) \geq \frac{1}{2} f(m) + \frac{1}{2} f(m + 1)$$

lo que implica,

$$\left(\frac{2m + 1}{2}\right)^2 \geq \frac{1}{2} m^2 + \frac{1}{2}(m + 1)^2$$

lo cual es equivalente a,

$$\left(m+\frac{1}{2}\right)^2 \geq \left(m+\frac{1}{2}\right)^2+\frac{1}{4}$$

Por tanto, se observa que f no satisface la definición de concavidad.

La definición anterior nos permite afirmar que una función de utilidad sobre una relación de preferencia convexa es una función cuasi-cóncava. Por tanto, enunciemos el siguiente teorema.

Teorema 2.2 *Para un conjunto convexo C de un espacio vectorial y una función $u : C \rightarrow R$ entonces se satisfacen los siguientes enunciados:*

1. *La función u es cuasi-cóncava si y solo si, la relación de preferencia definida por u es convexa.*
2. *La función u es estrictamente cuasi-cóncava si y solo si, la relación de preferencia definida por u es estrictamente convexa.*

Demostración.

1. Supongamos que u es una función cuasi-cóncava y que

$$y \precsim x \ \wedge y \precsim z \ en \ C.$$

luego,

$$u(y) \leq u(x) \ \wedge \ u(y) \leq u(z),$$

usando la hipótesis de cuasi-concavidad de u tenemos

$$u(y) \leq \min\{u(x), u(z)\} \leq u(\alpha x + (1-\alpha) z), \forall \in [0,1].$$

De aquí,

$$u(y) \leq u(\alpha x + (1-\alpha) z) \ \forall, \ \alpha \in [0,1].$$

De la definición de función de utilidad se tiene

$$y \precsim \alpha x + (1-\alpha) z, \ \forall \ \alpha \in [0,1]$$

Así, queda probado la convexidad de $\precsim$.
Reciprocamente, consideremos que la relación de preferencia $\precsim$ definida por u es convexa. Sean $x, y \in C$, sin pérdida de generalidad supongamos que $u(y) \leq u(x)$ por definición de función de utilidad $y \precsim x$. Así, tenemos que

$$y \precsim x \ \wedge \ y \precsim y$$

De la convexidad de $\precsim$ tenemos

$$y \precsim \alpha x + (1-\alpha) y, \ \forall \ \alpha \in [0,1]$$

Por tanto, de la definición de función de utilidad

$$\min\{u(x), u(y)\} = u(y) \leq u(\alpha x + (1-\alpha) y), \ \forall \alpha \in [0,1]$$

Entonces, queda probado que la función u es cuasi-cóncava.

2. Supongamos que u es una función estrictamente cuasi-cóncava y consideremos en C, $x \neq z$ tal que

$$y \precsim x \wedge y \precsim z$$

Por definición de función de utilidad se obtiene

$$u(y) \leq u(x) \ \wedge \ u(y) \leq u(z)$$

desde que u es estrictamente cuasi-cóncava

$$u(y) \leq \min\{u(x), u(z)\} < u(\alpha x + (1-\alpha) z), \ \forall \alpha \in (0,1)$$

Por transitividad en $\leq$

$$u(y) < u(\alpha x + (1-\alpha) z), \ \forall \ \alpha \in (0,1)$$

De la definición de función de utilidad

$$y \prec \alpha x + (1-\alpha) z, \ \forall \ \alpha \in (0,1)$$

Entonces, queda probado que $\precsim$ es estrictamente convexa.
Reciprocamente, asumiremos que la relación de preferencia $\precsim$ definida por u es estrictamente convexa. Sean $x, y \in C$, con $x \neq y$ sin pérdida de generalidad supongamos que $u(y) \leq u(x)$ por definición de función de utilidad $y \precsim x$. Considerando

$$y \precsim x \ \wedge \ y \precsim y$$

De la convexidad estricta de $\precsim$ resulta

$$y \prec \alpha x + (1-\alpha) y$$

Por lo tanto, de la definición de función de utilidad

$$\min\{u(x), u(y)\} = u(y) < u(\alpha x + (1-\alpha) y)$$

Entonces, queda demostrado que la función u es estrictamente cuasi-cóncava.

■

Definición 2.8 *Un espacio vectorial ordenado E es un espacio vectorial con la relación de orden $\leq$ tal que satisface las siguientes propiedades:*

- *Si $y \leq x$ en E entonces $y + z \leq x + z$ para todo $z \in E$.*
- *Si $y \leq x$ en E entonces $\alpha y \leq \alpha x$ para todo $\alpha \in R$.*

2164

Definamos, el conjunto E^+ por:

$$E^+ = \{x \in E : 0 \leq x\}$$

el cual, es conocido como el cono positivo de E y sus elementos son llamados vectores positivos. Un ejemplo importante es el espacio vectorial ordenado $E = R^n$. El orden es definido por $y = (y_1, .., y_n) \leq (x_1, .., x_n) = x$ si y solo si $y_i \leq x_i$ para todo $i = 1, 2, ..., n$. El cono positivo de R^n es denotado por R^n_+, esto es,

$$R^n_+ = \{x = (x_1, .., x_n) : 0 \leq x_i, \ \forall i = 1, .., n\}.$$

Notemos que $y < x$ en R^n si y solo si $y_i \leq x_i$ para todo i y $y_i < x_i$ para algún i. A partir de estas definiciones definimos la monotonicidad de las preferencias.

Definición 2.9 *Una relación de preferencia $\precsim$ en un subconjunto no vacío X de un espacio vectorial ordenado, se dice que es*

1. *Monótona; si $x, y \in X$ y $x < y$ implica $x \precsim y$.*

2. *Estrictamente monótona; si $x, y \in X$ y $x < y$ implica $x \prec y$.*

Toda preferencia estrictamente monótona es monótona. Sin embargo, la preferencia monótona no es necesariamente estrictamente monótona. Por ejemplo, consideremos la preferencia en R^2_+ definida por la función de utilidad

$$u(x_1, x_2) = x_1 x_2.$$

Sean

$$\begin{aligned} x_1 &= 1 \qquad & x_2 &= 0, \\ y_1 &= 2 \qquad & y_2 &= 0 \end{aligned}$$

claramente, $(1, 0) < (2, 0)$. De la definición de u tenemos

$$u(1, 0) \leq u(2, 0)$$

entonces,

$$(1, 0) \precsim (2, 0).$$

Además, observamos que

$$(2, 0) \precsim (1, 0)$$

lo cual implica que no se satisface

$$(1, 0) \prec (2, 0).$$

Por tanto, se satisface $(1, 0) < (2, 0)$ y $\neg((1, 0) \prec (2, 0))$, es decir, la relación de preferencia no es estrictamente monótona sin embargo, es monótona.
Introducimos una nueva definición

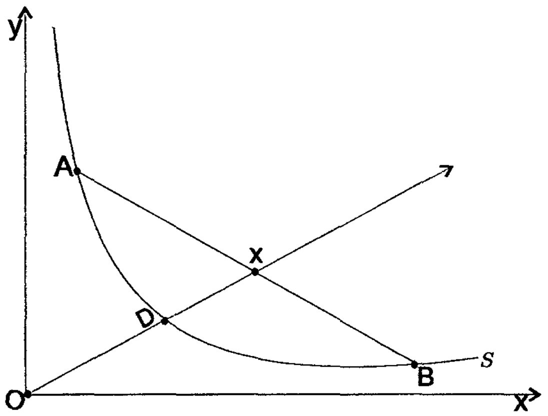

Figura 2.2: Convexidad por el origen

Definición 2.10 *(Convexidad por el origen) Una curva S es convexa por el origen si existe a lo más un punto de la curva entre el segmento formado por el origen y el punto de intersección del segmento formado por dos puntos de S.*

La curva mostrada en la gráfica es convexa por el origen, desde que A y B son dos puntos en la curva, entonces el rayo pasando por el origen O y x (punto del segmento AB) intercepta a la curva en el punto D.
En base a esta definición podemos dar mas resultados de funciones de utilidad cuasi-cóncavas.

Teorema 2.3 *Sea $u : C \to R$ una función de utilidad definida en un subconjunto convexo C del cono positivo de un espacio vectorial ordenado. Si u es estrictamente monótona y cuasi-cóncava entonces sus curvas de nivel son convexas por el origen.*

Demostración. Sean $x, y \in C$ talque $u(x) = u(y) = c$ consideremos

$$z = \alpha x + (1 - \alpha) y$$

para algún $0 < \alpha < 1$.
Desde que u es cuasi-cóncava se tiene

$$u(z) = u(\alpha x + (1 - \alpha) y) \geq \min \{u(x), u(y)\} = c.$$

Por la monotonicidad estricta de u se concluye que el rayo $\{\lambda z : \lambda \geq 0\}$ no intercepta al conjunto de nivel $\{a \in C : u(a) = c\}$ en un punto fuera del segmento formado por 0 y z.
En efecto, supongamos que el rayo $\{\lambda z : \lambda \geq 0\}$ intercepta al conjunto de nivel $\{a \in C : u(a) = c\}$ fuera del segmento $\overline{0z}$ entonces existe $\overline{\lambda} > 1$ tal que $u(\overline{\lambda} z) = c$
Por otro lado, se tiene $0 < z < \overline{\lambda} z$ entonces desde que u estrictamente monótona y creciente por ser función de utilidad $u(z) < u(\overline{\lambda} z)$ entonces

$$c \leq u(z) < u(\overline{\lambda} z) = c$$

Así $c > c$, lo cual es una contradicción. Por lo tanto, el teorema queda demostrado.
■

Definición 2.11 *Sea $\precsim$ una relación de preferencia en un subconjunto X de un espacio vectorial E. Entonces un vector $v \in E$ se dice extremadamente preferible para $\precsim$ cuando:*

1. $x + \alpha v \in X$ *para todo* $x \in X$ *y todo* $\alpha > 0$.
2. $x \prec x + \alpha v$ *para todo* $x \in X$ *y todo* $\alpha > 0$.

Gracias a esta definición enunciamos el siguiente teorema que es importante para preferencias definidas en un cono convexo de un espacio vectorial finito dimensional.

Teorema 2.4 *Para una preferencia continua $\precsim$ definida en el cono positivo R^n_+ del espacio vectorial R^n se satisfacen los siguientes enunciados:*

1. *Si $\precsim$ es convexa, monótona con un vector extremadamente preferible, entonces $\precsim$ puede ser representada por una función de utilidad continua monótona y cuasi-cóncava.*
2. *Si $\precsim$ es estrictamente convexa y estrictamente monótona, entonces $\precsim$ puede ser representada por una función de utilidad continua, estrictamente monótona y estrictamente cuasi-cóncava.*

Demostración. Sea $\precsim$ una relación de preferencia continua, monótona y convexa la cual tiene un vector extremedamente preferible v, es decir:

- $x + \alpha v \in R^n_+, \ \forall x \in R^n_+, \ \forall \alpha > 0.$
- $x \prec x + \alpha v, \ \forall x \in R^n_+, \ \forall \alpha > 0.$

Considerando $e = v + (1, ..., 1)$

- $\alpha (1, ..., 1) \in R^n_+ \wedge x + \alpha v \in R^n_+$ entonces
$$x + \alpha v + \alpha (1, ..., 1) \in R^n_+,$$
así,
$$x + \alpha e \in R^n_+.$$
- Desde que,
$$x + \alpha v < x + \alpha v + \alpha (1, ..., 1),$$
por monotonicidad de $\precsim$ tenemos
$$x + \alpha v \precsim x + \alpha v + \alpha (1, ..., 1)$$
y como
$$x \prec x + \alpha v$$
entonces por transitividad tenemos,
$$x \prec x + \alpha e, \ \forall \alpha > 0.$$

Así, e también es extremadamente preferible donde $e = (e_1, ..., e_n)$.
De esta manera, podemos asumir que existe un vector extremadamente preferible e satisfaciendo $e_i > 0$ para todo i.
Ahora para cada $x \in R^n_+$ definimos el conjunto

$$\{\alpha > 0 : x \precsim \alpha e\}$$

desde que $e_i > 0$ y $x_i > 0$ por el Lema Arquimediano existe $\alpha \in N$ talque $x_i < \alpha e_i$ entonces $x < \alpha e$ por monotonicidad $x \precsim \alpha e$ se concluye que el conjunto es no vacío.
Observamos que el conjunto anteriormente definido es no vacío y además tiene como cota inferior 0 y por el axioma del supremo este conjunto tiene ínfimo.
A partir de esto definimos

$$u(x) = \inf \{\alpha > 0 : x \precsim \alpha e\}$$

el cual está bien definido.
Afirmamos que $x \sim u(x)\, e$.
En efecto, desde que $\precsim$ es continua entonces $\{y \in R^n_+ : x \precsim y\}$ es cerrado.
Por otro lado, si $u(x) > 0$ entonces para todo ϵ suficientemente pequeño tal que $u(x) - \epsilon > 0$ se tiene $(u(x) - \epsilon)\, e \precsim x$. Supongamos que la afirmación anterior no es cierta, es decir, $x \precsim (u(x) - \epsilon)\, e$
Desde que $u(x) - \epsilon > 0$ podemos considerar $\alpha = u(x) - \epsilon$ talque $x \precsim \alpha e$, de la definición de $u(x)$ como el ínfimo de este conjunto

$$u(x) \leq u(x) - \epsilon.$$

Donde se concluye $\epsilon < 0$ lo cual es una contradicción. Entonces no es cierto que $x \precsim (u(x) - \epsilon)\, e$ por lo que queda demostrada la afirmación.
Si $u(x) = 0$ entonces de la hipótesis de $x \geq 0$ y la monotonicidad de $\precsim$ implican $o \precsim x$ Por tanto,

$$0 = u(x)\, e \precsim x$$

Por lo tanto, $u(x)\, e \sim x$.
Hacemos la siguiente observación la que usaremos para demostrar resultados posteriores. Si $0 \leq \alpha$ y $0 \leq \beta$, entonces $\beta e \precsim \alpha e$ si y solo si $\beta \leq \alpha$.
En efecto, supongamos que $\beta e \precsim \alpha e$ implica $\alpha < \beta$.
Entonces $\beta - \alpha > 0$ y desde que e es extremadamente preferible cosideremos $\overline{\alpha} = \beta - \alpha$ y αe un vector de R^n obtenemos

$$\alpha e \prec \alpha e + (\beta - \alpha)\, e = \beta e$$

Lo cual es una contradicción, por lo que queda demostrada la afirmación.
En particular, se ha demostrado que para cada $x \in R^n$ existe un escalar $u(x)$ tal que $x \sim u(x)\, e$.

Es claro que la función $u : R^n_+ \to R$ definida anteriomente es una función de utilidad representando $\precsim$.
En efecto, si $x \precsim y$ entonces $x \sim u(x)$ y $y \sim u(y)$ por la definición de indiferencia $\sim$ obtenemos:

$$u(x)\, e \precsim x \precsim y \precsim u(y)\, e$$

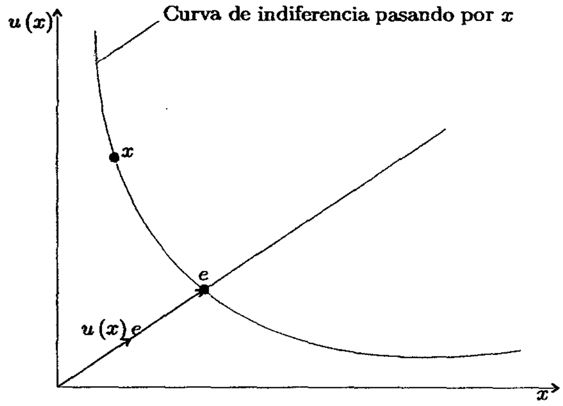

Figura 2.3: Curva de indiferencia

Probaremos la cotinuidad de la función de utilidad a partir de las siguientes identidades:
La primera identidad es la siguiente:

$$\left\{x \in R_+^n : u(x) \leq r\right\} = \left\{x \in R_+^n : x \precsim re\right\}$$

En efecto, sea $z \in \left\{x \in R_+^n : u(x) \leq r\right\}$, es decir $u(z) \leq r$ entonces $u(z)e \leq re$ de monotonicidad $u(z)e \precsim re$, desde que $z \sim u(z)e$ obtenemos $z \precsim u(z)e \precsim re$. Por lo tanto, de la observación anterior

$$z \in \left\{x \in R_+^n : x \precsim re\right\}$$

Reciprocamente, sea $z \in \left\{x \in R_+^n : x \precsim re\right\}$, es decir $z \precsim re$, desde que $u(z)e \sim z$ obtenemos $u(z)e \precsim z \precsim re$. Por lo tanto, de la observación anterior

$$z \in \left\{x \in R_+^n : u(x) \leq r\right\}$$

La segunda identidad es la siguiente:

$$\left\{x \in R_+^n : r \leq u(x)\right\} = \left\{x \in R_+^n : re \precsim x\right\}$$

En efecto, sea $z \in \left\{x \in R_+^n : r \leq u(x)\right\}$, es decir $r \leq u(z)$ entonces $re \leq u(z)e$ de monotonicidad $re \precsim u(z)e$, desde que $u(z)e \sim z$ obtenemos $re \precsim u(z)e \precsim z$. Por lo tanto, de la observación anterior

$$z \in \left\{x \in R_+^n : re \precsim x\right\}$$

Reciprocamente, sea $z \in \left\{x \in R_+^n : re \precsim x\right\}$, es decir $re \precsim z$, desde que $z \sim u(z)e$ obtenemos $re \precsim z \precsim u(z)e$. Por lo tanto, de la observación anterior

$$\left\{x \in R_+^n : r \leq u(x)\right\}$$

Ahora probaremos la continuidad de la función de utilidad a partir de estas identidades y de la hipótesis de continuidad de $\precsim$.
Consideremos una sucesión (x_n) en R_+^n tal que $x_n \to x$ por demostrar que

$$u(x_n) \to u(x).$$

Primero estudiaremos la existencia de tal límite, como u es una función de utilidad es monótona creciente además por su definición es acotada inferiormente por 0 y superiormente por un $\alpha > 0$. Desde que $u(x_n)$ es monótona y acotada entonces es convergente. Desde que $x \sim u(x)\, e$ obtenemos

$$x_n \precsim u(x_n) \ \wedge \ u(x_n) \precsim x_n$$

De la continuidad de $\precsim$

$$x \precsim \lim u(x_n)\, e \ \wedge \ \lim u(x_n)\, e \precsim x$$

Usando las identidades anteriores

$$u(x) \leq \lim u(x_n) \ \wedge \ \lim u(x_n) \leq u(x)$$

Por tanto, $\lim u(x_n) = u(x)$ entonces u es continua.
La monotonicidad se obtiene desde u es una función de utilidad es monótona creciente.
Por último probaremos que u es cuasi-cóncava, sin pérdida de generalidad consideremos $x \leq y$ entonces por monotonicidad de $\precsim$ se tiene $x \precsim y$ desde que u es función de utilidad $u(x) \leq u(y)$.
De la convexidad de $\precsim$ como $x \precsim y$ y $x \precsim x$ implica $x \precsim \alpha x + (1-\alpha)$ y desde que u es una función de utilidad

$$\min\{u(x), u(y)\} = u(x) \leq u(\alpha x + (1-\alpha)\, y)$$

Lo que queda entonces demostrado. ∎

El siguiente ejemplo, muestra que no toda relación de preferencia puede ser representada por una función de utilidad.

Ejemplo 2.2 *Sea $X = [0,1] \times [0,1]$ y consideremos la siguiente relación*

$$(x_1, x_2) \succ (y_1, y_2)$$

si $x_1 > y_1$ o si $x_1 = y_1$ y $x_2 > y_2$. Esta es una relación de preferencia que se conoce como la relación de preferencia lexicográfica. Veremos que esta relación no puede representarse mediante una función de utilidad.

En efecto, supongamos que tal representación es posible, y por lo tanto existe una función de utilidad. Entonces, para cada $r \in [0,1]$ se tiene

$$(r, 0) \prec (r, 1) \text{ pues } r = r \text{ y } 1 > 0 \ \text{ entonces}$$

Por definición de función de utilidad,

$$u[(r,0)] < u[(r,1)]$$

Consideremos $d(r) = u[(r,1)] - u[(r,0)] > 0$ para todo r.
En consecuencia, afirmamos que $[0,1] = \bigcup_{n=1}^{\infty} \{r \ / \ d(r) > \frac{1}{n}\}$,
Primero, probaremos que

$$[0,1] \subset \bigcup_{n=1}^{\infty} \left\{r \ / \ d(r) > \tfrac{1}{n}\right\}$$

Consideremos $x \in [0,1]$, de ahí $(x,0) \prec (x,1)$

$$u[(x,0)] < u[(x,1)]$$

Lo cual es equivalente,

$$u[(x,1)] - u[(x,0)] > 0$$

luego,

$$d(x) > 0$$

entonces, por el Lema Arquimediano $d(x) > \frac{1}{n}$ para algún $n \in N$ se tiene

$$x \in \bigcup_{n=1}^{\infty} \left\{r \ / \ d(r) > \tfrac{1}{n}\right\}.$$

Reciprocamente, probaremos que

$$\bigcup_{n=1}^{\infty} \left\{r \ / \ d(r) > \frac{1}{n}\right\} \subset [0,1]$$

Sea $x \in \bigcup_{n=1}^{\infty} \left\{r \ / \ d(r) > \frac{1}{n}\right\}$ entonces $d(x) > \frac{1}{n}$ para algún $n \in N$. Lo cual implica,

$$u[(x,1)] - u[(x,0)] > \frac{1}{n} > 0$$

para algún $n \in N$. Por tanto, $u[(x,1)] > u[(x,0)]$.
Notamos que, $x \in [0,1]$ pues u está definida de la siguiente manera

$$u : [0,1] \times [0,1] \to R$$

Como $[0,1]$ es no numerable entonces para algún $n \in N$, el conjunto $\left\{r \ / \ d(r) > \frac{1}{n}\right\}$ es no numerable. Supongamos que el conjunto,

$$\left\{r \ / \ d(r) > \frac{1}{m}\right\} = S_m$$

es no numerable. Tomando $u[(1,1)] - u[(0,0)] = k$ y sea N entero mayor que $km+1$, es decir, $N > km+1$ lo cual implica $\frac{N-1}{m} > k$.
Elijamos un conjunto de N elementos $\{r_1, r_2, ..., r_N\}$ en el conjunto $S_m = \left\{r \ / \ d(r) > \frac{1}{m}\right\}$ de manera que $r_1 < r_2 < ... < r_N$.
Puesto que,

$$(r_n, 0) > (r_{n-1}, 1) \text{ pues } r_n > r_{n-1}$$

luego,

$$u[(r_n, 0)] > u[(r_{n-1}, 1)].$$

Por lo tanto,

$$u\left[(r_n,0)\right]-u\left[(r_{n-1},0)\right]>u\left[(r_{n-1},1)\right]-u\left[(r_{n-1},0)\right]>\tfrac{1}{m}.$$

En consecuencia,

$$\begin{aligned}
k=u\left[(1,1)\right]-u\left[(0,0)\right] &= \{u\left[(1,1)\right]-u\left[(r_N,0)\right]\}+\{u\left[(r_N,0)\right]-u\left[(r_{N-1},0)\right]\}\\
&+ \{u\left[(r_{N-1},0)\right]-u\left[(r_{N-2},0)\right]\}+\{u\left[(r_{N-2},0)\right]-u\left[(r_{N-3})\right]\}\\
&+ \ldots+\{u\left[(r_2,0)\right]-u\left[(r_1,0)\right]\}+\{u\left[(r_1,0)\right]-u\left[(0,0)\right]\}\\
&> 0+\frac{1}{m}+\frac{1}{m}+\ldots\frac{1}{m}>\frac{N-1}{m}>k,
\end{aligned}$$

lo cual es una contradicción, por lo tanto, $\prec$ no se puede representar por una función de utilidad.

Capítulo 3

Funciones Cuasi-convexas

En la actualidad existe un buen número de definiciones equivalentes y caracterizaciones de funciones cuasi-convexas. Formalmente las definimos por la convexidad de sus conjuntos de nivel inferior. Debemos resaltar, que las funciones cuasi-convexas son usadas como la generalización de funciones convexas en teoría económica.

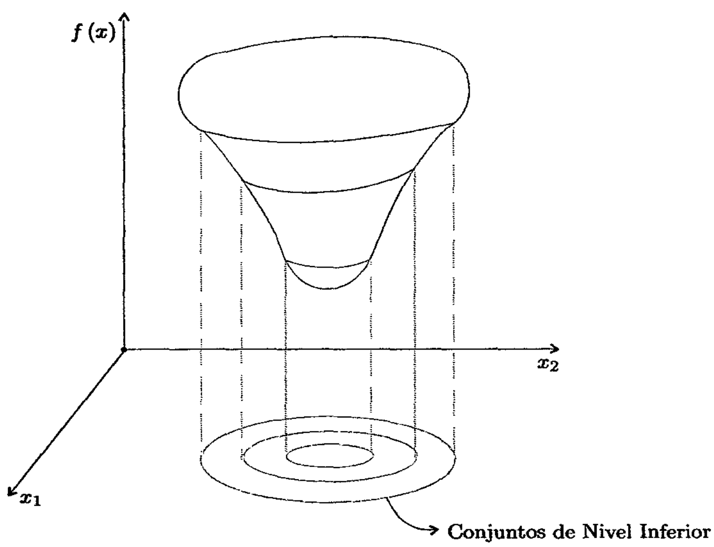

Figura 3.1: Función Cuasi-convexa

Definición 3.1 *Una función $f : C \subset R^n \to R$ definida sobre un conjunto convexo C, se dice que es cuasi-convexa si sus conjuntos de nivel inferior*

$$L(f, \alpha) = \{x \in C : f(x) \leq \alpha\}$$

son conjuntos convexos para todo $\alpha \in R$ (ver figura 3.1). Similarmente, f se dice que es cuasi-cóncava si sus conjuntos de nivel superior

$$U(f, \alpha) = \{x \in C : \alpha \leq f(x)\}$$

son convexos para todo $\alpha \in R$.

Las funciones definidas de esta manera fueron enunciados por primera vez por De Finetti [22]. Las propiedades de funciones cuasi-convexas, las relaciones entre ellas y las funciones convexas fueron estudiadas por Frenchel [20].
Resultados equivalentes pueden obtenerse fácilmente para funciones cuasi-cóncavas, desde que f es cuasi-cóncava si y solo si $-f$ es cuasi-convexa.

Teorema 3.1 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto convexo C. Entonces f es cuasi-convexa si y solo si*

$$f(\alpha x + (1-\alpha) y) \leq \max\{f(x), f(y)\}$$

$\forall\ x, y \in C, \forall\ \lambda \in [0,1]$.

Demostración. Supongamos que f es cuasi-convexa, entonces

$$L(f, \alpha) = \{x \in C : f(x) \leq \alpha\}$$

es convexo para todo α real. Sean $x, y \in C$ y $\overline{\alpha} = \max\{f(x), f(y)\}$.
De la convexidad del conjunto $L(f, \bar{\alpha})$ tenemos que

$$f(\lambda x + (1-\lambda) y) \leq \overline{\alpha} = \max\{f(x), f(y)\}, \ \forall\ \lambda \in [0,1].$$

Reciprocamente, sea $L(f, \alpha)$ el conjunto de nivel inferior de f y sean $x, y \in L(f, \alpha)$ entonces

$$f(x) \leq \alpha \ \wedge \ f(y) \leq \alpha$$

por hipótesis y por las desigualdades anteriores tenemos

$$f(\lambda x + (1-\lambda) y) \leq \max\{f(x), f(y)\} \leq \alpha$$

Así, el conjunto $L(f, \alpha)$ es un conjunto convexo. ■

En las figuras 3.2 y 3.3 se presentan ejemplos gráficos de funciones cuasi-convexas.

Observación 3.1 *Toda función convexa es cuasi-convexa. En efecto, sea f convexa, entonces*

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) \leq \max\{f(x), f(y)\}, \ \forall x, y \in C, \forall \lambda \in [0,1].$$

El recíproco no es cierto, para ello considere por ejemplo la función $f(x) = -x^2$ que no es convexa pero es cuasi-convexa (ver ejemplo 2.1). Algunos resultados de la composición de funciones cuasi-convexas fueron desarrolladas por Schaible [44].
Uno de estos resultados se muestra en la siguiente proposición.

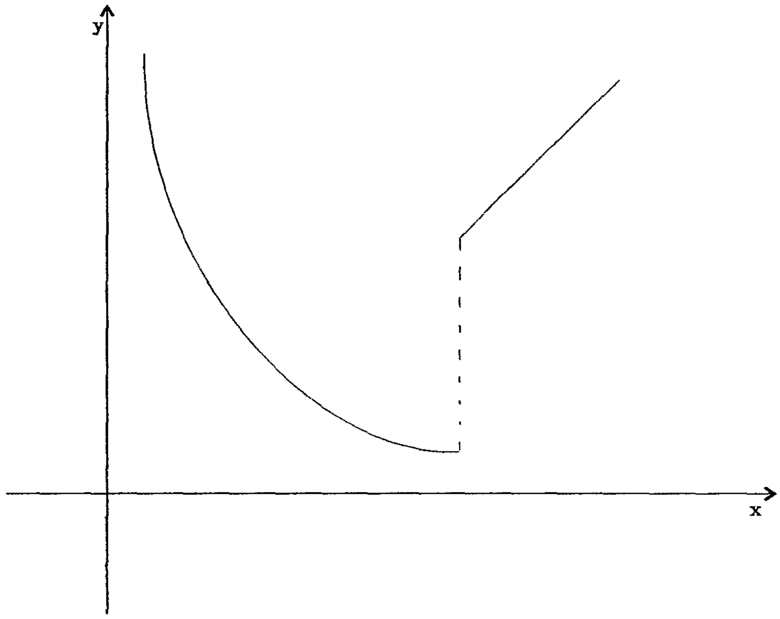

Figura 3.2: Función cuasi-convexa discontinua

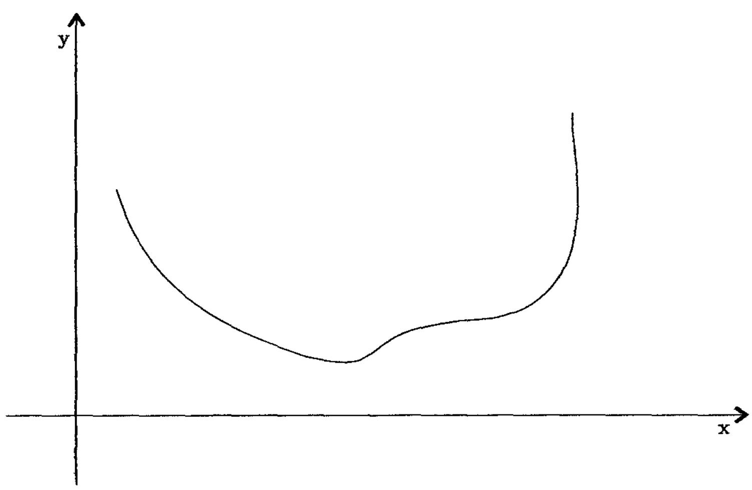

Figura 3.3: Funcion cuasi-convexa continua

Proposición 3.1 *(Fenchel, 1951) Sea $\phi : C \subset R^n \rightarrow R$ una función cuasi-convexa y sea $f : D \subset R \rightarrow R$ una función no decreciente en D conteniendo la imagen de ϕ, entonces la función composición $f \circ \phi$ es también cuasi-convexa.*

Demostración. Sea ϕ una función cuasi-convexa en C, entonces para todo $x, y \in C$ y $\lambda \in [0,1]$, tenemos

$$\phi(\lambda x + (1-\lambda) y) \leq \max\{\phi(x), \phi(y)\}.$$

Como f es no decreciente entonces

$$f(\phi(\lambda x + (1-\lambda) y)) \leq f(\max\{\phi(x), \phi(y)\}).$$

Lo cual es equivalente a

$$f(\phi(\lambda x + (1-\lambda) y)) \leq \max\{f(\phi(x)), f(\phi(y))\}.$$

Por lo tanto queda demostrado que $f \circ \phi$ es cuasi-convexa. ∎

Observación 3.2 *La proposición anterior no puede enunciarse para funciones convexas, esto es, si ϕ es convexa, entonces $f \circ \phi$ no necesariamente es convexa. Para ver esto, consideremos el siguiente la función*

$$\begin{array}{rcl} f : R & \rightarrow & R \\ y & \rightarrow & f(y) = y^3 \end{array}$$

la cual es no decreciente y la función afin ϕ

$$\begin{array}{rcl} \phi : C & \rightarrow & R \\ x & \rightarrow & \phi(x) = x. \end{array}$$

A pesar que ϕ es convexa y f es no decreciente se tiene que $f(\phi(x)) = x^3$ no es convexa, es decir $f \circ \phi$ no es convexa.

De la proposición anterior se deduce que toda transformación no decreciente de funciones convexas es cuasi-convexa.

Observación 3.3 *Cuando la función es convexa es bien conocido que todo mínimo local es global. Sin embargo, para funciones cuasi-convexas este resultado no es verdadero. Basta considerar la función cuasi-convexa.*

$$f(x) = \begin{cases} x^2 & -1 \leq x \leq 1 \\ 2 & 1 < x \leq 3 \\ -(x-4)^2 + 6 & 3 < x \leq 4. \end{cases}$$

Observamos que $\bar{x} = 2$ es un mínimo local, pero no un mínimo global. El siguiente resultado muestra que resultados globales pueden ser obtenidos para mínimos locales estrictos.

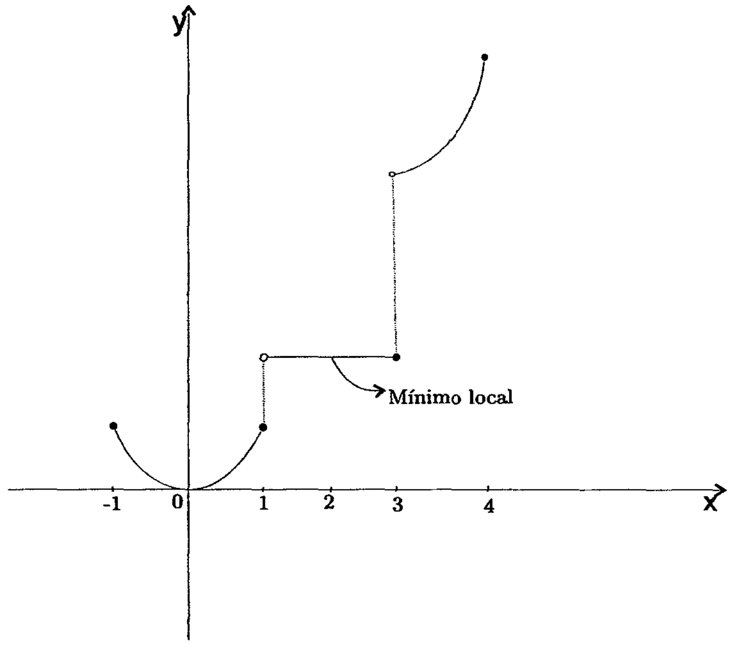

Figura 3.4: Mínimo local, que no es mínimo global

Proposición 3.2 *Sea $f : C \subset R^n \to R$ una función cuasi-convexa definida en el conjunto convexo C. Si x_* es un mínimo local estricto de f entonces x_* es un mínimo global estricto de f en C.*

Demostración. Si x_* es un mínimo local estricto entonces, existe $\delta > 0$, tal que para todo $x \neq x_*$ en el conjunto $C \cap B_\delta(x_*)$ se tiene

$$f(x_*) < f(x). \tag{3.1}$$

Supongamos que x_* no es un mínimo global estricto de f entonces existe un $\bar{x} \in C$, $\bar{x} \neq x_*$ tal que

$$f(x_*) > f(\bar{x}).$$

Por cuasi-convexidad de f

$$f(\lambda\bar{x} + (1-\lambda)x_*) \leq \max\{f(\bar{x}), f(x_*)\} = f(x_*), \ \forall\, \lambda \in [0,1]. \tag{3.2}$$

En particular consideremos

$$\lambda \in \left(0, \frac{\delta}{\|\bar{x} - x_*\|}\right) \cap (0,1),$$

entonces,

$$\hat{x} = \lambda\bar{x} + (1-\lambda)x_* \in C \cap B_\delta(\bar{x}).$$

de (3.2) obtenemos que

$$f(\hat{x}) \leq f(x_*), \ \hat{x} \in B_\delta(x_*) \cap C$$

Lo cual contradice (3.1). Por lo tanto, queda demostrado que si x_* es un mínimo local estricto, también es un mínimo global estricto. ∎

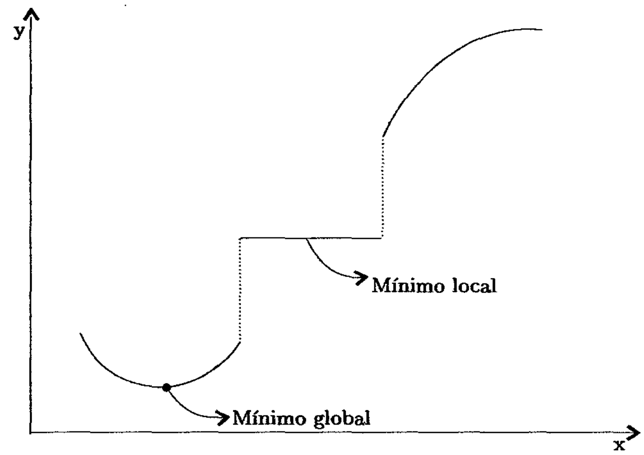

Figura 3.5: Función cuasi-convexa en una variable

Definición 3.2 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto convexo C, entonces f tiene la propiedad de segmento máximo si f alcanza su máximo sobre todo segmento cerrado $[x_1, x_2] \in C$ en algún punto $x \in [x_1, x_2]$.*

Proposición 3.3 *Una función cuasi-convexa definida en el conjunto convexo $C \subset R^n$ tiene la propiedad de segmento máximo.*

Demostración. Sea $[x_1, x_2] \subset C$ un segmento cerrado entonces para todo punto $x \in [x_1, x_2]$ tenemos

$$x = \lambda x_1 + (1 - \lambda) x_2 \in C, \ \forall \lambda \in [0, 1],$$

desde que f es cuasi-convexa

$$f(\lambda x_1 + (1 - \lambda) x_2) \leq \max\{f(x_1), f(x_2)\}$$

De esta manera, el máximo de f sobre $[x_1, x_2]$ es alcanzado por x_1 o x_2. ∎

El siguiente concepto, concierne a funciones de variable real y fue introducido por Diewert, Avriel y Zang [18].

Definición 3.3 *Sea $f : I \subset R \to R$ una función definida en el intervalo abierto I, entonces f alcanza un máximo local semi-estricto en $x_0 \in I$ si existen $x_1, x_2 \in I$, $x_1 < x_0 < x_2$ tal que*

$$f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x_0), \ \forall \lambda \in [0, 1]$$

y

$$\max\{f(x_1), f(x_2)\} < f(x_0)$$

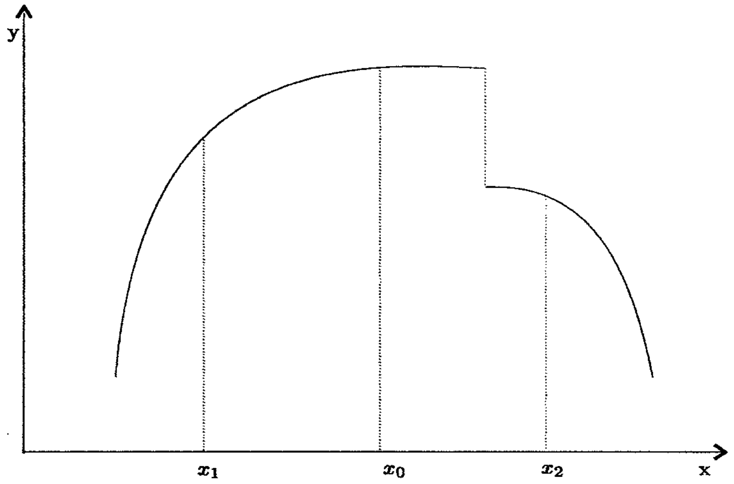

Figura 3.6: Máximo local semi-estricto

Proposición 3.4 *Sea $f : I \subset R \to R$ una función definida en el intervalo I. Si f es cuasi-convexa entonces no existe un punto x_0 que sea un máximo local semi-estricto de f.*

Demostración. Si f es cuasi-convexa, entonces para cualquier par de puntos $x_1, x_2 \in I$ tenemos

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \max\{f(x_1), f(x_2)\}, \ \forall \lambda \in [0,1].$$

Sea $x_0 = \lambda^* x_1 + (1-\lambda^*) x_2$, para $\lambda^* \in [0,1]$ entonces

$$f(x_0) \leq \max\{f(x_1), f(x_2)\}$$

Desde que f no satisface la segunda condición de máximo local semi-estricto entonces f no tiene un máximo local semi-estricto. ∎

Teorema 3.2 *(Diewert, Avriel y Zang, 1981) Una función $f : C \subset R^n \to R$ definida en el conjunto abierto y convexo C es cuasi-convexa, si y solo si,*

1. *Tiene la propiedad de segmento máximo.*
2. *Para todo $x \in C$ y todo $v \in R^n$ tal que $v^T v = 1$, además $\bar{t} \in R$, $\bar{t} > 0$ satisfaciendo $x + \bar{t}v \in C$ la función $F(t) = f(x + tv)$ no alcanza un máximo local semi-estricto en $t \in (0, \bar{t})$*

Demostración. Primero, supongamos que f es cuasi-convexa en un conjunto C abierto y convexo.
El item 1 se satisface, desde que f es cuasi-convexa y por Proposición 3.3.
Para demostrar el item 2, afirmamos que $F(t) = f(x + tv)$ es cuasi-convexa en $(0, \bar{t})$. Probaremos inicialmente

$$\forall\, t \in (0, \bar{t}) \Rightarrow x + tv \in C$$

En efecto, sea $0 < t < \bar{t}$ entonces,

$$0 < \frac{t}{\bar{t}} < 1$$

de aquí,

$$\begin{aligned} x + tv = x + \frac{t}{\bar{t}}(\bar{t}v) &= \left(\frac{t}{\bar{t}} + 1 - \frac{t}{\bar{t}}\right)x + \frac{t}{\bar{t}}(\bar{t}v) \\ &= \frac{t}{\bar{t}}(x + \bar{t}v) + \left(1 - \frac{t}{\bar{t}}\right)x \in C. \end{aligned}$$

La pertenencia anterior se cumple desde que C es un conjunto convexo.
Ahora probaremos que F es cuasi-convexa en $(0, \bar{t})$. Por tanto, debemos demostrar que

$$F(\lambda t_1 + (1-\lambda)t_2) \leq \max\{F(t_1), F(t_2)\}, \ \forall\, t_1, t_2 \in (0, \bar{t}).$$

En efecto, sean $t_1, t_2 \in (0, \bar{t})$ luego

$$x + t_1 v, x + t_2 v \in C,$$

de la cuasi-convexidad de f tenemos,

$$f(\lambda(x + t_1 v) + (1-\lambda)(x + t_2 v)) \leq \max\{f(x + t_1 v), f(x + t_2 v)\}, \ \forall\, \lambda \in [0,1].$$

La expresión anterior es equivalente a

$$F(\lambda t_1 + (1-\lambda)t_2) \leq \max\{F(t_1), F(t_2)\}.$$

La cual demuestra que F es cuasi-convexa en $(0, \bar{t})$.
Por tanto, usando la Proposición 3.4 la función $F(t) = f(x + tv)$ no alcanza un máximo local semi-estricto en $t \in (0, \bar{t})$.
Reciprocamente, sean $x_1, x_2 \in C$ tal que $x_1 \neq x_2$, consideremos

$$v = \frac{x_2 - x_1}{\|x_2 - x_1\|} \ \wedge \ \bar{t} = \|x_2 - x_1\|. \tag{3.3}$$

Como $x_1 + tv \in C$, definamos

$$F(t) = f(x_1 + tv), \ \forall t \in (0, \bar{t}).$$

Desde que f tiene la propiedad de segmento máximo sobre $[0, \bar{t}]$, F alcanza su máximo en algún $t_0 \in [0, \bar{t}]$, es decir,

$$F(t) \leq F(t_0), \ \forall\, t \in [0, \bar{t}]. \tag{3.4}$$

Ya que F no alcanza un máximo local semi-estricto en $t_0 \in (0, \bar{t})$, entonces

$$F(t_0) \leq \max\{F(0), F(\bar{t})\}. \tag{3.5}$$

De (3.4) y (3.5),

$$F((1-\lambda)\bar{t}) = F(t) \leq \max\{F(0), F(\bar{t})\}. \tag{3.6}$$

Usando (3.3) y reemplazando en (3.6) tenemos

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{f(x_1), f(x_2)\}, \forall\lambda \ \in [0,1].$$

Por lo tanto, f es cuasi-convexa. ∎

Proposición 3.5 *Sea $f : C \subset R^n \to R$ una función semicontinua superior definida en C. Entonces f tiene la propiedad de segmento máximo.*

Demostración. Sea $[x_1, x_2] \subset C$ un segmento cerrado. Afirmamos que f es acotada superiormente. Procediendo por contradicción, supongamos que existe

$$\{x_k\} \subset [x_1, x_2] \ \wedge \ \lim f(x_k) = +\infty.$$

Como $\{x_k\} \subset [x_1, x_2]$ y $[x_1, x_2]$ es cerrado entonces,

$$\lim x_k = \bar{x} \ \wedge \ \bar{x} \in [x_1, x_2].$$

Por la semicontinuidad superior de f

$$+\infty = \limsup f(x_k) \leq f(\bar{x})$$

Lo cual es una contradicción. Asi, f es acotada superiormente.
Desde que f es acotada superiormente, por el axioma del supremo, f posee supremo, es decir, existe $\alpha \in R$ tal que

$$\alpha = \sup_{x \in [x_1, x_2]} f(x).$$

Por lo tanto, existe $\{x_l\} \subset [x_1, x_2]$ tal que

$$\lim f(x_l) = \alpha.$$

Por la cerradura de C, existe $\{x_l\} \subset [x_1, x_2]$ tal que

$$\lim x_l = x_* \ \wedge \ x_* \in [x_1, x_2].$$

De la semicontinuidad superior de f tenemos

$$\alpha = \limsup f(x_l) \leq f(x_*).$$

Entonces, existe $\{x_*\} \in [x_1, x_2]$ tal que

$$f(x) \leq \alpha \leq f(x_*) \ \forall \ x \in [x_1, x_2].$$

Por tanto, el máximo de f en $[x_1, x_2]$ es alcanzado en x_*. ∎

Definición 3.4 *Sea $f : I \subset R \to R$ definida en el intervalo I. Esta función es unimodal en I si existe $x_* \in I$ donde f alcanza su mínimo en I y para $x_1, x_2 \in I$ tal que $x_1 < x_2$ se tiene*

$$x_2 \leq x_* \text{ implica que } f(x_2) \leq f(x_1)$$

y

$$x_* \leq x_1 \text{ implica que } f(x_1) \leq f(x_2)$$

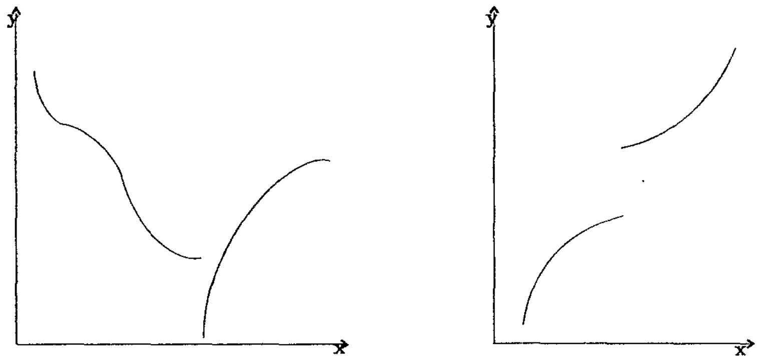

Figura 3.7: Funciones Unimodales

La definición anterior significa que una función de variable real unimodal es dividida en dos subintervalos (uno de ellos puede ser vacío) por el punto que alcanza el mínimo en I, donde el subintervalo de la izquierda de x hace que la función sea no creciente y el subintervalo de la derecha de x hace que la función sea no decreciente.

Proposición 3.6 *(Martos, 1975) Sea $f : I \subset R \to R$ definida en el intervalo abierto I y supongamos que alcanza su mínimo en un punto $x_* \in I$. Entonces f es cuasi-convexa si y solo si es unimodal en I.*

Demostración. Supongamos que f es cuasi-conveva y sean $x_1, x_2 \in I$ talque $x_1 < x_2$. Si $x_2 \leq x_*$ entonces

$$x_1 < x_2 \leq x_*$$

luego,

$$f(x_2) \leq \max\{f(x_1), f(x_*)\}$$

desde que f alcanza su mínimo en x_* tenemos,

$$f(x_2) \leq f(x_1). \tag{3.7}$$

Por otro lado, si $x_* \leq x_1$ entonces

$$x_* < x_1 \leq x_2$$

luego,

$$f(x_1) \leq \max\{f(x_2), f(x_*)\}$$

desde que f alcanza su mínimo en x_* tenemos,

$$f(x_1) \leq f(x_2). \tag{3.8}$$

De (3.7) y (3.8) concluimos que f es unimodal.
Reciprocamente, supongamos que f es unimodal, sean $x_1, x_2 \in I$ tal que $x_1 < x_2$ y definamos

$$x(\lambda) = \lambda x_1 + (1 - \lambda) x_2, \ \forall \ \lambda \in (0, 1).$$

Si tenemos $x(\lambda) \leq x_*$ y como $x_1 < x(\lambda)$ entonces

$$f(x(\lambda)) \leq f(x_1). \tag{3.9}$$

Ahora, si $x_* \leq x(\lambda)$ y como $x(\lambda) < x_2$ entonces

$$f(x(\lambda)) \leq f(x_2). \tag{3.10}$$

De (3.9) y (3.10) tenemos

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \max\{f(x_1), f(x_2)\}, \ \forall \ \lambda \in [0,1].$$

Lo cual demuestra que f es cuasi-convexa. ■

Note que si I es un intervalo compacto no vacío y f es semicontinua inferior alcanza su mínimo sobre I. Por tanto, bajo ciertas condiciones, cuasi-convexidad y unimodalidad de f son equivalentes. Una versión mas elaborada del anterior teorema se puede encontrar en Martos [33]. La idea de unimodalidad fue generalizada por Newman [36] y Wilde [50] para el caso n dimensional.

Definición 3.5 *Sea $f : C \subset R^n \to R$ definida en el conjunto convexo C. Se dice que f es unimodal en C si para todo intervalo cerrado $[x_1, x_2] \subset C$ en el cual f alcanza su mínimo en el punto x_* tenemos $f(x_*) \leq f(x) \leq f(x_1)$ donde $x = \lambda x_* + (1-\lambda) x_1$ para $0 \leq \lambda \leq 1$.*

Usando esta definición se puede mostrar que las funciones semicontinuas inferiores cuasi-convexas y unimodales son equivalentes.

Proposición 3.7 *La función $f : C \subset R^n \to R$ definida en el conjunto convexo C es cuasi-convexa si y solo si para todo $x_1, x_2 \in C$ la función de variable real F definida por*

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

es cuasi-convexa en el conjunto convexo

$$J = \{\lambda \in R : 0 \leq \lambda \leq 1\}.$$

Demostración. Sean $x_1, x_2 \in C$ y supongamos que f es cuasi-convexa.
Sean $\lambda_0, \lambda_1, \lambda_2$ tal que $0 \leq \lambda_1 < \lambda_0 < \lambda_2 \leq 1$ y definamos

$$y_i = \lambda_i x_1 + (1-\lambda_i) x_2, \ i = 0, 1, 2.$$

Entonces, existe $\alpha \in (0,1)$ tal que $y_0 = \alpha y_1 + (1-\alpha) y_2$. En efecto, como $\lambda_1 < \lambda_0 < \lambda_2$, existe $\alpha \in (0,1)$ tal que

$$\lambda_0 = \alpha \lambda_1 + (1-\alpha) \lambda_2.$$

De la definición se tiene

$$y_0 = \lambda_0 x_1 + (1-\lambda_0) x_2.$$

Reemplazando el valor de λ_0 obtenemos

$$y_0 = \alpha(\lambda_1 x_1 + (1-\lambda_1)x_2) + (1-\alpha)(\lambda_2 x_1 + (1-\lambda_2)x_2).$$

Usando de nuevo la definición se tiene que

$$y_0 = \alpha y_1 + (1-\alpha) y_2.$$

Luego, de la cuasi-convexidad de f se tiene que

$$\max\{f(y_1), f(y_2)\} \geq f(y_0).$$

Lo cual implica que

$$\max\{F(\lambda_1), F(\lambda_2)\} \geq F(\lambda_0).$$

Entonces, F es una función cuasi-convexa.
Reciprocamente, sean $x_1, x_2 \in C$ y supongamos que F es cuasi-convexa en J, entonces

$$\max\{F(0), F(1)\} \geq F(\lambda)$$

y

$$\max\{f(x_1), f(x_2)\} \geq f(\lambda x_1 + (1-\lambda)x_2)$$

Lo cual implica que f es cuasi-convexa. ∎

De esta manera f es una función cuasi-convexa en su dominio convexo $C \subset R^n$ si y solo si es cuasi-convexa en todo segmento contenido en C.

Definición 3.6 *Una función $f : C \subset R^n \to R$ definida sobre un conjunto convexo C es cuasi-afin si y solo si f es cuasi-convexa y cuasi-cóncava.*

Como los conjuntos de nivel superior y los conjuntos de nivel inferior de las funciones cuasi-afines, son conjuntos convexos, se satisface, para todo x_1, x_2 en el dominio convexo de f

$$\max\{f(x_1), f(x_2)\} \geq f(\lambda x_1 + (1-\lambda)x_2) \geq \min\{f(x_1), f(x_2)\},\ \lambda \in [0,1].$$

Para funciones de variable real las funciones monótonas (funciones crecientes o decrecientes) y cuasi-afines son equivalentes.
Cabe resaltar que una función es convexa y cóncava si y solo si es afin. Las funciones cuasi-afin son generalizaciones de funciones afines en el caso en que reemplazamos convexidad y concavidad por cuasi-convexidad y cuasi-concavidad.
Un interesante resultado de funciones cuasi-afines en términos de superficie de nivel se da en las superficies de indiferencia usadas en el análisis económico.

Proposición 3.8 *Sea $f : C \subset R^n \to R$ una función continua definida en el conjunto convexo C. La superficie de nivel definida por*

$$Y(f,\alpha) = U(f,\alpha) \cap L(f,\alpha) = \{x \in C : f(x) = \alpha\}$$

es un conjunto convexo para todo $\alpha \in R$ si y solo si f es cuasi-afin.

Demostración. Asumamos que el conjunto $Y(f,\alpha)$ es convexo. Probaremos que f es cuasi-afin. Supongamos que f no es cuasi-convexa, entonces existen $\bar{x_1}, \bar{x_2} \in C$ y $\lambda \in (0,1)$ tal que

$$\bar{\alpha} = \max\{f(\bar{x_1}), f(\bar{x_2})\} < f(\lambda\bar{x_1} + (1-\lambda)\bar{x_2}). \tag{3.11}$$

Sin pérdida de generalidad asumiremos que $\bar{\alpha} = f(\bar{x_1})$, luego

$$f(\bar{x_2}) \leq f(\bar{x_1}) < f(x)$$

donde $x = \lambda\bar{x_1} + (1-\lambda)\bar{x_2}$. Desde que f es continua en el segmento compacto $[x, \bar{x_2}]$ existe

$$\bar{x} = \bar{\lambda}x + \left(1-\bar{\lambda}\right)\bar{x_2} \tag{3.12}$$

para algún $\bar{\lambda} \in [0,1)$ tal que $f(\bar{x}) = \bar{\alpha}$.
Sustituyendo el valor de x en (3.12) tenemos

$$\bar{x} = \bar{\lambda}\lambda\bar{x_1} + \left(1-\lambda\bar{\lambda}\right)\bar{x_2}.$$

Definiendo $\lambda\bar{\lambda} = r$ se tiene

$$\bar{x} = r\bar{x_1} + (1-r)\bar{x_2}, \text{ donde } r \in [0,1].$$

Desde que $x_2 = \frac{x-\lambda x_1}{1-\lambda}$ obtenemos

$$\bar{x} = rx_1 + (1-r)\frac{x-\lambda x_1}{1-\lambda}$$

multiplicamos por $(1-\lambda)$

$$(1-\lambda)\bar{x} = (r-\lambda)\bar{x_1} + (1-r)x$$

multiplicamos por $\frac{1}{1-r}$

$$\frac{1-\lambda}{1-r}\bar{x} - \frac{r-\lambda}{1-\lambda}\bar{x_1} = x$$

consideremos $\gamma = \frac{1-\lambda}{1-r}$, entonces

$$x = \gamma\bar{x} + (1-\gamma)x_1 \quad \gamma \in (0,1).$$

Como $Y(f,\bar{\alpha})$ es convexo y $\bar{x}, x_1 \in Y(f,\bar{\alpha})$ entonces $x \in Y(f,\bar{\alpha})$ luego $f(x) = \bar{\alpha}$ contradiciendo (3.11). Por lo tanto, f es cuasi-convexa.
Por otro lado, probaremos que f es cuasi-cóncava, es decir, $-f$ es cuasi-convexa.
Supongamos que $-f$ no es cuasi-convexa, entonces existen $\bar{x_1}, \bar{x_2} \in C$ y $\lambda \in (0,1)$ tal que

$$\bar{\alpha} = \max\{-f(\bar{x_1}), -f(\bar{x_2})\} < -f(\lambda\bar{x_1} + (1-\lambda)\bar{x_2}). \tag{3.13}$$

Sin pérdida de generalidad asumiremos que $\bar{\alpha} = -f(\bar{x_1})$ luego

$$-f(\bar{x_2}) \leq -f(\bar{x_1}) < -f(x)$$

donde $x = \lambda\bar{x_1} + (1-\lambda)\bar{x_2}$. Desde que f es continua en el segmento compacto $[x, \bar{x_2}]$ existe

$$\bar{x} = \bar{\lambda}x + (1-\bar{\lambda})\bar{x_2}$$

para algún $\bar{\lambda} \in [0,1)$ tal que $f(\bar{x}) = \bar{\alpha}$.
Analogamente, realizando operaciones elementales como en el caso anterior tenemos

$$x = \gamma\bar{x} + (1-\gamma)x_1,\ \gamma \in (0,1).$$

Como $Y(-f,\bar{\alpha})$ es convexo y $\bar{x}, x_1 \in Y(-f,\bar{\alpha})$ entonces $x \in Y(-f,\bar{\alpha})$ luego $-f(x) = \bar{\alpha}$ contradiciendo (3.13). Por lo tanto, f es cuasi-cóncava.
Reciprocamente, si f es cuasi-afin entonces es cuasi-convexa y cuasi-cóncava. Por tanto, $U(f,\alpha)$ y $L(f,\alpha)$ son conjuntos convexos para todo $\alpha \in R$.
En consecuencia, la intersección de $U(f,\alpha)$ y $L(f,\alpha)$ también es convexa. ■

Concluimos esta sección con dos observaciones dadas por Ponstein [40] y Thompson y Parke [48]. Primero, para funciones continuas la definición de cuasi-convexidad se presenta de la manera siguiente: Una función f continua definida en el conjunto convexo C es cuasi-convexa si sus conjuntos de nivel inferior estricto

$$L^0(f,\alpha) = \{x \in C : f(x) < \alpha\}$$

son conjuntos convexos para todo real α.
La segunda observación se refiere a la desigualdad de Frenchel, que es equivalente a la siguiente expresión

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

Entonces, la caracterización de funciones cuasi-convexas se puede enunciar de la siguiente forma. Sea f una función continua definida en un conjunto convexo $C \subset R^n$ es cuasi-convexa si y solo si

$$\forall x_1, x_2 \in C\ f(x_1) < f(x_2) \Rightarrow f(\lambda x_1 + (1-\lambda)x_2) \leq f(x_2),\ \forall\ \lambda \in [0,1].$$

3.1 Funciones Diferenciables Cuasi-convexas

En esta sección trabajaremos con la caracterización de funciones cuasi-convexas diferenciables.

Teorema 3.3 *(Arrow y Enthoven, 1961) Sea $f : C \subset R^n \to R$ una función diferenciable en el conjunto abierto convexo C. Entonces, f es una función cuasi-convexa si y solo si para todo $x_1, x_2 \in C$*

$$f(x_1) \leq f(x_2) \textit{ implica que } (x_1 - x_2)^T \nabla f(x_2) \leq 0 \tag{3.14}$$

Demostración. Supongamos que f es cuasi-convexa, si

$$f(x_1) \leq f(x_2)$$

entonces

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{f(x_1), f(x_2)\} = f(x_2), \ \forall \lambda \in [0,1].$$

En particular, para $\lambda \in (0,1)$ tenemos

$$\frac{f(\lambda x_1 + (1-\lambda)x_2) - f(x_2)}{\lambda} \leq 0$$

Lo cual es equivalente a

$$\frac{f(x_2 + \lambda(x_1 - x_2)) - f(x_2)}{\lambda} \leq 0$$

considerando $v = x_1 - x_2$, tenemos

$$\frac{f(x_2 + \lambda v) - f(x_2)}{\lambda} \leq 0$$

Tomando límite cuando $\lambda \to 0^+$ entonces

$$\frac{\partial f}{\partial v}(x_2) \leq 0$$

Por definición,

$$\frac{\partial f}{\partial v}(x_2) = \langle \nabla f(x_2), v \rangle \leq 0$$

Por lo tanto,

$$(x_1 - x_2)^T \nabla f(x_2) \leq 0$$

Reciprocamente, sean $x_1, x_2 \in C$ sin pérdida de generalidad asumimos que

$$f(x_1) \leq f(x_2)$$

definiendo,

$$F(\lambda) = f(\lambda x_1 + (1-\lambda)x_2), \ \forall \ \lambda \in [0,1]$$

entonces

$$F(1) \leq F(0).$$

Supongamos que la función F no es cuasi-convexa, es decir existe $\lambda_* \in (0,1)$ tal que

$$F(0) < F(\lambda_*), \ \lambda_* \in [0,1].$$

Desde que $F(1) \leq F(0)$, se tiene que

$$F(1) < F(\lambda_*). \tag{3.15}$$

Por el teorema del valor medio en el intervalo $[\lambda_*, 1]$, existe $\lambda_0 \in (\lambda_*, 1)$ tal que

$$F'(\lambda_0) = \frac{F(1) - F(\lambda_*)}{1 - \lambda_*} < 0$$

Entonces,

$$F'(\lambda_0) = (x_1 - x_2)^T \nabla f(x_0) < 0 \tag{3.16}$$

donde $x_0 = \lambda_0 x_1 + (1 - \lambda_0) x_2$. De (3.15) se tiene que $f(x_2) < f(x_0)$ luego de la hipótesis

$$(x_2 - x_0)^T \nabla f(x_0) \leq 0 \tag{3.17}$$

Reemplazando

$$x_0 - x_2 = \lambda_0 (x_1 - x_2),$$

en (3.17) tenemos

$$\lambda_0 (x_2 - x_1)^T \nabla f(x_0) \leq 0$$

Lo cual es equivalente a

$$\lambda_0 (x_1 - x_2)^T \nabla f(x_0) \geq 0,$$

desde que $\lambda_0 > 0$ tenemos

$$(x_1 - x_2)^T \nabla f(x_0) \geq 0.$$

Lo cual contradice (3.16). Por lo tanto, queda demostrado que f es una función cuasi-convexa. ∎

La condición (3.14) tiene una simple interpretación geométrica cuando $\nabla f(x_0) \neq 0$. El vector $\nabla f(x_0)$ define un hiperplano para el conjunto de nivel

$$\{y : f(y) \leq f(x_0)\}$$

en el punto x_0. Notemos, que la caracterización de funciones convexas diferenciables

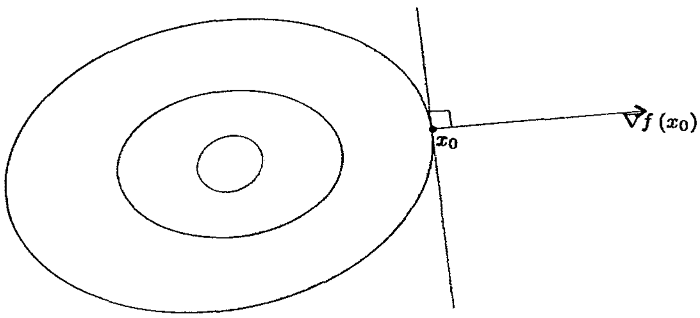

Figura 3.8: Conjuntos de nivel de una función cuasi-convexa

implica la caracterización de funciones diferenciables cuasi-convexas. Sin embargo, existen algunas diferencias inportantes, si f es convexa y $\nabla f(x_0) = 0$, entonces x_0 es un mínimo global de f, pero, este enunciado no se cumple para funciones cuasi-convexas, es posible que $\nabla f(x_0) = 0$, pero x_0 no es un mínimo global de f.

Observación 3.4 *La función $f : R \to R$ definida como $f(x) = x^3$ es una función cuasi-convexa. Su derivada está dada por $f'(x) = 3x^2$ en el punto $x_0 = 0$ tenemos que $f'(x_0) = 0$. Sin embargo, $x_0 = 0$ no es un mínimo global de f.*

Proposición 3.9 *Sea $f : C \subset R^n \to R$ diferenciable en el conjunto abierto y convexo C. Entonces f es cuasi-afin si y solo si para todo $x_1, x_2 \in C$*

$$f(x_2) \leq f(x_1) \text{ implica que } 0 \leq (x_1 - x_2)^T \nabla f(x_2)$$

y

$$f(x_1) \leq f(x_2) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) \leq 0$$

Demostración. Desde que f es cuasi-afin entonces es cuasi-cóncava y cuasi-convexa. Entonces f es cuasi-convexa si y solo si

$$f(x_2) \leq f(x_1) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) \leq 0$$

Además f es cuasi-cóncava si y solo si

$$f(x_2) \leq f(x_1) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) \geq 0$$

Por lo tanto, queda demostrado. ∎

Como consecuencia de esta proposición tenemos el siguiente corolario.

Corolario 3.1 *(Martos, 1965) Sea $f : C \subset R^n \to R$ una función diferenciable en el conjunto convexo y abierto C. Entonces f es cuasi-afin si y solo si para todo $x_1, x_2 \in C$ y $x = \lambda x_1 + (1 - \lambda) x_2$, con $0 \leq \lambda \leq 1$ tal que*

$$f(x_2) \leq f(x_1) \text{ implica que } 0 \leq (x_1 - x_2)^T \nabla f(x)$$

Demostración. Supongamos que f es afin, si $f(x_1) \geq f(x_2)$ entonces de la cuasi-convexidad de f se tiene que

$$f(x) = f(\lambda x_1 + (1 - \lambda) x_2) \leq \max\{f(x_1), f(x_2)\} = f(x_1), \ \forall \ \lambda \in [0, 1].$$

Desde que f es cuasi-cóncava y del teorema 3.3 obtenemos

$$(x_1 - x)^T \nabla f(x) \geq 0,$$

lo cual es equivalente a

$$(1 - \lambda)(x_1 - x_2)^T \nabla f(x) \geq 0, \ \forall \ \lambda \in [0, 1].$$

Por tanto,

$$(x_1 - x_2)^T \nabla f(x) \geq 0.$$

Reciprocamente, sean $x_1, x_2 \in C$ y $x = \lambda x_1 + (1 - \lambda) x_2$ tal que $f(x_1) \geq f(x_2)$ implica que

$$(x_1 - x_2)^T \nabla f(x) \geq 0.$$

Desde que la propiedad anterior se cumple para todo $\lambda \in [0,1]$, tomando en particular, a $\lambda = 0$ obtenemos

$$(x_1 - x_2)^T \nabla f(x_2) \geq 0$$

entonces del teorema 3.3 se tiene que f es cuasi-cóncava.
Por otro lado, considerando $\lambda = 1$ tenemos

$$(x_1 - x_2)^T \nabla f(x_1) \geq 0$$

lo cual es equivalente a

$$(x_2 - x_1)^T \nabla f(x_1) \leq 0$$

por tanto, del teorema 3.3 f es cuasi-convexa. Así, f es cuasi-afin. ∎

Otra caracterización de funciones diferenciables cuasi-convexas puede darse en términos del máximo local semi-estricto. Iniciaremos, considerando el caso de funciones de una variable.

Proposición 3.10 *(Diewert, Avriel y Zang,* 1981*) Sea $f : I \subset R \to R$ una función diferenciable definida en el intervalo abierto I. Entonces f es cuasi-convexa si y solo si para todo $x_0 \in I$ tal que $f'(x_0) = 0$ la función f no alcanza un máximo local semi-estricto en x_0.*

Demostración. Sea f cuasi-convexa y $x_0 \in I$ tal que $f'(x_0) = 0$.Supongamos qye f tiene un máximo local semi-estricto en x_0 entonces existen $x_1, x_2 \in I$ con $x_1 < x_0 < x_2$ tal que

$$f(\lambda x_1 + (1-\lambda) x_2) \leq f(x_0), \ \forall \lambda \in [0,1]$$

y

$$\max\{f(x_1), f(x_2)\} < f(x_0). \tag{3.18}$$

Desde que f es cuasi-convexa tenemos

$$f(x_0) \leq \max\{f(x_1), f(x_2)\} \tag{3.19}$$

luego de (3.18) y (3.19)

$$\max\{f(x_1), f(x_2)\} < \max\{f(x_1), f(x_2)\}.$$

Lo cual es una contradicción, por tanto, f no alcanza un máximo local semi-estricto en $x_0 \in I$.
Reciprocamente, supongamos que f no es cuasi-convexa, entonces existen $x_1, x_2 \in I$ y $\bar{\lambda} \in (0,1)$ tal que

$$\max\{f(x_1), f(x_2)\} < f\left(\bar{\lambda} x_1 + \left(1 - \bar{\lambda}\right) x_2\right).$$

Desde que f no alcanza un máximo local semi-estricto en x_0, para $x_1 < x_0 < x_2$ tenemos

$$f(x_0) \leq \max\{f(x_1), f(x_2)\}.$$

Sin pérdida de generalidad supongamos

$$f(x_2) \leq f(x_1),$$

luego,

$$f(x_0) \leq f(x_1) < f(\bar{x}) \tag{3.20}$$

donde $\bar{x} = \bar{\lambda}x_1 + \left(1 - \bar{\lambda}\right)x_2$. Como $f'(x_0) = 0$ tenemos

$$\forall\, \epsilon > 0, \exists\, \delta > 0 \;/\; 0 < |x - x_0| < \delta \Rightarrow \left|\frac{f(x) - f(x_0)}{x - x_0}\right| < \epsilon$$

de aquí para $0 < |x - x_0| < \delta$ se satisface

$$0 < |f(x) - f(x_0)| < \epsilon,\ \forall\, \epsilon > 0.$$

Sea $\epsilon = f(\bar{x}) - f(x_1) > 0$, talque $|\bar{x} - x_0| < \delta$ entonces

$$f(\bar{x}) - f(x_0) < f(\bar{x}) - f(x_1)$$

por tanto,

$$f(x_1) < f(x_0).$$

Lo cual contradice (3.20), en consecuencia f es cuasi-convexa. ∎

Podemos generalizar la proposición 3.10 para funciones de n variables.

Teorema 3.4 *(Diewert, Avriel y Zang, 1981) Sea $f : C \subset R^n \to R$ una función diferenciable definida en el conjunto abierto y convexo C. Entonces f es cuasi-convexa si y solo si para todo $x_0 \in C$ y todo $v \in R^n$ tal que $v^T v = 1$ y $v^T \nabla f(x_0) = 0$ la función $F(t) = f(x_0 + tv)$ no alcanza un máximo local semi-estricto en $t = 0$.*

Demostración. Supongamos que f es cuasi-convexa en un conjunto C abierto y convexo. En forma analoga a la prueba del teorema 3.2 se puede probar que $F(t) = f(x_0 + tv)$ es cuasi-convexa en $[0, \bar{t}]$. Como

$$F'(0) = v^t \nabla f(x_0)$$

usando la Proposición (3.10) para la función $F(t) = f(x_0 + tv)$ tenemos que F no alcanza un máximo local semi-estricto en $t = 0$.
Reciprocamente, supongamos que para todo $x_0 \in C$ y todo $v \in R^n$ tal que $v^t v = 1$ y $v^t \nabla f(x_{0)} = 0$ se tiene que

$$F(t) = f(x_0 + tv)$$

no alcanza máximo local semi-estricto en $t = 0$ y con $F'(0) = v^t \nabla f(x_0) = 0$ entonces de la proposición 3.10 se tiene que F es cuasi-convexa en $t = [0, \bar{t}]$. Luego,

$$F((1 - \lambda)\bar{t}) = F(t) \leq \max\{F(0), F(\bar{t})\},\ \forall\, \lambda \in [0,1]. \tag{3.21}$$

lo que es equivalente a

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \max\{f(x_1), f(x_2)\}, \forall \lambda\ \in [0,1]$$

Por lo tanto, f es cuasi-convexa. ■

La interpretación geométrica de la condición en el teorema anterior es que la función f restricta para una recta pasando por x_0 en dirección de v, ortogonal a la gradiente de f en x_0, no alcanza un máximo local semi-estricto en x_0. En particular, si $\nabla f(x_0) = 0$ en algun punto $x_0 \in C$, entonces x_0 no es un mínimo local semi-estricto de f a lo largo de toda recta pasando por x_0.

Proposición 3.11 *(Arrow y Enthoven, 1961; Avriel, 1972) Sea una función $f : C \subset R^n \to R$ cuasi-convexa y dos veces continuamente diferenciable en el conjunto abierto y convexo C. Si $x_0 \in C$, $v \in R^n$ y $v^T \nabla f(x_0) = 0$ entonces $0 \leq v^T \nabla^2 f(x_0) v$.*

Demostración. Supongamos que f es cuasi-convexa y existe $v \neq 0$ tal que $v^T \nabla f(x_0) = 0$ y $v^T \nabla^2 f(x_0) v < 0$.
Por continuidad de f, existe $\epsilon > 0$ tal que $\|\bar{x} - x_0\| < \epsilon$ implica

$$v^T \nabla^2 f(\lambda \bar{x} + (1 - \lambda) x_0) v < 0, \ \forall \lambda \in [0,1] \tag{3.22}$$

Sea $\bar{x} = x_0 + \mu v$ con $0 < \mu < \epsilon$. Entonces reemplazando obtenemos

$$\left(\frac{\bar{x} - x_0}{\mu}\right)^T \nabla^2 f(\lambda \bar{x} + (1 - \lambda) x_0) \left(\frac{\bar{x} - x_0}{\mu}\right) < 0$$

de ahí,

$$(\bar{x} - x_0)^T \nabla^2 f(\lambda \bar{x} + (1 - \lambda) x_0)(\bar{x} - x_0) < 0.$$

Por el teorema de Taylor,

$$f(\bar{x}) = f(x_0) + (\bar{x} - x_0)^T \nabla f(x_0) + \frac{1}{2}(\bar{x} - x_0)^T \nabla^2 f\left(\bar{\lambda}\bar{x} + \left(1 - \bar{\lambda}\right) x_0\right)(\bar{x} - x_0)$$

para algún $0 < \lambda < 1$. Desde que $v^T \nabla f(x_0) = 0$ y de (3.22), tenemos

$$f(\bar{x}) < f(x_0). \tag{3.23}$$

Por otro lado, consideremos $\hat{x} = x_0 + \mu(-v)$

$$-v = \frac{\bar{x} - x_0}{\mu}$$

lo cual es equivalente a,

$$v = \frac{x_0 - \hat{x}}{\mu}$$

que implica,

$$\left(\frac{x_0 - \hat{x}}{\mu}\right)^T \nabla^2 f(\lambda \hat{x} + (1 - \lambda) x_0) \left(\frac{x_0 - \hat{x}}{\mu}\right) < 0$$

luego,

$$\left(\frac{\hat{x}-x_0}{\mu}\right)^T \nabla^2 f\left(\hat{x}+(1-\lambda)\,x_0\right)\left(\frac{\hat{x}-x_0}{\mu}\right)<0$$

de ahí,

$$(\hat{x}-x_0)^T\,\nabla^2 f\left(\hat{x}+(1-\lambda)\,x_0\right)(\hat{x}-x_0)<0.$$

Por el teorema de Taylor,

$$f(\hat{x})=f(x_0)+(\hat{x}-x_0)^T\,\nabla f(x_0)+\frac{1}{2}(\hat{x}-x_0)^T\,\nabla^2 f\left(\hat{\lambda}\hat{x}+\left(1-\bar{\lambda}\right)x_0\right)(\hat{x}-x_0)$$

entonces,

$$f(\hat{x})<f(x_0)\,. \tag{3.24}$$

Como $x_0=\frac{1}{2}(\bar{x}+\hat{x})$ y de (3.23) y (3.24) se tiene que

$$f(x_0)\nleq \max\{f(\bar{x}),f(\hat{x})\}$$

Lo que contradice que f es cuasi-convexa. Así, se tiene que $v^T\nabla^2 f(x_0)\,v\geq 0.$ ∎

Corolario 3.2 *Sea $f:C\subset R^n\rightarrow R$ una función cuasi-convexa dos veces continuamente diferenciable en el conjunto abierto C. Si $x_0\in C$ y $\nabla f(x_0)=0$, entonces $0\leq z^T\nabla^2 f(x_0)\,z$ para todo $z\in R^n$.*

Demostración. Si $x_0\in C$ y $\nabla f(x_0)=0$ entonces

$$z^T\nabla f(x_0)=0,\ \forall z\in R^n,$$

usando el resultado de la proposición (3.11) obtenemos

$$z^T\nabla^2 f(x_0)\,z\geq 0,\ \forall z\in R^n.$$

Finalmente, queda demostrado el corolario. ∎

Cabe destacar, que la condiciones necesarias y suficientes para cuasi-convexidad en términos de segunda derivada son mas complicadas que para convexidad. El primer resultado en esta dirección se presenta en el siguiente corolario.

Corolario 3.3 *(Gerencsér, 1973) Sea $f:C\subset R^n\rightarrow R$ una función cuasi-convexa dos veces continuamente diferenciable en el conjunto abierto y convexo C. Entonces $\nabla^2 f$, la Hesiana de f, tiene a lo más un autovalor negativo para todo $x\in C$.*

Demostración. Supongamos que $x_0\in C$ y $\nabla^2 f(x^0)$ tiene dos o más autovalores negativos.
Sean $e_1,\dots,e_n$ un conjunto de n autovectores ortogonales de $\nabla f(x^0)$ y sin pérdida de generalidad, asumimos que los autovalores, λ_1 y λ_2, asociados con e_1 y e_2 son

negativos.
Consideremos

$$T=\left\{b_1e_1+...+b_ne_n:\ (e_i)^T\nabla f(x_0)=0;\ b_i\in R,\ i=1,...,n\right\}$$

como el subespacio generado por el conjunto de vectores ortogonales a $\nabla f(x_0)$ y sea

$$E=\{a_1e_1+a_2e_2:\ a_1,a_2\in R\}$$

el subespacio generado por e_1 y e_2.
Entonces $T\cap E$ debe contener al menos un vector $v_0\neq 0$ dado por $v_0=\alpha_1e_1+\alpha_2e_2$, donde α_1,α_2 no son ceros al mismo tiempo. Claramente,

$$(\alpha_1e_1+\alpha_2e_2)^T\nabla f(x_0)=\alpha_1(e_1)^T\nabla f(x_0)+\alpha_2(e_2)^T\nabla f(x_0)=0$$

luego, $(v_0)^T\nabla f(x_0)=0$. Por tanto, como λ_1 y λ_2 son autovalores asociados con e_1 y e_2

$$(v_0)^T\nabla^2 f(x_0)v_0=\alpha_1^2(e_1)^T\lambda_1e_1+\alpha_2\alpha_1(e_2)^T\lambda_1e_1+\alpha_1\alpha_2(e_1)^T\lambda_2e_2+\alpha_2^2(e_2)^T\lambda_2e_2$$

desde que e_1 y e_2 son autovectores ortogonales tenemos,

$$(v_0)^T\nabla^2 f(x_0)v_0=\lambda_1\alpha_1^2\|e_1\|^2+\lambda_2\alpha_2^2\|e_2\|^2$$

como λ_1,λ_2 son autovalores negativos

$$(v_0)^T\nabla^2 f(x_0)v_0<0$$

concluimos que la forma cuadrática es negativa, lo cual contradice la proposición 3.11. ■

Recordemos que la Hesiana de una función convexa no puede tener autovalores negativos. Del corolario anterior, observamos que la generalización de convexidad para cuasi-convexidad, es equivalente a la existencia de a lo mas un autovalor negativo en la Hesiana.
A continuación mencionaremos el concepto de matrices subdefinidas positivas, introducida por Martos [32].

Definición 3.7 *La matriz simétrica B de números reales se dice subdefinida positiva si*

$$y^TBy<0\Rightarrow By\leq 0\ \vee\ By\geq 0,$$

para todo $y\in R^n$. Si B es subdefinida positiva, entonces $-B$ es llamada subdefinida negativa. En analogía, a esta definición B se dice semidefinida positiva si

$$y^TBy\leq 0\Rightarrow By=0$$

Se puede probar que B es semidefinida negativa, entonces B es subdefinida negativa y la familia de matrices subdefinidas negativa es precisamente la familia de matrices simétricas $n\times n$ que pueden tener hasta un autovalor positivo.
En el siguiente ejemplo mostraremos que la condición necesaria de la proposición anterior no es suficiente para garantizar cuasi-convexidad.

Ejemplo 3.1 *Considere la función* $f(x) = -x^4$ *en*

$$C = \{x : x \in R, -1 < x < 1\}.$$

Entonces

$$\nabla f(x) = f'(x) = -4x^3$$

y el único punto $x_0 \in C$ para $v \neq 0$ tal que $v^T \nabla f(x_0) = 0$ es $x_0 = 0$. En este punto $v^T \nabla^2 f(x_0) v = 0$ para todo $v \in R$. Por tanto, se satisface la condición de la proposición, pero f no es cuasi-convexa.
Supongamos que f es cuasi-convexa entonces se cumple

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \max\{f(x_1), f(x_2)\}, \ \forall \ \lambda \in [0,1]$$

consideremos $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{2}$ y $\lambda_0 = \frac{1}{2}$ entonces

$$f\left(-\frac{1}{4} + \frac{1}{4}\right) \leq \max\left\{f\left(-\frac{1}{2}\right), f\left(\frac{1}{2}\right)\right\}$$

de ahí,

$$f(0) \leq \max\left\{-\frac{1}{16}, -\frac{1}{16}\right\}$$

luego,

$$0 \leq -\frac{1}{16}$$

lo cual es una contradicción. Por lo tanto, f no es cuasi-convexa.
Notemos que si aumentamos condiciones adicionales a la proposición 3.11 son suficientes para garantizar cuasi-convexidad. El primer resultado en esta dirección fue realizado por Katzner [27], donde se demostró que la función dos veces continuamente diferenciable definida en un conjunto abierto $C \subset R^n$ es cuasi-convexa si

i La función f tiene gradiente cuyas componentes son positivas.

ii Se satisface la siguiente condición

$$x_0 \in C, \ \ v \in R^n, v^T \nabla f(x_0) = 0 \Rightarrow v^T \nabla^2 f(x_0) v \geq 0. \tag{3.25}$$

Este resultado ha sido estudiado por Diewert, Avriel y Zang [18], donde (i) fue reemplazado por la condición que el gradiente no debe anularse en C. Un resultado más fuerte está dado por el siguiente teorema.

Teorema 3.5 *(Crouzeix, 1980) Sea* $f : C \subset R^n \rightarrow R$ *una función dos veces diferenciable en el conjunto abierto y convexo* C. *Supongamos que se satisfacen (3.25) y*

$$x_0 \in C, \ \nabla f(x_0) = 0 \ \Rightarrow \nabla^2 f(x_0) \ es \ definida \ positiva$$

Entonces f *es cuasi-convexa.*

Demostración. Supongamos que f no es cuasi-convexa entonces, existen $x_1, x_2 \in C$; $x_1 \neq x_2$ y existe $\lambda \in (0,1)$ tal que

$$\max\{f(x_1), f(x_2)\} < f(x_1 + \lambda(x_2 - x_1)), \ \lambda \in [0,1].$$

Por otro lado, notemos que

$$F(\lambda) = f(x_1 + \lambda(x_2 - x_1))$$

es continua sobre el conjunto compacto $[0,1]$, usando el Teorema de Weierstrass concluimos que F alcanza su máximo en $[0,1]$, entonces

$$\max\{f(x_1), f(x_2)\} \leq \max_{\lambda\in[0,1]} \{f(x_1 + \lambda(x_2 - x_1))\} = M \tag{3.26}$$

Sea $\bar{x} \in (x_1, x_2)$ un máximo global de f en $[x_1, x_2]$, definida por $\bar{x} = x_1 + \bar{\lambda}(x_2 - x_1)$, para $\bar{\lambda} \in (0,1)$. Además, sea

$$h(\lambda) = f(\bar{x} + \lambda(x_2 - x_1))$$

notemos que,

$$h(\lambda) = f\left(x_1 + \left(\lambda + \bar{\lambda}\right)(x_2 - x_1)\right)$$

de modo que

$$\lambda + \bar{\lambda} \in [0,1],$$

lo cual implica que $\lambda \in \left[-\bar{\lambda}, 1 - \bar{\lambda}\right]$. Desde que f alcanza su máximo global en $\bar{x}$, $f(\bar{x}) = M$, es decir, $h(0) = M$. Por tanto, $h(\lambda)$ es una función continua en λ tal que

$$\begin{aligned} h(\lambda) < h(0) &= M, \ \forall\, \lambda \in \left[-\bar{\lambda}, 0\right) &(3.27)\\ h(\lambda) \leq h(0) &= M, \ \forall\, \lambda \in \left[0, 1 - \bar{\lambda}\right] &(3.28) \end{aligned}$$

Entonces,

$$h'(0) = (x_2 - x_1)^T \nabla f(\bar{x}) = 0 \tag{3.29}$$

Consideremos dos casos:
Primer Caso: Si $\nabla f(\bar{x}) = 0$ entonces por hipótesis,

$$h''(0) = (x_2 - x_1)^T \nabla^2 f(\bar{x})(x_2 - x_1) > 0 \tag{3.30}$$

Aplicando la expansión de Taylor de segundo orden para $h(\lambda)$ alrededor de $\lambda = 0$ tenemos

$$h(\lambda) = h(0) + \lambda h'(0) + \frac{1}{2}\lambda^2 h''(0) + \lambda^2 \alpha(\lambda) \tag{3.31}$$

donde $\alpha(\lambda) \to 0$ cuando $\lambda \to 0$. De (3.30) tenemos que

$$\lambda^2 \left[\frac{1}{2}h''(0) + \alpha(\lambda)\right] > 0$$

para λ suficientemente pequeño. Usando (3.29) obtenemos

$$h(0) < h(\lambda)$$

contradiciendo (3.27) y (3.28).
Segundo Caso: Si $\nabla f(\bar{x}) \neq 0$. Sin pérdida de generalidad supongamos que $\bar{x} = 0$, y que

$$f(\bar{x}) = h(0) = 0. \tag{3.32}$$

Además, consideremos

$$\nabla f(\bar{x}) = -(0, ..., 1)^T. \tag{3.33}$$

Sea $\hat{x} \in R^{n-1}$, representada por el vector

$$\hat{x} = \left(x^1, ..., x^{n-1}\right)^T$$

de ahí $x \in R^n$ está dada por $x = (\hat{x}^t, x^n)^T$. Usando los resultados de (3.29) y (3.33) tenemos

$$-(x_2^n - x_1^n) = 0$$

de aquí,

$$(x_2 - x_1) = \left((\hat{x_2} - \hat{x_1})^T, 0\right)^T.$$

Considerando el Teorema de la Función Implícita para la función $f : C \subset R^n \to R$, donde $\left(\hat{0}, 0\right) \in C$ tal que se satisface (3.32) entonces, existe una vecindad convexa $U \subset R^{n-1}$ de $\hat{0}$ donde

$$f(\hat{x}, g(\hat{x})) = 0, \quad \forall\, \hat{x} \in U. \tag{3.34}$$

Por otro lado, de (3.33) concluimos que existe una función $g : U \subset R^{n-1} \to R$ dos veces diferenciable tal que para todo $\hat{x} \in U$ tenemos

$$\frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x})) < 0 \tag{3.35}$$

derivando (3.34) con respecto a $\hat{x}$ obtenemos

$$\nabla g(\hat{x}) = -\frac{1}{\frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x}))} \nabla_x f(\hat{x}, g(\hat{x})) \tag{3.36}$$

donde $\nabla_x f \in R^{n-1}$ es el vector de derivadas parciales de f respecto sus $n-1$ primeros argumentos, y $\frac{\partial f}{\partial x_n}$ es la derivada parcial de f con respecto su n-ésimo argumento, entonces

$$\nabla f(\hat{x}, g(\hat{x})) = \left(\nabla_x f(\hat{x}, g(\hat{x})), \frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x}))\right). \tag{3.37}$$

También, observemos que $\bar{x} = 0$ implica que $g\left(\hat{0}\right) = 0$ y que por (3.33) y (3.36) obtenemos $\nabla g\left(\hat{0}\right) = 0$. Además, la diferenciación de (3.36) con respecto a $\hat{x}$

$$\nabla^2 g(\hat{x}) = -\frac{1}{\frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x}))}\nabla^2_{\hat{x}\hat{x}} f(\hat{x}, g(\hat{x})) + \nabla g(\hat{x}) \nabla^2_{n\hat{x}} f(\hat{x}, g(\hat{x}))^T \quad (3.38)$$

$$+ \nabla^2_{\hat{x}n} f(\hat{x}, g(\hat{x})) \nabla g(\hat{x})^T + \nabla^2_{nn} f(\hat{x}, g(\hat{x})) \nabla g(\hat{x}) \nabla g(\hat{x})^T \quad (3.39)$$

para todo $\hat{x} \in U$, donde la segunda derivada en (3.38)

$$\nabla^2 f = \begin{vmatrix} \nabla^2_{\hat{x}\hat{x}} f & \nabla^2_{\hat{x}n} f \\ \nabla^2_{n\hat{x}} f & \nabla^2_{nn} f \end{vmatrix} \quad (3.40)$$

Consideremos el vector $\bar{z} \in R^n$, dado por

$$\bar{z} = \left(\hat{z}^T, \hat{z}^T \nabla g(\hat{x})\right)^T. \quad (3.41)$$

Del producto interno de (4.3) y (3.41) tenemos

$$\hat{z}^T \nabla_x f(\hat{x}, g(\hat{x})) + \hat{z}^T \frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x})) \nabla g(\hat{x}) \quad (3.42)$$

lo cual es equivalente a,

$$\hat{z}^T \left\{ \nabla_x f(\hat{x}, g(\hat{x})) + \frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x})) \nabla g(\hat{x}) \right\}. \quad (3.43)$$

De (3.36) y (3.43)

$$\bar{z} \nabla f(\hat{x}, g(\hat{x})) = 0$$

y por (3.25) obtenemos

$$\bar{z}^T \nabla^2 f(\hat{x}, g(\hat{x})) \bar{z} \geq 0 \quad (3.44)$$

Usando la partición (3.40) y el resultado de (3.38)

$$\hat{z}^T \nabla^2 g(\hat{x}) \hat{z} = -\frac{\bar{z}^T \nabla^2 f(\hat{x}, g(\hat{x})) \bar{z}}{\frac{\partial f}{\partial x_n}(\hat{x}, g(\hat{x}))} \geq 0 \quad (3.45)$$

para todo $\hat{x} \in U$ y $\hat{z} \in R^{n-1}$, donde la desigualdad en (3.45) sigue de (3.44). Por tanto, g es convexa en U.
Consideremos $\hat{u} = \hat{x_2} - \hat{x_1}$ entonces de (3.27) en términos de nuevas coordenadas

$$f(\lambda \hat{u}, 0) < f\left(\hat{0}, 0\right) = 0 \quad (3.46)$$

lo cual de (3.46) implica que $g(\lambda \hat{u}) \neq 0$ para todo $\lambda \in \left[-\bar{\lambda}, 0\right)$ tal que $\lambda \hat{u} \in U$. Además, por el teorema de Taylor

$$g(\lambda \hat{u}) = g\left(\hat{0}\right) + \lambda \hat{u} \nabla g\left(\hat{0}\right) + \frac{1}{2}\lambda^2 \hat{u}^T \nabla^2 g(\hat{v}) \hat{u} \quad (3.47)$$

donde $\hat{v} \in (0, \lambda\hat{u})$. Desde que g es convexa y los dos primeros términos en el lado derecho de (3.75) se anulan. Luego,

$$g(\lambda\hat{u}) \geq 0$$

en particular,

$$g(\lambda\hat{u}) > 0$$

para todo $\lambda \in [-\lambda, 0) \cap \{\lambda : \lambda\hat{u} \in U\}$.
Igualmente, la continuidad de ∇f y (3.35) implica la existencia de una vecindad convexa V de $\hat{0}$, $V \subset U$ y $\delta > 0$ tal que

$$(\hat{x}, x^n) \in V \times [-\delta, \delta] \Rightarrow \frac{\partial f}{\partial x_n}(\hat{x}, x^n) < 0 \tag{3.48}$$

La continuidad de g implica la existencia de $\hat{\lambda} \in [-\bar{\lambda}, 0)$ tal que $\hat{\lambda}\hat{u} \in V$ y $0 < g\left(\hat{\lambda}\hat{u}\right) < \delta$.
Finalmente, tenemos de (3.46), (3.34) y (3.48), respectivamente,

$$f\left(\hat{\lambda}\hat{u}, \hat{0}\right) < 0 \tag{3.49}$$

$$f\left(\hat{\lambda}\hat{u}, g\left(\hat{\lambda}\hat{u}\right)\right) = 0 \tag{3.50}$$

y

$$(0, ..., 1)^T \nabla f\left(\hat{\lambda}\hat{u}, \lambda\right) = \frac{\partial f}{\partial x_n}\left(\hat{\lambda}\hat{u}, x^n\right) < 0 \tag{3.51}$$

para todo $x_n \in \left[0, g\left(\hat{\lambda}\hat{u}\right)\right]$. Pero (3.51), implica que f es decreciente en la n-ésima coordenada de $x^n = 0$ a $x^n = \hat{\lambda}\hat{u}$ entonces

$$f\left(\hat{\lambda}\hat{u}, g\left(\hat{\lambda}\hat{u}\right)\right) < f\left(\hat{\lambda}\hat{u}, \hat{0}\right)$$

contradiciendo (3.49) y (3.50); de esta manera (3.26) no se satisface, y la prueba está completa. ∎

Del teorema anterior y la proposición 3.11 obtenemos el siguiente resultado probado por Diewert, Avriel y Zang [18].

Corolario 3.4 *Sea f una función dos veces continuamente diferenciable en el conjunto abierto y convexo $C \subset R^n$ y supongamos que $\nabla f(x) \neq 0$ para todo $x \in C$. Entonces, f es cuasi-convexa si y solo si*

$$x \in C,\ v \in R^n,\ v^T \nabla f(x) = 0 \ \Rightarrow\ v^T \nabla^2 f(x) v \geq 0.$$

Demostración. Supongamos que

$$\nabla f(x) \neq 0,\ \forall x \in C.$$

Sea v tal que $v^T \nabla f(x) = 0$. Por la cuasi-convexidad de f de la Proposición 3.11 se tiene que

$$v^T \nabla^2 f(x) v \geq 0.$$

Reciprocamente, supongamos que f no es cuasi-convexa tal que $\nabla f(x) \neq 0$, considerando el segundo caso del teorema (3.5) y procediendo analogamente llegamos a una contradicción. Lo cual implica que f es una función cuasi-convexa. ∎

Los siguientes dos resultados, aparecen en Diewert, Avriel y Zang [18] y Avriel y Schaible [7], los cuales complementan las proposiciones anteriores.

Proposición 3.12 *Sea $f : C \subset R \to R$ una función dos veces continuamente diferenciable definida en el conjunto abierto y convexo C. Entonces f es cuasi-convexa si y solo si para todo $x_0 \in C$ tal que $f'(x_0) = 0$ y $f''(x_0) > 0$ ó $f''(x_0) = 0$ y f no alcanza un máximo local semi-estricto en x_0.*

Demostración. Supongamos que f es cuasi-convexa tal que $f'(x_0) = 0$ por la proposición (3.11) tenemos $f''(x_0) \geq 0$.
Si $f''(x_0) \neq 0$ entonces $f''(x_0) > 0$. Además, por proposición (3.10) tenemos que f no alcanza un máximo local semi-estricto en x_0.
Reciprocamente, si $f'(x_0) = 0$ tal que $f''(x_0) > 0$, entonces por teorema (3.5) f es cuasi-convexa.
Si $f''(x_0) = 0$ tal que no es cierto $f''(x_0) > 0$, como f no alcanza un máximo local semi-estricto en x_0 y desde que $f'(x_0) = 0$ entonces por proposición (3.10), f es cuasi-convexa. ∎

Teorema 3.6 *Sea $f : C \subset R^n \to R$ una función dos veces diferenciable definida en el conjunto abierto y convexo C. Entonces f es cuasi-convexa si y solo si para todo $x_0 \in C$ y $v \in R^n$ tal que $v^T v = 1$ y $v^T \nabla f(x_0) = 0$ y $v^T \nabla^2 f(x_0) v > 0$ ó $v^T \nabla^2 f(x_0) v = 0$ y la función $F(t) = f(x_0 + tv)$ no alcanza una máximo local semi-estricto en $t = 0$.*

Demostración. Supongamos que f es cuasi-convexa en un conjunto C abierto y convexo. Si $v^T \nabla f(x_0) = 0$ entonces por proposición (3.11) tenemos

$$v^T \nabla^2 f(x_0) v \geq 0.$$

Si $v^T \nabla^2 f(x_0) v \neq 0$ entonces $v^T \nabla^2 f(x_0) v > 0$. Además, por proposición (3.10) f no alcanza un máximo local semi-estricto en x_0.
Reciprocamente, si $v^T \nabla f(x_0) = 0$ tal que $v^T \nabla^2 f(x_0) v > 0$ donde no se cumple $v^T \nabla^2 f(x_0) v = 0$ por teorema (3.5), f es cuasi-convexa.
Si $v^T \nabla^2 f(x_0) v = 0$ tal que $F'(0) = v^T f(x_0) = 0$ donde la función F no alcanza un máximo local semi-estricto en $t = 0$. Usando la proposición (3.10) obtenemos la cuasi-convexidad de F, es decir,

$$F(t) = f(x_0 + tv)$$

es cuasi-convexa en $t \in [0, \bar{t}]$. Consideremos

$$v = \frac{x_2 - x_0}{\|x_2 - x_0\|} \quad \wedge \quad \bar{t} = \|x_2 - x_0\| \tag{3.52}$$

Usando (3.52) y la cuasi-convexidad de F entonces

$$f(\lambda x_0 + (1-\lambda) x_2) \leq \max\{f(x_0), f(x_2)\}, \ \forall \ \lambda \in [0,1]$$

Por lo tanto, queda demostrada la cuasi-convexidad de f. ∎

Concluimos esta sección mencionando las condiciones necesarias y suficientes para funciones cuasi-convexas dos veces diferenciable las cuales estan dadas por "determinantes acotadas". Arrow y Enthoven [3] fueron los primeros en dar estas condiciones para funciones definidas en octante no negativo de R^n. Estas condiciones fueron extendidas por Ferland [21] para funciones definidas en conjuntos convexos más generales. Resultados relacionados fueron también realizados por Avriel y Schaible [7] y Crouzeix y Ferland [14].

El k-ésimo orden de la matriz Hessiana de una función f dos veces continuamente diferenciable en el punto $x \in R^n$ es definida como

$$D_k(x) = \begin{vmatrix} 0 & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \cdots & \frac{\partial f}{\partial x_k} \\ \frac{\partial f}{\partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_k} \\ \cdots & \cdots & & \cdots \\ \frac{\partial f}{\partial x_k} & \frac{\partial^2 f}{\partial x_k \partial x_1} & \frac{\partial^2 f}{\partial x_k \partial x_2} \cdots & \frac{\partial^2 f}{\partial x_k \partial x_k} \end{vmatrix}.$$

Una condición necesaria para que f sea cuasi-convexa en un conjunto sólido convexo $C \subset R^n$ es

$$(-1)^k \det D_k(x) \leq 0, \quad k = 1, ..., n$$

para todo $x \in C$, donde *det* denota la determinante. Una condición suficiente para una función $f : C \subset R^n \to R$ dos veces continuamente diferenciable en un conjunto convexo C sea cuasi-convexa en C es que

$$(-1)^k \det D_k(x) < 0, \quad k = 1, ..., n$$

para todo $x \in C$. Por otro lado, si la función $f : C \subset R^n \to R$ es cuasi-cóncava en el conjunto convexo C entonces, para todo $x \in C$ se satisface

$$(-1)^k \det D_k(x) \geq 0, \quad k = 1, ..., n.$$

Además, si para todo $x \in C$ tenemos

$$(-1)^k \det D_k(x) > 0$$

la función f es cuasi-cóncava.

Ejemplo 3.2 *Probar que*

$$f(x_1, x_2, x_3) = x_1 x_2 x_3$$

es cuasi-cóncava en R^3_{++}.

En efecto,

$$D_3(x) = \begin{vmatrix} 0 & x_2x_3 & x_1x_3 & x_1x_2 \\ x_2x_3 & 0 & x_3 & x_2 \\ x_1x_3 & x_3 & 0 & x_1 \\ x_1x_2 & x_2 & x_1 & 0 \end{vmatrix}$$

tenemos

$$\begin{aligned} |D_1| &= -x_2^2x_3^2 < 0 \\ |D_2| &= 2x_1x_2x_3^3 > 0 \\ |D_3| &= -x_1^2x_2^2x_3^2 < 0 \end{aligned}$$

entonces

$$\begin{aligned} &(-1)^1 |D_1| > 0 \\ &(-1)^2 |D_2| > 0 \\ &(-1)^3 |D_3| > 0 \end{aligned}$$

Por lo tanto, $f(x_1, x_2, x_3) = x_1x_2x_3$ es cuasi-cóncava en R^3_{++}.

3.2 Funciones Estrictamente Cuasi-convexas

Comenzamos esta sección con la caracterización de funciones cuasi-convexas dada en teorema 3.1. Sea f una función definida en el conjunto convexo $C \subset R^n$ es cuasi-convexa si para $x_1, x_2 \in C$

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{f(x_1), f(x_2)\}, \ \forall\lambda \in [0,1].$$

Considerando esta caracterización como definición, tenemos la siguiente definición.

Definición 3.8 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto convexo C. La función f se denomina estrictamente cuasi-convexa si*

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\}$$

para todo $x_1, x_2 \in C$, $x_1 \neq x_2$ y $\lambda \in (0,1)$.

Si f es estrictamente cuasi-convexa entonces $-f$ es estrictamente cuasi-cóncava.
Es importante resaltar que existe confusión en la literatura concerniente a la terminología de funciones convexas generalizadas, funciones estrictamente cuasi-convexas son llamadas fuertemente cuasi-convexas por Avriel [6]. Es claro, que toda función estrictamente cuasi-convexa es cuasi-convexa. Por esto, surge una interrogante muy importante:
ï¿½Cuál es la diferencia entre los dos tipos de funciones?
Una función que es cuasi-convexa, pero no estrictamente cuasi-convexa es constante

en algún intervalo de su dominio. Los conjuntos de nivel inferior de funciones estrictamente cuasi-convexas semicontinuas inferiores (continuas) son conjuntos estrictamente convexos en el interior de su dominio. Una función estrictamente cuasi-convexa alcanza su máximo en a no más que un punto.
De la misma manera que en la cuasi-convexidad, discutiremos las propiedades de funciones estrictamente cuasi-convexas comenzando para el caso de variable real.

Proposición 3.13 *Sea $f : I \subset R \to R$ una función definida en el intervalo abierto I. Si f es una función estrictamente cuasi-convexa, entonces no existe un punto $x_0 \in C$ que es un máximo local de f.*

Demostración. Si f es estrictamente cuasi-convexa, entonces para todo $x_1, x_2 \in C$, $x_1 \neq x_2$, y $0 < \lambda < 1$

$$f(\lambda x_1 + (1-\lambda) x_2) < \max\{f(x_1), f(x_2)\}$$

De esta manera se concluye, que para $x_1, x_2 \in C$ no existe $x_0 \in (x_1, x_2)$ que sea máximo local de f. ■

Proposición 3.14 *(Diewert, Avriel y Zang, 1981) Sea $f : C \subset R \to R$ una función que satisface la propiedad de segmento máximo en el conjunto convexo y abierto C. Si no existe un punto $x_0 \in C$ que es un máximo local de f, entonces f es una función semi-estricta cuasi-convexa.*

Demostración. Sean $x_1, x_2 \in C$, $x_1 < x_2$. Por hipótesis, consideramos que f alcanza la propiedad de segmento máximo en el intervalo cerrado $[x_1, x_2]$.
Si el conjunto de puntos maximizadores solo posee a los puntos extremos, entonces

$$f(\lambda x_1 + (1-\lambda) x_2) < \max\{f(x_1), f(x_2)\}, \ \forall\, \lambda \in (0,1).$$

Supongamos que el conjunto de puntos maximizadores contiene al punto x_0 tal que $x_1 < x_0 < x_2$, luego

$$f(\lambda x_1 + (1-\lambda) x_2) < f(x_0), \ \forall\, \lambda \in (0,1).$$

Pero esto contradice la hipótesis que no existe una máximo local, x_0 en C. Por tanto, nuestra suposición es falsa y f es una función estrictamente cuasi-convexa. ■

Similarmente, para el caso cuasi-convexo podemos concluir que para funciones de variable real teniendo la propiedad de segmento máximo, la no existencia de máximo local caracteriza a las funciones estrictamente cuasi-convexas.

Teorema 3.7 *(Diewert, Avriel y Zang, 1981) Una función $f : C \subset R^n \to R$ definida en el conjunto convexo y abierto C es estrictamente cuasi-convexa si y solo si*

1. *Tiene la propiedad de segmento máximo.*

2. *Para todo $x \in C$, todo $v \in R^n$ tal que $v^T v = 1$ y $\bar{t} \in R$, $\bar{t} > 0$ satisfaciendo $x + \bar{t}v \in C$ la función $F(t) = f(x + tv)$ no alcanza un máximo local en $t \in (0, \bar{t})$.*

Demostración. Primero, supongamos que f es estrictamente cuasi-convexa en un conjunto C abierto y convexo.
El item 1 se satisface, desde que f es estrictamente cuasi-convexa, sean $x_1, x_2 \in C$ tal que $x_1 \neq x_2$ entonces

$$f(\lambda x_1 + (1-\lambda) x_2) < \max\{f(x_1), f(x_2)\}, \ \forall \lambda \in (0,1).$$

De esta manera, el máximo de f sobre $[x_1, x_2]$ es alcanzado por x_1 o x_2. Por tanto, f tiene la propiedad de segmento máximo.
Para demostrar el item 2, afirmamos que $F(t) = f(x + tv)$ es estrictamente cuasi-convexa en $(0, \bar{t})$. Probaremos inicialmente

$$\forall\ t \in (0, \bar{t}) \Rightarrow x + tv \in C$$

En efecto, sea $0 < t < \bar{t}$ entonces,

$$0 < \frac{t}{\bar{t}} < 1$$

de aquí,

$$\begin{aligned} x + tv = x + \frac{t}{\bar{t}}(\bar{t}v) &= \left(\frac{t}{\bar{t}} + 1 - \frac{t}{\bar{t}}\right) x + \frac{t}{\bar{t}}(\bar{t}v) \\ &= \frac{t}{\bar{t}}(x + \bar{t}v) + \left(1 - \frac{t}{\bar{t}}\right) x \in C \end{aligned}$$

La última igualdad se cumple desde que C es un conjunto convexo.
Ahora probaremos que F es estrictamente cuasi-convexa en $(0, \bar{t})$. Por tanto, debemos demostrar que

$$F(\lambda t_1 + (1-\lambda) t_2) < \max\{F(t_1), F(t_2)\}, \ \forall\ t_1, t_2 \in (0, \bar{t}).$$

En efecto, sean $t_1, t_2 \in (0, \bar{t})$ luego

$$x + t_1 v, x + t_2 v \in C,$$

de la cuasi-convexidad estricta de f tenemos,

$$f(\lambda(x + t_1 v) + (1-\lambda)(x + t_2 v)) < \max\{f(x + t_1 v), f(x + t_2 v)\}, \ \forall\ \lambda \in (0,1).$$

La expresión anterior es equivalente a

$$\begin{aligned} F(\lambda t_1 + (1-\lambda) t_2) = f(x + (\lambda t_1 + (1-\lambda) t_2) v) &< \max\{f(x + t_1 v), f(x + t_2 v)\} \\ &= \max\{F(t_1), F(t_2)\} \end{aligned}$$

La cual demuestra que F es estrictamente cuasi-convexa en $(0, \bar{t})$.
Usando la Proposición 3.13, la función $F(t) = f(x + tv)$ no alcanza un máximo

local en $t \in (0, \bar{t})$.
Reciprocamente, sean $x_1, x_2 \in C$ tal que $x_1 \neq x_2$, consideremos

$$v = \frac{x_2 - x_1}{\|x_2 - x_1\|} \ \wedge \ \bar{t} = \|x_2 - x_1\| . \tag{3.53}$$

Como $x_1 + tv \in C$, definamos

$$F(t) = f(x_1 + tv), \ \forall t \in (0, \bar{t}).$$

Desde que f tiene la propiedad de segmento máximo sobre $[0, \bar{t}]$, F alcanza su máximo en algún $t_0 \in [0, \bar{t}]$ y como F no alcanza un máximo local en $t \in (0, \bar{t})$, entonces por proposición 3.14, F es estrictamente cuasi-convexa, es decir,

$$F((1-\lambda)\bar{t}) = F(t) < \max\{F(0), F(\bar{t})\}, \ \forall \ \lambda \ \in (0,1) \tag{3.54}$$

Usando (3.53) y reemplazando en (3.54) tenemos

$$f(\lambda x_1 + (1-\lambda) x_2) < \max\{f(x_1), f(x_2)\}, \ \forall \ \lambda \ \in (0,1).$$

Por lo tanto, f es estrictamente cuasi-convexa. ■

A continuación, analizaremos la relación entre cuasi-convexidad estricta y unimodalidad estricta, definida por Newman [36] y Wilde [50]. Por consiguiente, daremos la siguiente definición de unimodalidad estricta.

Definición 3.9 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto convexo C. Esta función es llamada estrictamente unimodal en C si para todo intervalo cerrado $[x_1, x_2] \subset C$ donde f alcanza su mínimo en el punto $x_* \neq x_1$ tenemos $f(x_*) < f(x) < f(x_1)$ donde $x = \lambda x_1 + (1-\lambda) x_*$.*

Proposición 3.15 *(Elkin, 1968) Sea $f : C \subset R^n \to R$ una función semicontinua inferior definida en el conjunto convexo C. Entonces f es estrictamente cuasi-convexa si y solo si f es estrictamente unimodal.*

Demostración. Supongamos que f es estrictamente cuasi-convexa, entonces f es cuasi-convexa. Sean $x_1, x_2 \in C$ y sea x^* un mínimo de f en el intervalo cerrado $[x_1, x_2]$, tal que $x^* \neq x_1$.
Desde que f es cuasi-convexa, f es unimodal y $x = \lambda x_1 + (1-\lambda) x^*$, $0 < \lambda < 1$ tenemos

$$f(x^*) \leq f(x) \leq f(x_1).$$

Si $f(x_1) = f(x)$ entonces f debe ser constante en $[x_1, x]$ ó $[x, x^*]$.
En efecto, supongamos que f no es constante en ningún intervalo, entonces existe $\bar{x} \in [x_1, x]$ tal que

$$f(\bar{x}) < \max\{f(x), f(x_1)\} = f(x). \tag{3.55}$$

De la misma manera, existe $\hat{x} \in [x, x^*]$

$$f(\hat{x}) < \max\{f(x), f(x_*)\} = f(x). \tag{3.56}$$

De (3.55) y (3.56) tenemos

$$\max\{f(\hat{x}), f(\bar{x})\} < f(x)$$

desde que $x \in [\bar{x}, \hat{x}]$, se concluye que f no es cuasi-convexa.
Pero, por definición una función estrictamente cuasi-convexa no puede ser constante en todo intervalo, entonces $f(x) < f(x_1)$.
Si $f(x) = f(x^*)$ entonces por cuasi-convexidad estricta de f para todo

$$x_0 = \lambda_0 x + (1 - \lambda_0) x^*, \quad \lambda_0 \in (0, 1)$$

entonces

$$f(x_0) < \max\{f(x^*), f(x)\} = f(x^*)$$

lo cual es equivalente a

$$f(x_0) < f(x^*)$$

contradiciendo la hipótesis que x^* es un máximo en $[x_1, x_2]$.
Reciprocamente, supongamos que f es estrictamente unimodal, entonces es unimodal, por tanto f es cuasi-convexa.
Consideremos $x_1, x_2 \in C$, $x_1 \neq x_2$ tal que

$$x_0 = \lambda x_1 + (1 - \lambda) x_2, \quad \forall \lambda \in (0, 1).$$

Por la cuasi-convexidad de f tenemos

$$f(x_0) \leq \max\{f(x_1) f(x_2)\}$$

Supongamos que

$$f(x_0) = \max\{f(x_1), f(x_2)\} = f(x_1)$$

entonces f debe ser constante en $[x_1, x_0]$ o $[x_0, x_2]$, contradiciendo la unimodalidad estricta de f. Por tanto,

$$f(\lambda x_1 + (1 - \lambda) x_2) < \max\{f(x_1), f(x_2)\}, \ \forall \lambda \in (0, 1)$$

luego, f es estrictamente cuasi-convexa. ■

Centraremos nuestra atención en los conjuntos de nivel de las funciones estrictamente cuasi-convexas. Necesitamos caracterizar este tipo de funciones por medio de sus conjuntos de nivel, así, como en el caso cuasi-convexo. Desafortunadamente, si los conjuntos de nivel inferior son estrictamente convexos no es condición suficiente para garantizar cuasi-convexidad estricta. Presentamos la siguiente definición.

Definición 3.10 *Un conjunto convexo $C \subset R^n$ se dice estrictamente convexo si para dos puntos x_1, x_2 en su frontera, todo punto x está dado por $x = \lambda x_1 + (1 - \lambda) x_2$*

Proposición 3.16 *(Elkin, 1968) Sea f una función continua, definida en R^n. Si f es estrictamente cuasi-convexa entonces sus conjuntos de nivel inferior son estrictamente convexos.*

Demostración. Si el conjunto de nivel inferior de f

$$L(f,\alpha) = \{x : x \in C, f(x) \leq \alpha\}$$

no es convexo implica que la función no es cuasi-convexa, de ahí la función no es estrictamente cuasi-convexa.
De esta manera, $L(f,\bar{\alpha})$ es convexo, sin pérdida de generalidad consideremos $\bar{\alpha} \in R$.
Sean $x_1, x_2 \in L(f,\bar{\alpha})$ tal que

$$f(x_1) = \bar{\alpha} \ \wedge \ f(x_2) = \bar{\alpha},$$

desde que f es estrictamente cuasi-convexa

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\} = \bar{\alpha}, \ \forall \lambda \in (0,1). \tag{3.57}$$

Supongamos que existe $0 < \lambda_* < 1$ tal que

$$x_* = \lambda_* x_1 + (1-\lambda_*)x_2,$$

no es un punto interior de $L(f,\bar{\alpha})$, entonces

$$f(x_*) \geq \bar{\alpha}$$

Por tanto,

$$f(\hat{x}) < \bar{\alpha} \leq f(x_*)$$

donde $\hat{x} = \hat{\lambda}x_1 + \left(1-\hat{\lambda}\right)x_2$. Desde que f es continua en R^n, existe

$$\bar{x} = \bar{\lambda}x_* + \left(1-\bar{\lambda}\right)\hat{x}, \ \bar{\lambda} \in (0,1]$$

tal que $f(\bar{x}) = \bar{\alpha}$.
Luego, de (3.57)

$$\bar{\alpha} = f(\bar{x}) < \bar{\alpha}.$$

Lo cual es una contradicción. En consecuencia, todo punto

$$x = \lambda x_1 + (1-\lambda)x_2, \ \forall \lambda \in (0,1)$$

es un punto interior de $L(f,\bar{\alpha})$.
Por lo tanto, $L(f,\bar{\alpha})$ es un conjunto estrictamente convexo. ■

Aunque, la proposición anterior ha sido probada unicamente para funciones definidas en el espacio R^n, los resultados también se satisfacen para un subconjunto convexo y propio $C \subset R^n$, provisto de definiciones topológicas (especialmente las definiciones de frontera e interior) son interpretadas en la topología relativa para C.

En el siguiente ejemplo se observa que convexidad estricta de los conjuntos de nivel inferior no es condición suficiente para la cuasi-convexidad estricta.

Ejemplo 3.3 *Consideremos la función f definida en R^n dada por*

$$f(x) = \begin{cases} 0 & si\,\|x\| \leq 1 \\ \|x\|^2 - 1 & si\,\|x\| > 1 \end{cases}$$

Esta es una función convexa con conjuntos de nivel estrictamente convexos, pero, no es estrictamente cuasi-convexa.

Proposición 3.17 *Una función f estrictamente cuasi-convexa definida en el conjunto convexo $C \subset R^n$ alcanza su mínimo en C en no mas que un punto.*

Demostración. Si x_* es un mínimo local entonces, existe $\delta > 0$, para todo $x \neq x_*$ en el conjunto $C \cap B_\delta(x_*)$ tal que

$$f(x_*) \leq f(x). \tag{3.58}$$

Supongamos que x_* no es un mínimo global de f entonces existe un $\bar{x} \in C$, $\bar{x} \neq x_*$ tal que

$$f(x_*) > f(\bar{x}).$$

Por cuasi-convexidad estricta de f

$$f(\lambda\bar{x} + (1-\lambda)x_*) < \max\{f(\bar{x}), f(x_*)\} = f(x_*), \ \forall\, \lambda \in (0,1). \tag{3.59}$$

En particular consideremos

$$\lambda \in \left(0, \frac{\delta}{\|\bar{x} - x_*\|}\right) \cap (0,1),$$

entonces,

$$\hat{x} = \lambda\bar{x} + (1-\lambda)x_* \in C \cap B_\delta(\bar{x}).$$

de (3.59) obtenemos que

$$f(\hat{x}) < f(x_*), \ \hat{x} \in B_\delta(x_*) \cap C$$

Lo cual contradice (3.58). Por lo tanto, queda demostrado que si x_* es un mínimo local, también es un mínimo global. ∎

3.3 Funciones Semi-estrictas Cuasi-convexas

Definición 3.11 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto convexo C. La función se denomina semi-estricta cuasi-convexa si*

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

donde $x_1, x_2 \in C$ y $0 < \lambda < 1$. Si f es semi-estricta cuasi-convexa entonces $-f$ es una función semi-estricta cuasi-cóncava.

Ejemplo 3.4 *Consideremos la función f definida en $C = [-1,1] \subset R$ dada por*

$$f(x) = \begin{cases} -1 & si\ x \neq 0 \\ -\frac{1}{2} & si\ x = 0 \end{cases}$$

Figura 3.9: Función semi-estricta cuasi-convexa

Esta función no es cuasi-convexa. En efecto, consideremos el conjunto de nivel inferior

$$L(f,-1) = \{x \in C, f(x) \leq -1\}$$

consideremos $f(x) = -1$ entonces

$$x < 0 \ \wedge \ x > 0$$

Supongamos que $L(f,\alpha)$ es un conjunto convexo.
Sean $-1, 1 \in C$ entonces para $\lambda_0 = \frac{1}{2}$ tenemos

$$-\frac{1}{2} + \frac{1}{2} = 0 \in L(f,-1)$$

lo cual es una contradicción. De ahí, $L(f,\alpha)$ no es convexo. Por lo tanto, f no es una función cuasi-convexa.
Notemos que, f no es estrictamente cuasi-convexa.
Supongamos que la función f es estrictamente cuasi-convexa, es decir,

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\} \quad \forall\, \lambda \in (0,1)$$

Tomemos $x_1 = -1$ y $x_2 = 1$ tal que $\lambda_0 = \frac{1}{2}$

$$f(0) < \max\{f(-1), f(1)\}$$

luego $-\frac{1}{2} < -1$, lo cual es una falsedad. Por tanto, f no es estricta cuasi-convexa.
Afirmamos, que f es semi-estricta cuasi-convexa, si $f(x_1) < f(x_2)$ entonces

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\}, \quad \forall\, \lambda \in (0,1).$$

En efecto, desde que $-1 < -\frac{1}{2}$ entonces $x_1 \neq 0$ y $x_2 = 0$.
Luego,

$$f(\lambda x_1 + (1-\lambda)x_2) = f(\lambda x_1) = -1 < \frac{1}{2} = \max\{f(x_1), f(x_2)\}$$

Por tanto,

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\} \quad \forall\, \lambda \in (0,1).$$

Proposición 3.18 *(Karamardian, 1967) Si $f : C \subset R^n \to R$ es una función semi-estricta cuasi-convexa y semicontinua inferior definida en el conjunto convexo C entonces f también es cuasi-convexa.*

Demostración. Debemos demostrar que para $x_1, x_2 \in C$ tal que $f(x_1) = f(x_2)$ tenemos

$$f(\lambda x_1 + (1-\lambda) x_2) \leq f(x_1) \tag{3.60}$$

para todo $0 < \lambda < 1$. Supongamos que (3.60) no se satisface, entonces existe x_0 tal que

$$f(x_0) = f(\lambda_0 x_1 + (1-\lambda_0) x_2) > f(x_1) \tag{3.61}$$

para algún $0 < \lambda_0 < 1$. Por cuasi-convexidad semi-estricta de f, $f(x_1) < f(x_0)$ implica

$$f\left(\bar{\lambda} x_1 + \left(1-\bar{\lambda}\right) x_0\right) < \max\{f(x_1), f(x_0)\}, \ \forall\, \bar{\lambda} \in (0,1) \tag{3.62}$$

donde $\bar{x} = \bar{\lambda} x_1 + \left(1-\bar{\lambda}\right) x_0$, es decir,

$$f(\bar{x}) < f(x_0). \tag{3.63}$$

Si $f(\bar{x}) < f(x_1) = f(x_2)$ desde que $x_0 \in [\bar{x}, x_2]$, por cuasi-convexidad semi-estricta de f tenemos

$$f(x_0) < \max\{f(\bar{x}), f(x_2)\} = f(x_2) = f(x_1) \tag{3.64}$$

luego,

$$f(x_0) < f(x_1).$$

Lo cual contradice (3.61).
Si $f(\bar{x}) > f(x_1) = f(x_2)$, entonces

$$f(x_0) < \max\{f(\bar{x}), f(x_2)\} = f(\bar{x})$$

lo cual contradice (3.63).
Luego, concluimos que

$$f(\bar{x}) = f(x_1) = f(x_2)$$

para todo $\bar{x} = \bar{\lambda} x_1 + \left(1-\bar{\lambda}\right) x_0$.
Por otro lado, desde que (3.60) no se satisface, entonces existe x_0 tal que

$$f(x_0) = f(\lambda_0 x_1 + (1-\lambda_0) x_2) > f(x_2) \tag{3.65}$$

para algún $0 < \lambda_0 < 1$. Por cuasi-convexidad semi-estricta de f, $f(x_2) < f(x_0)$ implica

$$f\left(\hat{\lambda} x_0 + \left(1-\hat{\lambda}\right) x_2\right) < \max\{f(x_0), f(x_2)\}, \ \forall\, \hat{\lambda} \in (0,1) \tag{3.66}$$

donde $\hat{x} = \hat{\lambda} x_0 + \left(1-\hat{\lambda}\right) x_2$, es decir,

$$f(\hat{x}) < f(x_0). \tag{3.67}$$

Si $f(\hat{x}) < f(x_2) = f(x_1)$ desde que $x_0 \in [x_1, \hat{x}]$, por cuasi-convexidad semi-estricta de f tenemos

$$f(x_0) < \max\{f(\hat{x}), f(x_1)\} = f(x_1) = f(x_2) \tag{3.68}$$

luego,

$$f(x_0) < f(x_2).$$

Lo cual contradice (3.65).
Si $f(\hat{x}) > f(x_2) = f(x_1)$, entonces

$$f(x_0) < \max\{f(\hat{x}), f(x_1)\} = f(\hat{x})$$

lo cual contradice (3.67).
Luego, concluimos que

$$f(\hat{x}) = f(x_1) = f(x_2)$$

para todo $\hat{x} = \hat{\lambda} x_0 + \left(1 - \hat{\lambda}\right) x_2$.
Consideremos el conjunto de nivel inferior $L(f, \alpha_1)$, donde $\alpha_1 = f(x_1)$. Notemos de (3.61) que

$$f(x_0) > \alpha_1,$$

es decir, $x_0 \notin L(f, \alpha_1)$. Sean,

$$x_n \in L(f, \alpha_1) \ \wedge \ x_n \to x_0$$

entonces,

$$f(x_n) \leq \alpha_1. \tag{3.69}$$

Además, por la semicontinuidad inferior de f y de (3.69) tenemos

$$f(x_0) \leq \liminf f(x_n) \leq \lim f(x_n) \leq \alpha_1.$$

Por tanto, $x_0 \in L(f, \alpha_1)$, lo cual es una contradicción. ■

El término semi-estricto cuasi-cóncava es dado por Elkin[19]. Otra terminología para funciones semi-estrictas cuasi-cóncavas fue dada por Martos[33] como funciones "explícitamente cuasi-cóncavas".

Definición 3.12 *Sea $f : I \subset R \to R$ una función definida en el intervalo abierto C. Entonces f alcanza un máximo local semi-estricto en el punto $x_0 \in I$ si existen dos puntos $x_1, x_2 \in I, x_1 < x_0 < x_2$ tal que*

$$f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x_0)$$

para todo $0 \leq \lambda \leq 1$ y

$$f(x_1) < f(x_0) \ o \ f(x_2) < f(x_0).$$

Proposición 3.19 *(Diewert, Avriel y Zang, 1981) Sea $f : I \subset R \to R$ una función semicontinua inferior definida en el intervalo abierto I. Si f es semi-estricta cuasi-convexa, entonces no existe un punto $x_0 \in I$ que sea un máximo local semi-estricto lateral de f.*

Demostración. Supongamos que $x_0 \in C$ es un máximo local semi-estricto lateral de f. Sin pérdida de generalidad, asumimos que existen $x_1, x_2 \in C$ tal que $x_1 < x_0 < x_2$, $f(x_2) < f(x_0)$ tal que

$$f(x) < f(x_0), \ \forall x \in (x_1, x_2) \tag{3.70}$$

Sea $\{x_k\} \subset (x_1, x_0)$ una sucesión convergiendo a x_0. De (3.70) tenemos

$$f(x_k) \leq f(x_0), \ \forall \ k. \tag{3.71}$$

Por la semicontinuidad inferior de f,

$$f(\lim x_k) \leq \liminf f(x_k)$$

entonces

$$f(x_0) \leq \liminf f(x_k) \tag{3.72}$$

De (3.71) y (3.72),

$$f(x_0) \leq \liminf f(x_k) \leq \lim f(x_k) \leq f(x_0)$$

Por lo tanto,

$$\lim f(x_k) = f(x_0).$$

Por definición tenemos

$$\forall \ \epsilon > 0, \ \exists \ k_0 \in N \ / \ k > k_0 \ \Rightarrow \ |f(x_k) - f(x_0)| < \epsilon$$

en particular para $\epsilon = f(x_0) - f(x_2) > 0$,

$$f(x_2) - f(x_0) < f(x_k) - f(x_0).$$

Por tanto, para $k > k_0$,

$$f(x_2) < f(x_k) \tag{3.73}$$

Considerando $x_k < x_0 < x_2$, de (3.71) y (3.73) obtenemos

$$f(x_2) < f(x_k) \leq f(x_0).$$

Luego,

$$\max\{f(x_k), f(x_2)\} \leq f(x_0)$$

lo cual contradice la definición de cuasi-convexidad semi-estricta. ■

Proposición 3.20 *Dada la función $f : I \subset R \to R$ con la propiedad de segmento máximo en el intervalo abierto I. Si no existe un punto $x_0 \in C$ que sea un máximo local semi-estricto lateral de f, entonces f es una función semi-estricta cuasi-convexa.*

Demostración. Supongamos que f no es una función semi-estricta cuasi-convexa entonces, existen $x_1, x_2 \in C$ tal que $f(x_1) > f(x_2)$ y para algún $0 < \bar{\lambda} < 1$ tenemos

$$f(\bar{x}) = f\left(\bar{\lambda} x_1 + \left(1 - \bar{\lambda}\right) x_2\right) \geq f(x_1) > f(x_2). \tag{3.74}$$

Sin pérdida de generalidad, consideremos que f alcanza la propiedad de segmento máximo en x_0 entonces,

$$f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x_0), \ \forall\, \lambda \in [0,1]. \tag{3.75}$$

De (3.74) y (3.75), f alcanza un máximo local semi-estricto lateral en un punto intermedio x_0 que se encuentra en el segmento $[x_1, x_2]$. ∎

Teorema 3.8 *Una función $f : C \subset R^n \to R$ semicontinua inferior definida en el conjunto abierto y convexo C es semi-estricta cuasi-convexa si y solo si*

1 *Tiene la propiedad de segmento máximo*

2 *Para todo $x \in C$, todo $v \in R^n$ tal que $v^T v = 1$ y $\bar{t} \in R$, $\bar{t} > 0$ satisfaciendo $x + \bar{t} v \in C$, la función $F(t) = f(x + tv)$ no alcanza un máximo local semi-estricto lateral en $t \in \left(0, \bar{t}\right)$.*

Demostración. Si f es una función semicontinua inferior y semiestricta cuasi-convexa en $C \subset R^n$ por proposición (3.18), f es cuasi-convexa. Por tanto, consideremos que f es cuasi-convexa en un conjunto $C \subset R^n$ abierto y convexo, usando el resultado de la proposición (3.3), f tiene la propiedad de segmento máximo. Lo cual permite demostrar el item 1.

Para demostrar el item 2, afirmamos que $F(t) = f(x + tv)$ es semi-estricta cuasi-convexa en $\left(0, \hat{t}\right)$. Probaremos inicialmente

$$\forall\, t \in \left(0, \hat{t}\right) \Rightarrow x + tv \in C$$

En efecto, sea $0 < t < \hat{t} \Rightarrow 0 < \frac{t}{\hat{t}} < 1$

$$\begin{aligned} x + tv = x + \frac{t}{\hat{t}}\left(\hat{t} v\right) &= \left(\frac{t}{\hat{t}} + 1 - \frac{t}{\hat{t}}\right) x + \frac{t}{\hat{t}}\left(\hat{t} v\right) \\ &= \frac{t}{\hat{t}}\left(x + \hat{t} v\right) + \left(1 - \frac{t}{\hat{t}}\right) x \in C \end{aligned}$$

La última igualdad se cumple desde que C es un conjunto convexo.

Ahora probaremos que F es semi-estricta cuasi-convexa en $\left(0, \bar{t}\right)$. Por tanto, debemos demostrar que

$$F(\lambda t_1 + (1 - \lambda) t_2) < \max\{F(t_1), F(t_2)\}, \ \forall\, t_1, t_2 \in \left(0, \bar{t}\right)$$

En efecto, sean $t_1, t_2 \in (0, \bar{t})$ luego

$$x + t_1 v, x + t_2 v \in C,$$

de la cuasi-convexidad semi-estricta de f tenemos,

$$f(\lambda(x + t_1 v) + (1 - \lambda)(x + t_2 v)) < \max\{f(x + t_1 v), f(x + t_2 v)\}, \ \forall\ \lambda \in (0, 1)$$

La expresión anterior es equivalente a

$$\begin{aligned} F(\lambda t_1 + (1 - \lambda) t_2) = f(x + (\lambda t_1 + (1 - \lambda) t_2) v) &< \max\{f(x + t_1 v), f(x + t_2 v)\} \\ &= \max\{F(t_1), F(t_2)\} \end{aligned}$$

La cual demuestra que F es cuasi-convexa en $(0, \bar{t})$.
Por tanto, usando la Proposición (3.19) la función $F(t) = f(x + tv)$ no alcanza un máximo local lateral semi-estricto en $t \in (0, \bar{t})$.
Reciprocamente, sean $x_1, x_2 \in C$ tal que $x_1 \neq x_2$, consideremos

$$v = \frac{x_2 - x_1}{\|x_2 - x_1\|} \ \wedge\ \bar{t} = \|x_2 - x_1\|. \qquad (3.76)$$

Como $x_1 + tv \in C$, definamos

$$F(t) = f(x_1 + tv), \ \forall t \in (0, \bar{t}).$$

Desde que f tiene la propiedad de segmento máximo sobre $[0, \bar{t}]$, y como F no alcanza un máximo local lateral semi-estricto en $t_0 \in (0, \bar{t})$, entonces, usando la proposición (3.20) tenemos

$$F(t_0) < \max\{F(0), F(\bar{t})\}.$$

La cual es equivalente a,

$$F((1 - \lambda)\bar{t}) = F(t) < \max\{F(0), F(\bar{t})\} \qquad (3.77)$$

Usando (3.76) y reemplazando en (3.77) tenemos

$$f(\lambda x_1 + (1 - \lambda) x_2) < \max\{f(x_1), f(x_2)\}, \forall \lambda \ \in (0, 1).$$

Por lo tanto, f es semi-estricta cuasi-convexa. ■

Proposición 3.21 *Una función continua $f : R^n \to R$ definida en R^n es semi-estricta cuasi-convexa si y solo si $L(f, \alpha)$ es convexo y $Y(f, \alpha) \subset B(f, \alpha)$ ó $Y(f, \alpha) = L(f, \alpha)$ para todo $\alpha \in R$*

Demostración. Supongamos que f es una función semi-estricta cuasi-convexa y desde que f es continua entonces, usando la proposición 3.18 f es cuasi-convexa y $L(f, \alpha)$ es un conjunto convexo para todo $\alpha \in R$.
Consideremos que $B(f, \alpha) \subset Y(f, \alpha)$ entonces existe $x_0 \in Y(f, \alpha_0)$ que es un punto interior de $L(f, \alpha_0)$, y sea $x_1 \neq x_0 \in L(f, \alpha_0)$.

El segmento $[x_1, x_0]$ esta propiamente incluido en $[x_1, x_2] \subset L(f, \alpha)$ para algún $x_2 \neq x_0$.
Supongamos que $f(x_1) \neq f(x_2)$, sin pérdida de generalidad consideremos

$$f(x_2) < f(x_1)$$

por cuasi-convexidad semi-estricta de f tenemos

$$f(x_0) = f(\lambda x_1 + (1-\lambda) x_2) < f(x_1) \leq \alpha_0$$

contradiciendo que $x_0 \in Y(f, \alpha_0)$.
De la misma manera, si consideramos

$$f(x_1) < f(x_2)$$

por cuasi-convexidad semi-estricta

$$f(x_0) = f(\lambda x_1 + (1-\lambda) x_2) < f(x_1) \leq \alpha_0$$

contradiciendo el argumento que $x_0 \in Y(f, \alpha_0)$. De esta manera, $f(x_2) = f(x_1)$ por cuasi-convexidad de f tenemos

$$\alpha_0 \leq f(x_0) \leq f(x_1) \leq \alpha_0$$

de ahí,

$$f(x_1) = \alpha_0$$

Desde que x_0 es un punto arbitrario de $L(f, \alpha_0)$, f es constante en $L(f, \alpha_0)$, lo que implica

$$Y(f, \alpha_0) = L(f, \alpha_0)$$

Reciprocamente, sea $L(f, \alpha)$ conjunto convexo para todo α. Luego, f es cuasi-convexa y para todo $x_1, x_2 \in R^n$ y $0 < \lambda < 1$, si $f(x_2) < f(x_1) = \alpha_1$ implica que

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \alpha_1 \tag{3.78}$$

Por tanto, $x_2 \in L(f, \alpha_1)$ y f no es constante en $L(f, \alpha_1)$, por consiguiente,

$$Y(f, \alpha_1) \neq L(f, \alpha_1)$$

entonces

$$Y(f, \alpha_1) \subset B(f, \alpha_1)$$

sea el punto x_2 en el interior de $L(f, \alpha_1)$, por continuidad existe una vecindad $N_\epsilon(x_2) \subset L(f, \alpha_1)$ tal que para todo x en esta vecindad

$$f(x) < \alpha_1.$$

Consideremos el intervalo abierto (x_1, x_2). Pretendemos demostrar que todo punto en este intervalo es un punto interior de $L(f, \alpha_1)$. Sea $x = \lambda x_1 + (1-\lambda) x_2$, para $0 < \lambda < 1$, donde x es todo punto en el intervalo (x_1, x_2). Entonces la vecindad $B_{(1-\lambda)\epsilon}(x)$ esta en la cápsula convexa de $B_\epsilon(x_2) \cup x_1$ y de esta manera está en $L(f, \alpha_1)$. De ahí, se obtiene la desigualdad estricta de (3.78) y f es semi-estricta cuasi-convexa. ∎

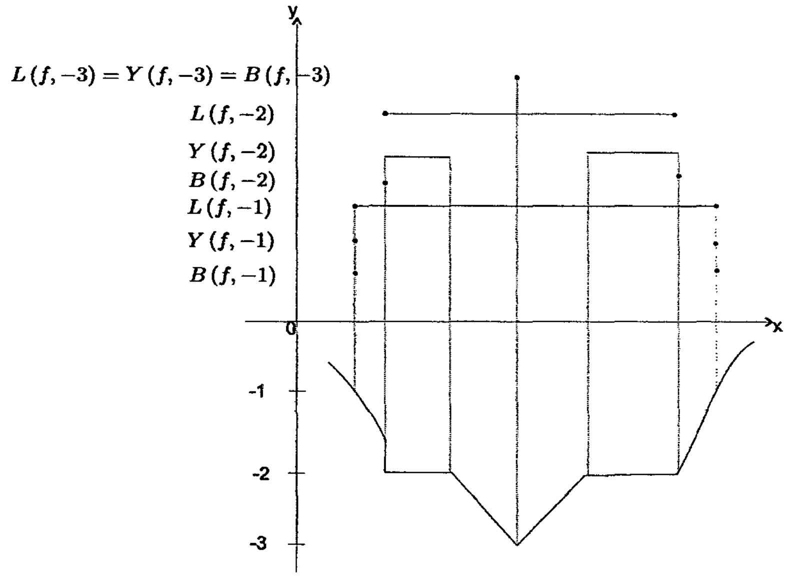

Figura 3.10: Conjunto de nivel inferior y conceptos relacionados

Corolario 3.5 *Sea f una función continua estrictamente cuasi-convexa en R^n. Entonces sus conjuntos de nivel superior $U(f,\alpha)$ son estrictamente convexos y $Y(f,\alpha) \subset B(f,\alpha)$ para todo $\alpha \in R$.*

Demostración. Supongamos que f es una función continua estrictamente cuasi-convexa entonces por proposición sus conjuntos de nivel $L(f,\alpha)$ son estrictamente convexos. Por otro lado, sabemos que toda función estricta cuasi-convexa es semi-estricta cuasi-convexa. De la proposición anterior consideremos

$$Y(f,\alpha) = L(f,\alpha), \ \forall\alpha \in R$$

es decir,

$$L(f,\alpha) \subset Y(f,\alpha)$$

sean $x_1, x_2 \in L(f,\alpha)$,

$$f(x_1) = \alpha \ \wedge \ f(x_2) = \alpha$$

por convexidad estricta de $L(f,\alpha)$ tenemos,

$$\lambda x_1 + (1-\lambda)x_2 \in L(f,\alpha)$$

de ahí,

$$f(\lambda x_1 + (1-\lambda)x_2) = \alpha, \ \forall\lambda \in (0,1)$$

como f es estrictamente cuasi-convexa

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

Por tanto, $\alpha < \alpha$, lo cual es una contradicción.
En consecuencia, si f es estrictamente cuasi-convexa no se satisface

$$Y(f,\alpha) = L(f,\alpha).$$

De la observación, f es semi-estricta cuasi-convexa y desde que $Y(f,\alpha) \neq L(f,\alpha)$ por la proposición anterior tenemos,

$$Y(f,\alpha) \subset B(f,\alpha)$$

Lo cual demuestra el corolario. ■

Teorema 3.9 *Sea f una función cuasi-convexa definida en R^n. Entonces f es semi-estricta cuasi-convexa si y solo si todo mínimo local x_* es también un mínimo global de f en C.*

Demostración. Si x_* es un mínimo local entonces, existe $\delta > 0$, para todo $x \neq x_*$ en el conjunto $C \cap B_\delta(x_*)$ tal que

$$f(x_*) \leq f(x). \tag{3.79}$$

Supongamos que x_* no es un mínimo global de f entonces existe un $\bar{x} \in C$, $\bar{x} \neq x_*$ tal que

$$f(x_*) > f(\bar{x}).$$

Por cuasi-convexidad de f

$$f(\lambda\bar{x} + (1-\lambda)x_*) < \max\{f(\bar{x}), f(x_*)\} = f(x_*), \ \forall\, \lambda \in (0,1). \tag{3.80}$$

En particular consideremos

$$\lambda \in \left(0, \frac{\delta}{\|\bar{x} - x_*\|}\right) \cap (0,1),$$

entonces,

$$\hat{x} = \lambda\bar{x} + (1-\lambda)x_* \in C \cap B_\delta(\bar{x}).$$

de (3.80) obtenemos que

$$f(\hat{x}) < f(x_*), \ \hat{x} \in B_\delta(x_*) \cap C$$

Lo cual contradice (3.79). Por lo tanto, queda demostrado que si x_* es un mínimo local, también es un mínimo global.
Reciprocamente, supongamos que f es cuasi-convexa y que todo mínimo local de f es un mínimo global.
Por proposición 3.21 es suficiente mostrar que $Y(f,\alpha)$ contiene un punto interior de $L(f,\alpha)$, entonces

$$Y(f,\alpha) = L(f,\alpha),$$

esto es, f es constante en $L(f,\alpha)$.
Si f es constante en R^n o $\alpha = \max\{f(x) : x \in R^n\}$ es claro que,

$$Y(f,\alpha) = L(f,\alpha).$$

Asumimos, que f no es constante en R^n y sea $x_0 \in Y(f,\alpha_0)$ un punto interior de $L(f,\alpha_0)$ para algún α_0 tal que el conjunto

$$L^0(f,\alpha_0) = \{x : x \in R^n, f(x) < \alpha_0\}$$

es no vacío. Consideremos

$$U(f,\alpha_0) = \{x : x \in R^n, f(x) \geq \alpha_0\}$$

Tomemos,

$$x_n \in U(f,\alpha_0) \quad \wedge \quad x_n \to x$$

por definición

$$f(x_n) \geq \alpha_0$$

aplicando límite cuando $n \to \infty$

$$\lim_{n\to\infty} f(x_n) \geq \alpha_0$$

desde que f es continua por tanto

$$x \in U(f,\alpha_0).$$

Concluimos que $U(f,\alpha_0)$ es cerrado, en consecuencia

$$R^n - U(f,\alpha_0)$$

es abierto, es decir,

$$L^0(f,\alpha_0)$$

es abierto. Además, desde que f es cuasi-convexa entonces el conjunto

$$L(f,\alpha_0) = \{x : x \in R^n : f(x) \leq \alpha_0\}$$

es convexo. Por el teorema de separación para el conjunto convexo $L^0(f,\alpha_0)$ y $x_0 \notin L^0(f,\alpha_0)$. Entonces existe un hiperplano H_c^a tal que para todo $x \in L^0(f,\alpha_0)$

$$\begin{array}{ll} \langle c,x\rangle \leq a & x \in L^0(f,\alpha_0) \\ \langle c,x\rangle = a & x = x_0 \notin L^0(f,\alpha_0) \end{array}$$

para $c \in R^n$ y $a \in R$. Desde que x_0 es un punto interior de $L(f,\alpha_0)$ podemos encontrar $\hat{x}$ en el interior de $L(f,\alpha_0)$ tal que,

$$\hat{x} \neq x_0 \ \wedge \ \langle c,\hat{x}\rangle > a$$

Por la continuidad de f, existe un $\epsilon > 0$ tal que

$$N_\epsilon(\hat{x}) \subset L(f,\alpha) \ \wedge \ \langle c,x\rangle > a \ \forall x \in N_\epsilon(\hat{x})$$

Consecuentemente,

$$N_\epsilon \cap L^0(f,\alpha_0) = \phi$$

y por lo tanto

$$N_\epsilon(\hat{x}) \subset Y(f,\alpha_0)$$

desde que $\hat{x}$ es un mínimo local de f entonces si $\hat{x} \in N_\epsilon(\hat{x})$

$$f(\hat{x}) \leq f(x) = \alpha_0$$

luego, por hipótesis $\hat{x}$ es un mínimo global de f. Si $x \in L(f,\alpha_0)$ entonces

$$f(\hat{x}) \leq f(x) = \alpha_0.$$

Por tanto,

$$Y(f,\alpha_0) = L(f,\alpha_0)$$

por proposición (3.21) f es semi-estricta cuasi-convexa. ■

$$f(x_1) \leq f(x_2) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) \leq 0$$

entonces f es semi-estricta cuasi-convexa. El recíproco de este resultado no es válido, el cual se puede observar en el siguiente ejemplo.

Ejemplo 3.5 *Sea f una función definida en el intervalo $C = [0,1] \subset R$ dada por*

$$f(x) = -x^2$$

Se observa que f es semi-estricta cuasi-convexa. En efecto, sean $x_1, x_2 \in C$,
$0 < \lambda < 1$ tal que $f(x_1) < f(x_2)$, entonces

$$-x_1^2 < -x_2^2$$

luego, desde que $0 < x_1 < 1$ y $0 < x_2 < 1$

$$x_2 < x_1.$$

Por tanto,

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

de ahí,

$$-(\lambda x_1 + (1-\lambda)x_2)^2 < -x_2^2$$

en consecuencia,

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

Lo que implica que f es una función estrictamente cuasi-convexa.
Por otro lado, sea $x_1 \in (0,1]$ y $x_2 = 0$ entonces

$$-x_1^2 = f(x_1) < f(x_2) = 0$$

no implica que

$$(x_1 - x_2) f'(x_2) < 0$$

3.4 Funciones Pseudo-convexas

En la sección de funciones cuasi-convexas se dijo que si una función cuasi-convexa diferenciable satisface $f(x_1) \leq f(x_2)$ implica que

$$(x_1 - x_2)^T \nabla f(x_2) \leq 0.$$

Además, debemos resaltar que para una función cuasi-convexa, $\nabla f(x_0) = 0$ no implica que x_0 sea un mínimo global. Esta observación permite introducir el concepto de función pseudo-convexa, la cual se obtiene de una pequeï¿½a modificación de la condición antes mencionada. La definición de pseudo-convexidad para funciones diferenciables fue introducida por Magasarian.

Definición 3.13 *Una función diferenciable $f : C \subset R^n \to R$ definida en el conjunto abierto y convexo C se dice pseudo-convexa, si para $x_1, x_2 \in C$*

$$f(x_1) < f(x_2) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) < 0.$$

Esta es llamada estrictamente pseudo-convexa si $x_1 \neq x_2$ y

$$f(x_1) \leq f(x_2) \text{ implica que } (x_1 - x_2)^T \nabla f(x_2) < 0$$

Definición 3.14 *Sea $f : C \subset R^n \to R$ una función definida en el conjunto abierto y convexo C. Esta función es llamada pseudo-convexa si para $x_1, x_2 \in C$ y $\lambda \in (0, 1)$*

$$f(x_1) < f(x_2) \text{ implica que } f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x_2) - (1 - \lambda) \lambda b(x_1, x_2)$$

donde $b(x_1, x_2)$ es un número positivo, dependiendo en general de x_1 y x_2. La función f es llamada estrictamente pseudo-convexa si $x_1 \neq x_2$ y

$$f(x_1) \leq f(x_2) \text{ implica que } f(\lambda x_1 + (1 - \lambda) x_2) \leq f(x_2) - (1 - \lambda) \lambda b(x_1, x_2)$$

Diewert[17] usa la definición de funciones pseudo-convexas usando derivadas direccionales.

Proposición 3.22 *Para funciones diferenciales las definiciones 3.13 y 3.14 son equivalentes.*

Demostración. Supongamos que se satisface la definición 3.14. Si f es una función diferenciable entonces para todo $x_1, x_2 \in C$ tenemos

$$f(x_1) = f(x_2) + (x_1 - x_2)^T \nabla f(x_2) + \alpha(x_1, x_2) \|x_1 - x_2\|$$

reemplazando $\lambda x_1 + (1 - \lambda) x_2$ en x_1, para $\lambda \in (0, 1)$

$$f(\lambda x_1 + (1 - \lambda) x_2) = f(x_2) + \lambda (x_1 - x_2)^T \nabla f(x_2) + \lambda \alpha (x_2 + \lambda (x_2 - x_1), x_2) \|x_1 - x_2\|$$

luego,

$$\lambda (x_1 - x_2)^T \nabla f(x_2) + \lambda \alpha (x_2 + \lambda (x_1 - x_2), x_2) \|x_1 - x_2\| = f(\lambda x_1 + (1 - \lambda) x_2) - f(x_2)$$

donde $\alpha(x_1,x_2)$ es una función tal que $\alpha(x_1,x_2) \longrightarrow 0$ cuando $x_1 \longrightarrow x_2$. Usando la definición 3.14

$$\lambda(x_1-x_2)^T \nabla f(x_2) + \lambda\alpha(x_2+\lambda(x_1-x_2),x_2)\|x_1-x_2\| \leq -(1-\lambda)\lambda b(x_1,x_2)$$

para todo $0 \leq \lambda \leq 1$. Dividiendo entre λ y tomando el límite cuando $\lambda \longrightarrow 0$

$$(x_1-x_2)^T \nabla f(x_2) \leq -b(x_1,x_2) < 0.$$

Entonces$f$ satisface la definición 3.13. Reciprocamente, supongamos que no se satisface la definición 3.14, si $f(x_1) < f(x_2)$ y para todo positivo $b(x_1,x_2)$ existe $\lambda \in (0,1)$ tal que

$$\frac{f(\lambda x_1+(1-\lambda)x_2)-f(x_2)}{\lambda} > -(1-\lambda)b(x_1,x_2) > -b(x_1,x_2).$$

En particular, existe $0 < \lambda_i < 1$ tal que

$$\frac{f(x_2+\lambda_i(x_1-x_2))-f(x_2)}{\lambda_i} > -\frac{1}{i}$$

La sucesión $\{\lambda_i\}$ es acotada, por tanto, posee una subsucesión convergente, esto es,

$$\lim_{i_k\to\infty} \lambda_{i_k} = \lambda_0$$

entonces,

$$\lim_{i_k\to\infty} \{x_2+\lambda_{i_k}(x_1-x_2)\} = x_0$$

donde $x_0 = x_2 + \lambda_0(x_1-x_2)$. Si $x_0 \neq x_2$, entonces, $\lambda_0 \neq 0$. En efecto, supongamos que $\lambda_0 = 0$ entonces $x_0 = x_2$ lo cual contradice la hipótesis.

$$f(x_0) \geq f(x_2)$$

f no es estrictamente cuasi-convexa, de aqui, f no es una función pseudo-convexa, satisfaciendo la definición 3.13.
Si $x_0 = x_2$ entonces $\lambda_{i_k} \longrightarrow 0$ y

$$\frac{f(x_2+\lambda_{i_k}(x_1-x_2))-f(x_2)}{\lambda_{i_k}} > -\frac{1}{i_k}$$

aplicando límite cuando $i_k \longrightarrow \infty$

$$\lim_{i_k\to\infty} \frac{f(x_2+\lambda_{i_k}(x_1-x_2))-f(x_2)}{\lambda_{i_k}} \geq 0$$

luego,

$$(x_1-x_2)^T \nabla f(x_2) \geq 0$$

Consecuentemente, f no satisface la definición 3.13. ■

La prueba del caso estrictamente pseudo-convexo, es idéntico. En efecto, supongamos que se satisface el segundo item de la definición 3.14. Si f es una función diferenciable entonces para todo $x_1, x_2 \in C$ tenemos para $0 < \lambda < 1$

$$f(x_1) = f(x_2) + (x_1-x_2)^T \nabla f(x_2) + \alpha(x_1,x_2)\|x_1-x_2\|$$

entonces

$$f(\lambda x_1 + (1-\lambda) x_2) = f(x_2) + \lambda (x_1 - x_2)^T \nabla f(x_2) + \lambda \alpha (x_2 + \lambda (x_2 - x_1), x_2) \|x_1 - x_2\|$$

luego,

$$\lambda (x_1 - x_2)^T \nabla f(x_2) + \lambda \alpha (x_2 + \lambda (x_1 - x_2), x_2) \|x_1 - x_2\| = f(\lambda x_1 + (1-\lambda) x_2) - f(x_2)$$

donde $\alpha(x_1, x_2)$ es una función tal que $\alpha(x_1, x_2) \to 0$ cuando $x_1 \to x_2$

$$\lambda (x_1 - x_2)^T \nabla f(x_2) + \lambda \alpha (x_2 + \lambda (x_1 - x_2), x_2) \|x_1 - x_2\| \leq -(1-\lambda) \lambda b(x_1, x_2)$$

para todo $0 \leq \lambda \leq 1$. Dividiendo entre λ y tomando el límite cuando $\lambda \to 0$

$$(x_1 - x_2)^T \nabla f(x_2) \leq b(x_1, x_2) < 0$$

Reciprocamente, supongamos que no se satisface la definición 3.14, si $f(x_1) \leq f(x_2)$ tal que $x_1 \neq x_2$ y para todo positivo $b(x_1, x_2)$ existe un $0 < \lambda < 1$ tal que

$$\frac{f(\lambda x_1 + (1-\lambda) x_2) - f(x_2)}{\lambda} > -(1-\lambda) b(x_1, x_2) > -b(x_1, x_2).$$

En particular, existe $0 < \lambda_i < 1$ tal que

$$\frac{f(x_2 + \lambda_i (x_1 - x_2))}{\lambda_i} > -\frac{1}{i}$$

La sucesión es acotada, por tanto, posee una subsucesión convergente

$$\lambda_{i_k} \to \lambda_0 \ \wedge \ \{x_2 + \lambda_{i_k} (x_1 - x_2)\} \to x_0$$

Si $x_0 \neq x_2$, esto es, $\lambda_0 \neq 0$

$$f(x_0) \geq f(x_2)$$

f no es estrictamente cuasi-convexa, de aquí, f no es una función estrictamente pseudo-convexa, satisfaciendo la definición 3.13.
Si $x_0 = x_2$ cuando $\lambda_{i_k} \to 0$ y

$$\frac{f(x_2 + \lambda_{i_k} (x_1 - x_2)) - f(x_2)}{\lambda_{i_k}} > -\frac{1}{i_k}$$

aplicando límite cuando $i_k \to \infty$

$$\lim_{i_k \to \infty} \frac{f(x_2 + \lambda_{i_k} (x_1 - x_2)) - f(x_2)}{\lambda_{i_k}} \geq 0$$

de ahí,

$$(x_1 - x_2) \nabla f(x_2) \geq 0$$

Consecuentemente, f no satisface la definición 3.13.

Teorema 3.10 *(Mangasarian,1965) Sea $f : C \subset R^n \to R$ una función diferenciable y (estrictamente) pseudo-convexa en el conjunto abierto y convexo C. Si $\nabla f(x_0) = 0$ para algún $x_0 \in C$, entonces x_0 es un mínimo global (estricto) de f sobre C.*

Demostración. Desde que f es pseudo-convexa, por la definición 3.13 tenemos que

$$(x_1 - x_2)^T \nabla f(x_2) \geq 0 \text{ implica que } f(x_1) \geq f(x_2)$$

para $x_1, x_2 \in C$. Por hipótesis $\nabla f(x_0) = 0$, considerando

$$x_2 = x_0 \ \wedge \ x_1 = x$$

entonces

$$(x - x_0)^T \nabla f(x_0) = 0$$

lo cual implica

$$f(x_0) \leq f(x), \ \forall x \in C$$

entonces x_0 es un mínimo global.
Igualmente, por la definición 3.13 desde que $\nabla f(x_0) = 0$ consideremos

$$(x - x_0)^T \nabla f(x_0) = 0$$

por pseudo-convexidad estricta

$$f(x_0) < f(x), \ \forall x \in C$$

entonces x_0 es un mínimo global estricto. ■

Es interesante resaltar, que no se conoce caracterización de pseudo-convexidad en términos de conjunto de nivel. Los conjuntos de nivel superior de funciones pseudo-convexas no son necesariamente estrictamente convexas. Sin embargo, Ponstein [40] probó que si f es diferenciable, pseudo-convexa y estrictamente pseudo-convexa (sus conjuntos de nivel son estrictamente convexos) entonces f es estrictamente pseudo-convexa.

Ejemplo 3.6 *Consideremos la función f definida en el conjunto convexo*

$$C = \left\{x : x \in R^2, 0 < x_1 < 1, 0 < x_2 < 1\right\}$$

es dada por $f(x) = -x_1^2 - x_1$.

Esta función es pseudo-convexa, en efecto, supongamos que $f(x) < f(y)$ entonces $x_1 > y_1$. En efecto, supongamos que

$$x_1 \leq y_1$$

de ahí,

$$-x_1 \geq -y_1 \ \wedge -x_1^2 \geq -y_1^2$$

luego

$$-x_1 - x_1^2 \geq -y_1 - y_1^2$$

es decir,

$$f(x) \geq f(y)$$

lo cual es una contradicción.

$$(x-y)\nabla f(y)(x_1-y_1,x_2-y_2)(2y_1-1,0) = (x_1-y_1)(-2-y_1) < 0$$

La función no es convexa, es cóncava. Además, notemos que f no es estrictamente pseudo-convexa, consideremos

$$x = (\alpha,\beta) \ \wedge \ y = (\alpha,\gamma)$$

donde α,β,γ estan en $(0,1)$ tal que

$$\alpha < \beta \ \wedge \ \alpha < \gamma, \ \beta \neq \gamma$$

entonces $f(x) = f(y)$, pero

$$(x-y)^T \nabla f(y) = 0$$

Esta función no es estrictamente cuasi-convexa. Supongamos que f es estrictamente cuasi-convexa, es decir,

$$f(\lambda x + (1-\lambda)y) < \max\{f(x), f(y)\}$$

por tanto,

$$0 < -\alpha^2 - \alpha$$

desde que $0 < \alpha < 1$ tenemos

$$0 < 0$$

lo cual es una contradicción.
Sin embargo, la función f es semi-estricta cuasi-convexa, si $f(x) < f(y)$ de ahi

$$x_1^2 - x_1 < -y_1^2 - y_1$$

entonces $x_1 > y_1$, luego

$$\lambda x_1 + (1-\lambda)y_1 > y_1$$

lo cual es equivalente a

$$-\lambda x_1 - (1-\lambda)y_1 < -y_1$$

Por tanto,

$$f(\lambda x + (1-\lambda)y) < f(y)$$

Lo cual implica que f, es semi-estricta cuasi-convexa.
Además, la función f es cuasi-convexa.

Proposición 3.23 *(Diewert, Avriel y Zang, 1981) Sea $f : I \subset R \longrightarrow R$ una función continuamente diferenciable definida en el intervalo abierto I. Entonces f es pseudo-convexa si y solo si para todo $x_0 \in I$ tal que $f'(x_0) = 0$ es un mínimo local de f.*

Demostración. Si f es pseudo-convexa, entonces $f'(x_0) = 0$ implica que x_0 es un mínimo local de f. Reciprocamente, sean $x_1, x_2 \in C$, $x_1 \neq x_2$ dos puntos tal que

$$(x_1 - x_2) f'(x_2) \geq 0$$

y supongamos que $f(x_1) < f(x_2)$. Si $(x_1 - x_2) f'(x_2) > 0$ entonces la función

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

alcanza al menos un máximo local sobre el intervalo $(0, 1)$. Sea λ_0 el mayor punto que maximiza g sobre $(0, 1)$. Entonces,

$$g'(\lambda_0) = f'(\lambda_0 x_1 + (1 - \lambda_0) x_2) = f'(x_0) = 0,$$

pero x_0 no es un mínimo local de f, contradiciendo la hipótesis.
Si $(x_1 - x_2) f'(x_2) = 0$ entonces $f'(x_2) = 0$ y por la hipótesis, x_2 es un mínimo local de f. Desde que $f(x_1) < f(x_2)$ concluimos que la función $g(\lambda)$ debe alcanzar al menos un máximo local en el intervalo abierto $(0, 1)$, llegando a una contradicción. Por tanto, $f(x_1) \leq f(x_2)$ y f es pseudo-convexa. ∎

Teorema 3.11 *Sea $f : C \subset R^n \to R$ una función continuamente diferenciable definida en el conjunto convexo C. Entonces f es pseudo-convexa si y solo si para todo $x_0 \in C$ y $v \in R^n$ tal que $v^T v = 1$ y $v^T \nabla f(x_0) = 0$ la función $F(t) = f(x_0 + tv)$ alcanza un mínimo local en $t = 0$.*

Demostración. Consideremos $F(t) = f(x_0 + tv)$, en $t = 0$ tenemos

$$F'(0) = f(x_0) = 0$$

Desde que F es continuamente diferenciable y usando la proposición (3.23), F alcanza su máximo local en $t = 0$ si y solo si F es pseudo-convexa. ∎

Si una función es pseudo-convexa y pseudo-cóncava es llamada pseudo-afin. Las funciones pseudo-afines fueron caracterizadas por Thompson y Parke (1973). Concluimos esta sección introduciendo las clases de funciones fuertemente pseudo-convexas, una subclase de las funciones diferenciales estrictamente pseudo-convexas.

Definición 3.15 *(Diewert, Avriel y Zang, 1981) Una función diferenciable $f : C \subset R^n \to R$ definida en el conjunto abierto y convexo C se dice fuertemente pseudo-convexa si la función es estrictamente pseudo-convexa y para todo $x_0 \in C$ y $v \in R^n$ tal que $v^T v = 1$ y $v^T \nabla f(x_0) = 0$ existen números positivos ϵ y α tal que $x_0 + \epsilon v \in C$ y*

$$F(t) = f(x_0 + tv) \geq F(0) + \frac{1}{2}\alpha(t)^2$$

para $0 \leq t < \epsilon$.

Las funciones fuertemente pseudo-convexas tienen una muchas aplicaciones en economía. El siguiente concepto presenta un mínimo local fuerte en un segmento.

Definición 3.16 *(Diewert, Avriel y Zang, 1981) Una función $f : C \subset R \to R$ definida en el intervalo abierto C alcanza un mínimo local fuerte en x_0 si existen números positivos ϵ y α tal que $x_0 \pm \epsilon \in C$ y*

$$f(x) \geq f(x_0) + \frac{1}{2}\alpha (x - x_0)^2$$

para $|x| < \epsilon$

Proposición 3.24 *(Diewert, Avriel y Zang, 1981) Sea $f : C \subset R^n \to R$ una función diferenciable definida en el conjunto abierto y convexo C. Entonces f es fuertemente pseudo-convexa si y solo si para todo $x_0 \in C$ y $v \in R^n$ tal que $v^T v = 1$ y $v^T \nabla f(x_0) = 0$ la función $F(t) = f(x_0 + tv)$ alcanza un mínimo local fuerte en $t = 0$.*

Demostración. Supongamos que f es fuertemente pseudo-convexa, entonces para todo $x_0 \in C$ tal que $v^T \nabla f(x_0) = 0$ existen ϵ y δ positivos tal que $x_0 \pm \epsilon \in C$ y

$$F(t) = f(x_0 + tv) \geq F(0) + \frac{1}{2}\alpha t^2$$

para $0 \leq t < \epsilon$, donde $F(0) = f(x_0)$ y $F'(0) = v^T \nabla f(x_0) = 0$.
Afirmamos que, $F(0) \leq F(-t)$ para $0 < t < \epsilon$
En efecto, procediendo por el absurdo consideremos que $F(-t) < F(0)$, de ahí

$$F(-t) \leq F(0)$$

por pseudo-convexidad estricta

$$F'(0)(-t) < 0$$

lo cual es una contradicción, desde que $F'(0) = 0$. Por tanto,

$$F(0) - \frac{1}{2}\alpha t^2 \leq F(0) \leq F(-t)$$

para $0 \leq t < \epsilon$. Concluimos,

$$F(t) \geq F(0) + \frac{1}{2}\alpha t^2$$

para $|t| \leq \epsilon$. Finalmente, $F(t) = f(x_0 + tv)$ alcanza alcanza un mínimo local fuerte en $t = 0$.
Reciprocamente, supongamos que la función alcanza un mínimo local fuerte en $t = 0$, es decir,

$$F(t) \geq F(0) + \frac{1}{2}\alpha t^2$$

para $|t| < \epsilon$. Entonces, por definición f es fuertemente pseudo-convexa. ∎

Capítulo 4

Optimización Cuasi-convexa

Consideremos el problema de minimización:

$$\begin{cases} \min f(x) \\ s.a \\ g(x) \leq 0 \\ \quad x \geq 0 \end{cases}$$

donde f y g son funciones cuasi-convexas y diferenciables de un vector n-dimensional $x = (x_1, ..., x_n)$.
H. W. Kuhn y A. W. Tucker en su artículo de programación no lineal, demuestran quelas condiciones de cualificación, son condiciones necesarias para que x^0 minimice $f(x)$ sujeto a $g(x) \leq 0$ y $x \geq 0$ (las condiciones de Karush-Kuhn-Tucker, o KKT) son

$$\begin{aligned} \frac{\partial f(x^0)}{\partial x_i} + \lambda^0 \frac{\partial g(x^0)}{\partial x_i} &\geq 0 \\ x^0 \left(\frac{\partial f(x^0)}{\partial x_i} + \lambda^0 \frac{\partial g(x^0)}{\partial x_i} \right) &= 0 \\ \lambda^0 g(x^0) &= 0 \\ \lambda^0 &\geq 0 \end{aligned}$$

Kuhn y Tucker también probaron que si $f(x)$ y $g(x)$ son funciones convexas, las condiciones KKT son condiciones suficientes para hallar un mínimo.

4.1 Condiciones Suficientes para un Mínimo

Consideremos una variable relevante la cual puede tomar un valor positivo satisfaciendo las restricciones. Formalmente, x_i es una variable relevante cuando existe algún punto en el conjunto restricción, x^* el cual $x_i^* > 0$.

Teorema 4.1 *Sean $f : R^n \to R$, $g : R^n \to R$ funciones cuasi-convexas y diferenciables, definidas para $x \geq 0$.*
Si x^0 y λ^0 satisfacen las condiciones KKT y al menos una de las siguientes condiciones:

a. $\frac{\partial f(x^0)}{\partial x_i} > 0$ *para al menos una componente de la variable* x_i^0.

b. $\frac{\partial f(x^0)}{\partial x_j} < 0$ *para al menos una componente de la variable relevante* x_j^1.

c. $\frac{\partial f(x^0)}{\partial x_i} \neq 0$ *para todo* i *y* $f(x)$ *es dos veces diferenciable en la vecindad de* x^0.

d. $f(x)$ *es convexa.*

Entonces x^0 *minimiza* $f(x)$ *sujeto a las restricciones* $g(x) \leq 0$ *donde* $x \geq 0$.

Demostración. Consideremos la siguiente identidad

$$\frac{\partial f(x^0)}{\partial x_i}\left(x_i^1 - x_i^0\right) = \left(x_i^1 - x_i^0\right)\left(\frac{\partial f(x^0)}{\partial x_i} + \lambda_i^0 \frac{\partial g(x^0)}{\partial x_i}\right) - \lambda_i^0 \frac{\partial g(x^0)}{\partial x_i}\left(x_i^1 - x_i^0\right), \ \forall i = 1, ..., n.$$

Si x^0 satisface las condiciones de KKT, entonces

$$\frac{\partial f(x^0)}{\partial x_i} + \lambda_i^0 \frac{\partial g(x^0)}{\partial x_i} \geq 0,$$

desde que $x_i^1 \geq 0$, tenemos

$$x_i^1\left(\frac{\partial f(x^0)}{\partial x_i} + \lambda_i^0 \frac{\partial g(x^0)}{\partial x_i}\right) \geq 0. \tag{4.1}$$

Por otro lado, afirmamos que

$$\lambda_i^0 \frac{\partial g(x^0)}{\partial x_i}\left(x_i^1 - x_i^0\right) \leq 0$$

En efecto, si $\lambda_j^0 = 0$ entonces la j-ésima componente del término se anula.
Si $\lambda_j^0 > 0$ se tiene $g^j(x^0) = 0$. Pero, considerando x^1 en el conjunto restricción implica que

$$g^j\left(x^1\right) \leq 0$$

luego,

$$g^j\left(x^1\right) \leq g^j\left(x^0\right).$$

Por la cuasi-convexidad de g tenemos

$$\frac{\partial g(x^0)}{\partial x_j}\left(x_i^1 - x_i^0\right) \leq 0.$$

Como $\lambda_i^0 \geq 0$, entonces

$$-\lambda_i^0 \frac{\partial g(x^0)}{\partial x_j}\left(x_i^1 - x_i^0\right) \geq 0. \tag{4.2}$$

De (4.1) y (4.2) tenemos,

$$\frac{\partial f(x^0)}{\partial x_i}\left(x_i^1 - x_i^0\right) \geq 0, \ \ \forall\, i = 1, ..., n. \tag{4.3}$$

Por lo tanto, para $f(x)$ y $g(x)$ funciones diferenciables y cuasi-convexas con $g(x^1) \leq 0$, $x^1 \geq 0$ se satisface

$$\left(x^1 - x^0\right)^T \nabla f\left(x^0\right) \geq 0.$$

a. $\frac{\partial f(x^0)}{\partial x_i} > 0$ para al menos una coordenada de la variable x_0.
Consideremos h como un vector unitario en la i-ésima dirección $x^2 = x^0 + h$

$$\left(x^2 - x^0\right)^T \nabla f\left(x^0\right) = \nabla f\left(x^0\right) h = \frac{\partial f\left(x^0\right)}{\partial x_i} > 0. \tag{4.4}$$

Para x^1 en el conjunto de restricciones

$$x^1(\theta) = (1-\theta)\, x^1 + \theta x^2 \ \wedge \ x^0(\theta) = (1-\theta)\, x^0 + \theta x^2$$

entonces

$$\left[x^0(\theta) - x^0\right]^T \nabla f\left(x^0\right) = \left[\theta x^2 - \theta x^0\right]^T \nabla f\left(x^0\right)$$

lo cual es equivalente a,

$$\left(x^2 - x^0\right)^T \theta \nabla f\left(x^0\right) > 0$$

para $\theta > 0$. De (4.4), tenemos

$$\left[x^0(\theta) - x^0\right]^T \nabla f\left(x^0\right) > 0. \tag{4.5}$$

Usando el resultado (4.3) obtenemos

$$\left[x^1(\theta) - x^0(\theta)\right]^T \nabla f\left(x^0\right) = (1-\theta)\left(x^1 - x^0\right)^T \nabla f\left(x^0\right) \geq 0 \tag{4.6}$$

para $\theta \leq 1$. Sumando las expresiones (4.5) y (4.6)

$$\left[x^1(\theta) - x^0\right]^T \nabla f\left(x^0\right) > 0, \ \forall\, \theta \in [0,1]\,. \tag{4.7}$$

Por la cuasi-convexidad de la función f tenemos

$$f\left(x^1(\theta)\right) > f\left(x^0\right) \tag{4.8}$$

Aplicando límite cuando θ se aproxima a 0, entonces $x^1(\theta) \to x^1$. De donde se concluye que

$$\lim_{\theta \to 0} f\left(x^1(\theta)\right) \geq f\left(x^*\right) \tag{4.9}$$

Por lo tanto, $f\left(x^1\right) \leq f\left(x^0\right)$.

b. $\frac{\partial f(x_0)}{\partial x_j} < 0$ para alguna variable relevante x_j^1.
Si excluimos el caso (a), $x^0 \nabla f\left(x^0\right) < 0$ y (b) son equivalentes.
En efecto, si $x^0 \nabla f\left(x^0\right) < 0$ entonces

$$x_j^0 \frac{\partial f\left(x^0\right)}{\partial x_j} < 0$$

desde que $x^0 \geq 0$ existe x_j^0, en el conjunto restricción tal que $x_j^0 > 0$ luego,

$$\frac{\partial f\left(x^0\right)}{\partial x_j} < 0 \tag{4.10}$$

para alguna variable relevante x_j.

Reciprocamente, notemos que $g(x^1) \leq 0$ para $x^1 \geq 0$ implica que

$$\nabla f(x^0)^T (x^1 - x^0) \geq 0$$

entonces,

$$\nabla f(x^0)^T x^1 \geq \nabla f(x^0)^T x^0 \tag{4.11}$$

para todo x^1 en el conjunto de restricciones Excluyendo (a)

$$\nabla f(x^0) \leq 0 \tag{4.12}$$

Además, (b) implica que para algun x^* en el conjunto de restricciones y para algún i, $\frac{\partial f(x^0)}{\partial x_i} < 0$ y $x_i^* > 0$, lo cual implica que $\frac{\partial f(x^0)}{\partial x_i} x_i^* < 0$. Consideremos $x^1 = x^*$ en

$$\frac{\partial f(x^0)}{\partial x_i} x_i^0 \leq \frac{\partial f(x^0)}{\partial x_i} x_i^* < 0 \tag{4.13}$$

entonces $\nabla f(x^0) x^0 < 0$ para alguna variable relevante.

Para x^1 en el conjunto de restricciones, consideremos

$$x^1(\theta) = (1-\theta) x^1 \quad \wedge \quad x^0(\theta) = (1-\theta) x^0$$

entonces

$$\nabla f(x^0) \left[x^0(\theta) - x^0\right] = \nabla f(x^0) \left(-\theta x^0\right)$$

lo cual es equivalente a,

$$\theta \nabla f(x^0) \left(-x^0\right) > 0$$

para $\theta > 0$. Por tanto,

$$\nabla f(x^0) \left[x^0(\theta) - x^0\right] > 0. \tag{4.14}$$

Usando el resultado (4.3) obtenemos

$$\nabla f(x^0) \left[x^1(\theta) - x^0(\theta)\right] = (1-\theta) \nabla f(x^0) \left(x^1 - x^0\right) \geq 0 \tag{4.15}$$

para $\theta \leq 1$. Sumando las expresiones (4.14) y (4.15)

$$\nabla f(x^0) \left[x^1(\theta) - x^0\right] > 0, \ \forall\, \theta \in [0,1]. \tag{4.16}$$

Por la cuasi-convexidad de la función f tenemos

$$f\left(x^1(\theta)\right) > f\left(x^0\right) \tag{4.17}$$

Aplicando límite cuando θ se aproxima a 0, entonces $x^1(\theta) \to x^1$. De donde se concluye que

$$\lim_{\theta \to 0} f\left(x^1(\theta)\right) \geq f(x^*) \tag{4.18}$$

Por lo tanto, $f(x^1) \leq f(x^0)$.

c. $\nabla f(x^0) \neq 0$ y $f(x)$ es dos veces diferenciable en una vecindad de x^0.

En efecto, la partición del vector x^0 en dos subvectores y^0 y z^0 correspondientes a las variables relevantes e irrelevantes respectivamente. Entonces, si excluimos los casos ya estudiados asumiremos $\nabla f(x^0) \neq 0$ entonces

$$\nabla f(x^0) = (\nabla f(y^0), \nabla f(z^0)) \neq 0 \tag{4.19}$$

tal que $\nabla f(y^0) = 0$ y $\nabla f(z^0) \leq 0$ donde $\frac{\partial f(z^0)}{\partial z_i} < 0$ para algún z_i. Por la definición de la varible irrelevante, $z^0 = 0$ y $z^1 = 0$ para todo $x^1 = (y^1, z^1)$ en el conjunto de restricciones. Por lo tanto, para probar el teorema es suficiente probar que

$$f(y^0, 0) \leq f(y^1, 0) \tag{4.20}$$

Definamos la función

$$\phi(u, v) = f\left[(1-u)y^0 + uy^1, v\bar{z}\right] - f(y^0, 0) \tag{4.21}$$

para $0 \leq u \leq 1$ y $v \geq 0$, para todo $\bar{z} \geq 0$ tal que $\bar{z}_i^0 > 0$. Esta función es esencialmente $f(x)$ con el rango de variación de x restricto para un subconjunto convexo del octante no negativo, $\phi(u, v)$ es cuasi-convexa. Entonces tenemos

$$\phi(0,0) = f(y^0, 0) - f(y^0, 0) = 0 \tag{4.22}$$

$$\frac{\partial \phi(0,0)}{\partial u} = \nabla f(y^0)^T (y^1 - y^0) = 0 \tag{4.23}$$

y

$$\frac{\partial \phi(0,0)}{\partial v} = \nabla f(z^0)^T \bar{z} < 0 \tag{4.24}$$

Debemos probar $\phi(1,0) \geq 0$ o contradecir $\phi(1,0) > 0$. Primero, si para algún $\bar{u} > 0$ establecemos una pequeï¿½a vecindad de cero, $\phi(u,0)$ es positivo, cero o negativo (pero no mas de una de las condiciones). Entonces, mostraremos que $\phi(u,0) = 0$ y $\phi(u,0) > 0$ en una vecindad de cero es incompatible con $\phi(1,0) < 0$ mientras que $\phi(u,0) < 0$ contradice la hipótesis del teorema.

Consideremos, para algún $\bar{u} > 0$, $\phi(\bar{u}, 0) \leq 0$ entonces por cuasi-convexidad $\phi(u, 0) \leq 0$ para todo u tal que $0 \leq u \leq \bar{u}$. Además, $\phi(\bar{u}, v) \leq 0$ entonces por cuasi-convexidad y $\phi(u,0) \leq 0$ para todo u tal que $0 \leq u \leq \bar{u}$. De esta manera,

$$\phi(\lambda(0,v) + (1-\lambda)[u(v), 0]) \leq \max\{\phi(0,v), \phi[u(v), 0]\} = \phi(0,v), \ \forall\ \lambda \in [0,1]$$

luego,

$$\phi((1-\lambda)u(v), \lambda v) \leq \phi(0, v), \ \forall\ \lambda \in [0,1]. \tag{4.25}$$

Por cuasi-convexidad de ϕ,

$$((1-\lambda)\,u\,(v)\,,-(1-\lambda)\,v)^{T}\,\nabla\phi\,(0,v)\leq 0 \tag{4.26}$$

lo cual es equivalente a,

$$((1-\lambda)\,u\,(v)\,,-(1-\lambda)\,v)^{T}\left(\frac{\partial\phi}{\partial u}(0,v)\,,\frac{\partial\phi}{\partial v}(0,v)\right)\leq 0 \tag{4.27}$$

desde que $(1-\lambda)\geq 0$ tenemos

$$(u\,(v)\,,-v)^{T}\left(\frac{\partial\phi}{\partial u}(0,v)\,,\frac{\partial\phi}{\partial v}(0,v)\right)\leq 0 \tag{4.28}$$

Por tanto,

$$u\,(v)\,\frac{\partial\phi}{\partial u}(0,v)-v\frac{\partial\phi}{\partial v}(0,v)\leq 0. \tag{4.29}$$

Esto puede escribirse como

$$\frac{u\,(v)}{v}\frac{\partial\phi}{\partial u}(0,v)\leq\frac{\partial\phi}{\partial v}(0,v)\,. \tag{4.30}$$

Tomando límite cuando v tiende a cero

$$\lim_{v\to 0}\frac{u\,(v)}{v}\frac{\partial\phi}{\partial u}(0,v)\leq\lim_{v\to 0}\frac{\partial\phi}{\partial v}(0,v) \tag{4.31}$$

desde que, $\frac{\partial\phi}{\partial v}$ es continua

$$\lim_{v\to 0}\frac{u\,(v)}{v}\frac{\partial\phi}{\partial u}(0,v)\leq\frac{\partial\phi}{\partial v}(0,0) \tag{4.32}$$

De (4.24) tenemos

$$\lim_{v\to 0}\frac{u\,(v)}{v}\frac{\partial\phi}{\partial u}(0,v)\leq\frac{\partial\phi}{\partial v}(0,0)<0 \tag{4.33}$$

Además,

$$\lim_{v\to 0}\frac{1}{v}\left(\frac{\partial\phi}{\partial u}(0,v)-\frac{\partial\phi}{\partial u}(0,0)\right)=\frac{\partial^{2}\phi}{\partial v\partial u}(0,0)\,. \tag{4.34}$$

Como $\frac{\partial\phi}{\partial u}(0,0)=0$ entonces

$$\lim_{v\to 0}\frac{1}{v}\left(\frac{\partial\phi}{\partial u}(0,v)-\frac{\partial\phi}{\partial u}(0,0)\right)=\lim_{v\to 0}\frac{1}{v}\frac{\partial\phi}{\partial u}(0,v) \tag{4.35}$$

De (4.34) y (4.35) en (4.33) tenemos

$$\frac{\partial^{2}\phi}{\partial v\partial u}\lim_{v\to 0}u\,(v)<0 \tag{4.36}$$

Por tanto, de (4.24) y (4.36) tenemos

$$0<0,$$

lo cual es una contradicción. Por lo tanto, no es cierto que $\phi\,(u,0)<0$ para $u>0$ suficientemente pequeï¿$\frac{1}{2}$o.

d. La función es convexa

$$f(x_0) + (x_1 - x_0)^T \nabla f(x_0) \leq f(x_1)$$

y desde que

$$(x_1 - x_0)^T \nabla f(x_0) > 0$$

entonces

$$f(x_0) \leq f(x_1), \quad \forall\, x_1 \geq 0.$$

Lo cual demuestra el teorema.

A partir de la prueba, presentamos un ejemplo para una función $f(x,y)$ con

$$\frac{\partial f}{\partial x}(0,0) = 0 \ \wedge\ \frac{\partial f}{\partial y}(0,0) > 0$$

con $f(x,0)$ positivo en una vecindad lateral derecha.
El ejemplo será elegido considerando $f(0,y) = y$ tenemos

$$\frac{\partial f}{\partial x}(x,0) = \begin{cases} \frac{-1}{logx} & \text{para } x \leq \frac{1}{2} \\ \frac{1}{log2} & \text{para } x > \frac{1}{2} \end{cases}$$

Considerando $f(x,y)$ en los dos ejes, complementamos la definición haciendo que las curvas de nivel sean rectas, lo cual asegura la cuasi-concavidad de f. Formalmente, para todo valor fijo de $f(x,y) = z$ definimos $x = X(z)$ como solución de la ecuación $f(x,0) = z$.
Entonces la curva de nivel $f(x,y) = z$ intercepta al eje x en $x = X(z)$ y al eje y en $y = z$.
Si la curva de nivel es una línea recta y $f(x,y)$ es todo punto en esta tenemos

$$\frac{x}{X(z)} + \frac{y}{z} = 1. \tag{4.37}$$

Dados x, y fijos, entonces $f(x,y)$ es el único valor positivo de z para el cual (4.37) se satisface (excepto si $f(x,y) = 0$ para $x = y = 0$). Desde que

$$\frac{\partial f}{\partial x}(0,0) = 0 \ \wedge\ \frac{\partial f}{\partial y}(0,0) = 1$$

se satisface las condiciones de KKT para la restricción $-y \geq 0$ con $\lambda_0 = 1$. Pero el origen no es la restricción máxima.

4.2 Modelos Económicos Cuasi-convexos

La microeconomía se estudia de forma matemática, usando modelos matemáticos. La mayor parte de los desarrollos y estudios de la teoría del consumidor tienen como base el siguiente problema de maximizar, esto es, obtener el valor máximo de una función, el más alto de todos los que puede dar, así como qué valores son los que producen ese máximo. En este caso sería el de u, que es la función de utilidad de un consumidor, que se supone que depende de los valores de las cantidades de los n bienes (representados por las variables x_1 hasta x_n).

4.2.1 Demanda del Consumidor

La propiedad fundamental de la función de utilidad en la teoría de demanda del consumidor consiste en que las curvas de indiferencia definen conjuntos convexos o una disminución porcentual de sustitución. Así, la propiedad de toda función de utilidad es la cuasi-concavidad. El estudio de la teoría de demanda del consumidor así, como el Axioma Débil de Preferencia Relevada.

Consideremos el problema de maximizar la función de utilidad $u(x)$

$$\begin{array}{rcl} \max & u(x) & \\ s.a & & \\ p^T x & \leq & B \\ x & \geq & 0 \end{array}$$

donde p es el vector de precios, x es el vector de unidades y B es el presupuesto del consumidor. Este problema de maximización se puede expresar como el siguiente problema de minimización

$$\begin{array}{rcl} \min & -u(x) & \\ s.a & & \\ p^T x & \leq & B \\ x & \geq & 0 \end{array}$$

Supongamos que x^0 satisface las condiciones KKT, es decir,

$$\begin{aligned} -\frac{\partial u}{\partial x_i}\left(x^0\right) + \lambda_i^0 \frac{\partial p}{\partial x_i}\left(x^0\right) &\geq 0 \\ x_i^0 \left(-\frac{\partial u}{\partial x_i} + \lambda_i^0 \frac{\partial p}{\partial x_i}\right) &= 0 \\ \lambda_i^0 \left(B - \sum_{i=1}^{n} x_i^0 p_i\right) &= 0 \end{aligned}$$

Además, asumamos la no satisfacción, esto es,

$$\frac{\partial u\left(x^0\right)}{\partial x_i} > 0$$

para algún x_i. Entonces, desde que se satisface la condición b del teorema 4.1 tenemos que x^0 minimiza $-u(x)$ sujeto a las restricciones $B - \sum_{i=1}^{n} x_i p_i \geq 0$ y $x_i \geq 0$. Además, si $\lambda_i > 0$, y la hipótesis de no satisfacción implica que x^0 minimiza el costo en $u\left(x^0\right)$, para esto

$$\begin{array}{rcl} \max & -\sum_{i=1}^{n} x_i p_i & \\ & s.a & \\ u(x) - u\left(x^0\right) & \geq & 0 \\ x & \geq & 0 \end{array}$$

Por otro lado, es importante considerar que las empresas o cualesquiera otras unidades que en una economía se encarguen de producir bienes y servicios, comparte semejanzas desde cierto punto de vista con el del consumidor. En el caso del consumidor, la microeconomía lo reduce al problema de maximizar una función de utilidad, la cual es cuasi-cóncava, con una restricción presupuestaria. En el caso de la del productor, se trata de maximizar la función de beneficios teniendo en cuenta restricciones tecnológicas.

4.2.2 Producción

La teoría de la producción eficiente puede extenderse para funciones de producción que son cuasi-cóncavas pero no cóncavas, esto es para los casos donde existe un ingreso creciente pero un porcentaje de disminución marginal de sustitución.
Supongamos, por ejemplo, que una empresa alcanza una producción en un conjunto de procesos independientes los cuales transforman los insumos commprados en bienes intermedios los cuales no son tratados en el mercado, y ambos en producto.
Consideremos que la escala o intensidad del i-ésimo proceso es medida por x_i. Además, consideremos que el j-ésimo insumo o producto en el i-ésimo proceso es una función monótona $g_{ij}(x_i)$ la cual es positiva si el bien es un producto del proceso, negativa si es un insumo.
Enumeramos el producto final por $j = 1, ..., m_1$, los insumos comprados por $j = m_1 + 1, ..., m_2$, los bienes intermedios por $j = m_2 + 1, ..., m$ y consideremos n procesos.
Entonces el conjunto de insumos o productos del j-ésima bien está dada por

$$g_j(x) = \sum_{i=1}^{n} g_{ij}(x_i)$$

Ahora, consideremos el problema de minimizar el costo de producción en un conjunto de precios de productos dados por el precio de los insumos.
Sea p_j el precio del j-ésimo bien. Entonces el problema es

$$\max \sum_{j=m_1+1}^{m_2} p_j g_j(x)$$

$$s.a$$

$$g_j(x) - g_j\left(x^0\right) \geq 0$$

para $j = 1, ..., m_1$ y la restricción que el producto del bien intermedio es no negativo, o si consideremos ϵ_j representa el stock inicial, esto es el conjunto de consumo del bien intermedio no excede al stock inicial, es decir

$$g_j(x) + \epsilon_j \geq 0 \quad j = m_2 + 1, ..., m$$

ï¿½Bajo que condiciones este problema satisface la hipótesis del Teorema?
Sabemos que toda función monótona de una variable es cuasi-cóncava. Pero, aquí hemos encontrado una diferencia entre concavidad y cuasi-concavidad la cual es importante para las aplicaciones de la teoría económica.

Es evidente, que la combinación lineal no negativa de funciones cóncavas es cóncava, mientras la combinación lineal no negativa de funciones cuasi-cóncavas no necesariamente es cuasi-cóncava.
Como consecuencia, la hipótesis de cuasi-concavidad no puede ser reemplazada por la hipótesis de concavidad en muchas partes de la economía.
Consideremos uno de los productos restricción

$$g_j(x) - g_j(x^0) \geq 0$$

lo que es equivalente a

$$\sum_{j=1}^{n} g_{ij}(x_i) - g_j(x^0) \geq 0.$$

Para productos, tenemos $g'_{ij}(x_i) \geq 0$. Si $g''_{ij}(x_i) \leq 0$ es cóncava y por lo tanto, $g_j(x_i)$ es cuasi-cóncava.
Si $g''_{ij}(x_i) > 0$ para algún proceso, con $g''_{ij}(x_i) \leq 0$ para todos los otros procesos, entonces $g_j(x_i)$ no es necesariamente cuasi-cóncava.
Si $g_{ij}(x_i) > 0$ para dos o mas actividades $g_j(x)$ no puede ser cuasi-cóncava. Si $g_j(x)$ es cuasi-cóncava, el porcentaje marginal de sustitución entre todo par de insumos debe disminuir entre otros insumos que se mantienen constantes. Esto es, por ejemplo, $x_3, ..., x_n$ constantes

$$\left(g'_{2j}\right)^2 g''_{ij} + \left(g'_{1j}\right)^2 g''_{2j} \leq 0.$$

De esta manera, es posible que g''_{1j} ó g''_{2j} es positivo sin violar la condición anterior, pero claramente ambos no pueden ser positivos, similarmente, para todos otros pares de procesos.
Por lo tanto, $g''_{ij}(x_i) > 0$ para al menos un proceso i si $g_j(x_i)$ es cuasi-cóncava. Notemos que el mismo se satisface para las restricciones en el uso de bienes intermedios.
ï¿½Que pasa con el maximando $\sum_{j=m_1+1}^{n} p_j g_j(x)$?
Si las funciones $g_j(x)$ son cóncavas, su combinación lineal también es cóncava. Pero, si las funciones son cuasi-cóncavas y no cóncavas, no se puede garantizar la cuasi-concavidad de

$$\sum_{j=m_1+1}^{m} p_j g_j(x).$$

independientemente de los precios. De esta manera, la única forma para aplicar el teorema para todo conjunto de precios haciendo una disminución o un ingreso constante en el uso de productos medidos en terminos monetarios.
Por otro lado, para aplicar el Teorema y maximizar la ganancia

$$\max \sum_{j=1}^{m_2} p_j g_j(x)$$

$$s.a$$

$$g_j(x) + \epsilon_j \geq 0$$

para $j = m_2 + 1, ..., m$. Podemos considerar una cantidad total limitada de porcentaje creciente considerando bienes intermedios, pero no, en general, consideramos productos o insumos comprados en el mercado (al menos uno es producto y no insumo comprado).
Nuevamente, para todo conjunto de precios una cierta cantidad de porcentaje creciente en productos o compra de insumos medido en dinero puede ser tolerado.
Alternativamente, sea la función de producción

$$y = K^{\alpha}L^{\beta} \ \alpha > 0, \beta > 0$$

Esta función es cuasi-cóncava pero no cóncava para $\alpha + \beta > 1$. Entonces, el teorema sera aplicado para el problema de determinación de la eficiente combinación de insumos, dado algun producto especificado, pero no es aplicado para el problema de maximización de la ganancia. Esto es, el problema de $\min rK + wL$ o de $\max -(rK + wL)$ donde r y w son el costo de unidad de K y L respectivamente, sujeto a las restricciones $Y - Y^0 \geq 0, L \geq 0, K \geq 0$ satisface la hipótesis del teorema. Pero el problema de

$$\begin{aligned} \max \prod (K, L) &= pK^{\alpha}L^{\beta} - rK - wL \\ s.a & \\ K &\geq 0 \\ L &\geq 0 \end{aligned}$$

no satisface la hipótesis del teorema porque $\prod(K, L)$ no es cuasi-cóncava.

4.2.3 Bienestar Económico

Supongamos que toda función de producción sobre toda sociedad es cuasi-cóncava. El problema de encontrar una eficiente distribución de recursos (un óptimo de Pareto) puede ser formulado como un problema de maximización de la utilidad de una familia sujeto a las restricciones (también cuasi-cóncavas) de productos que estan dentro de las posibilidades de la producción de la sociedad y las utilidades de las otras familias son al menos iguales para niveles específicos. ∎

Capítulo 5

Un Método Proximal para Minimizar Funciones Cuasi-convexas

5.1 Motivación del Sistema Dinámico

Considere el problema de minimización convexa

$$(P) \quad \begin{cases} \min f(x) \\ x \in R^n_+ \end{cases}$$

donde $f : R^n \to (-\infty, \infty]$ es una función convexa propia y cerrada. Además,

$$R^n_+ = \{x \in R^n : x_j \geq 0, \ j = 1, ..., n\}.$$

Para resolver (P) consideramos el siguiente esquema proximal iterativo.
Iniciamos con

$$x_0 \in R^n_{++} = \{x \in R^n : x_j > 0, \ j = 1, ..., n\}$$

y generamos la sucesión $\{x_k\}$ por

$$x_k = \arg\min \left\{ f(x) + \lambda_k^{-1} d(x, x_{k-1}) \right\} \tag{5.1}$$

donde λ_k es una sucesión de números positivos y $d : R^n_{++} \times R^n_{++} \to R$ es una función distancia satisfaciendo las siguientes propiedades:

1. $d(x, y) \geq 0, \ \forall x, y \in R^n_{++}$ y $d(x, y) = 0$ si y solo si $x = y$;
2. $x \to d(x, y)$ es convexa $\forall y \in R^n_{++}$;
3. d satisface propiedades de diferenciabilidad cual veremos más adelante.

La función distancia, la cual será base de nuestro estudio del sistema dinámico propuesto (S) dependerá de una regularización logarítmica. En efecto, dado dos parámetros positivos $\mu > 0$ y $\nu > 0$ definimos para todo $(x,y) \in R^n_{++} \times R^n_{++}$ la siguiente función:

$$d(x,y) = \frac{\nu}{2} \|x-y\|^2 + \mu \sum_{j=1}^{n} y_j \varphi \left(y_j^{-1} x_j\right), \tag{5.2}$$

donde para todo $r > 0$ la función $\varphi(r)$ está dada por $\varphi(r) = r - \log r - 1$.
Desde que la función $h(x) = \log x$ es una función cóncava entonces se satisface

$$h(y) \leq h(x) + h'(x)(y-x), \ \forall x, y \in R_+$$

lo cual es equivalente a

$$\log y \leq \log x + \frac{1}{x}(y-x), \ \forall x, y \in R_+$$

Considerando $x = 1$ tenemos,

$$0 \leq y - \log y - 1.$$

A partir de este resultado obtenemos que $\varphi\left(\frac{x_j}{y_j}\right) \geq 0$, desde que $x_j, y_j \in R_{++}$.
Como los parámetros $\mu > 0$, $\nu > 0$ y $\varphi\left(\frac{x_j}{y_j}\right) \geq 0$ entonces se satisface la primera condición de la función distancia, es decir, $d(x,y) \geq 0$ y $d(x,y) = 0$ si y solamente si $x = y$.
Observamos también que $d(x,y)$ satisface la segunda condición de la función distancia definida en (5.2).
Debemos probar que

$$d(\lambda x_1 + (1-\lambda) x_2, y) \leq \lambda d(x_1, y) + (1-\lambda) d(x_2, y)$$

Primero, notemos que

$$\|\lambda x_1 + (1-\lambda) x_2 - y\|^2 \leq \lambda \|x_1 - y\|^2 + (1-\lambda) \|x_2 - y\|^2.$$

Desde que $\|x_1 - x_2\|^2 \geq 0$ entonces

$$\langle x_1, x_1 \rangle - 2 \langle x_1, x_2 \rangle + \langle x_2, x_2 \rangle \geq 0$$

multiplicamos por $\lambda(\lambda - 1) < 0$

$$\langle \lambda x_1, \lambda x_1 \rangle - \langle \lambda x_1, x_1 \rangle + 2 \langle \lambda x_1, (1-\lambda) x_2 \rangle + \lambda (1-\lambda) \langle x_2, x_2 \rangle \leq 0$$

lo cual es equivalente a

$$\langle \lambda x_1, \lambda x_1 + (1-\lambda) x_2 \rangle + \langle \lambda x_1, (1-\lambda) x_2 \rangle - \langle \lambda x_2, (1-\lambda) x_2 \rangle \leq \langle \lambda x_1, x_1 \rangle$$

sumando $\langle x_2, (1-\lambda) x_2 \rangle - \langle \lambda x_1, y \rangle$ en ambos miembros

$$\langle \lambda x_1, \lambda x_1 + (1-\lambda) x_2 - y \rangle + \langle \lambda x_1 + (1-\lambda) x_2, (1-\lambda) x_2 \rangle \leq \lambda \langle x_1, x_1 - y \rangle + (1-\lambda) \langle x_2, x_2 \rangle$$

luego, adicionamos $-\langle(1-\lambda)x_2, y\rangle + \langle -y, \lambda x_1 + (1-\lambda)x_2 - y\rangle$ en ambos miembros

$$\langle \lambda x_1 + (1-\lambda)x_2 - y, \lambda x_1 + (1-\lambda)x_2 - y\rangle \leq \lambda\langle x_1 - y, x_1 - y\rangle + (1-\lambda)\langle x_2 - y, x_2 - y\rangle$$

desde que $-y = -\lambda y - (1-\lambda)y$, entonces

$$\|\lambda x_1 + (1-\lambda)x_2 - y\|^2 \leq \lambda\|x_1 - y\|^2 + (1-\lambda)\|x_2 - y\|^2. \tag{5.3}$$

Por otro lado, desde que $\log x$ es una función cóncava y creciente obtenemos

$$\varphi(\lambda r_1 + (1-\lambda)r_2) < \lambda\varphi(r_1) + (1-\lambda)\varphi(r_2), \forall \ \lambda \in [0,1]. \tag{5.4}$$

De (5.3) y (5.4)

$$d(\lambda x_1 + (1-\lambda)x_2, y) \leq \lambda d(x_1, y) + (1-\lambda)d(x_2, y),$$

donde

$$d(\lambda x_1 + (1-\lambda)x_2, y) = \frac{\nu}{2}\|\lambda x_1 + (1-\lambda)x_2 - y\|^2 + \mu\sum_{j=1}^{n}\varphi\left(\frac{\lambda x_{1_j} + (1-\lambda)x_{2_j}}{y_j}\right).$$

Además, como $\varphi(r) = r - \log r - 1$ es una función estrictamente convexa la cual es no negativa en R_{++} entonces, para todo $r_1, r_2 \in R_{++}$

$$\varphi(\lambda r_1 + (1-\lambda)r_2) < \lambda\varphi(r_1) + (1-\lambda)\varphi(r_2), \ \forall\lambda \in (0,1), r_1 \neq r_2.$$

Desde que $\log x$ es una función cóncava tenemos

$$\log(\lambda r_1 + (1-\lambda)r_2) \geq \lambda\log r_1 + (1-\lambda)\log r_2, \ \forall\lambda \in (0,1).$$

Multiplicando ambos miembros por -1

$$-\log(\lambda r_1 + (1-\lambda)r_2) \leq -\lambda\log r_1 - (1-\lambda)\log r_2.$$

Sumando $\lambda r_1 + (1-\lambda)r_2$ en ambos miembros obtenemos

$$\varphi(\lambda r_1 + (1-\lambda)r_2) \leq \lambda\varphi(r_1) + (1-\lambda)\varphi(r_2).$$

Supongamos que se cumple la igualdad entonces

$$\varphi(\lambda r_1 + (1-\lambda)r_2) = \lambda\varphi(r_1) + (1-\lambda)\varphi(r_2).$$

lo cual es equivalente a

$$\lambda r_1 + (1-\lambda)r_2 - \log(\lambda r_1 + (1-\lambda)r_2) - 1 =$$
$$\lambda r_1 - \lambda\log r_1 - \lambda + (1-\lambda)r_2 - (1-\lambda)\log r_2 - (1-\lambda).$$

Sumando $-\lambda r_1 - (1-\lambda)r_2$ en ambos miembros

$$-\log(\lambda r_1 + (1-\lambda)r_2) = -\log r_1^{\lambda} - \log r_2^{1-\lambda}, \ \forall\lambda \in (0,1).$$

Lo cual se reduce a

$$\lambda r_1 + (1-\lambda) r_2 = r_1^{\lambda}.r_2^{1-\lambda}, \ \forall \lambda \in (0,1).$$

De donde se deduce que la igualdad se cumple si y solo si $r_1 = r_2$. Por tanto, φ es una función estrictamente convexa.
Desde que φ es una función estrictamente convexa la cual es no negativa en R_{++}, alcanza su único mínimo en $r = 1$, es decir,

$$\varphi(1) = \min \varphi = 0 = \varphi'(1)$$

desde que $\varphi'(r) = 1 - \frac{1}{r}$ y que

$$\lim_{r\to 0^+} \varphi(r) = \lim_{r\to +\infty} \varphi(r) = +\infty$$

En efecto, sabemos que $-\varphi(r) = \log r + 1 - r$ entonces

$$\lim_{r\to\infty} 10^{-\varphi(r)} = \lim_{r\to\infty} 10^{\log r + 1 - r}$$

Esta expresión se reduce, lo cual se vuelve un límite indeterminado $\frac{\infty}{\infty}$

$$= \lim_{r\to\infty} 10^{\log r}.10^{-(r-1)}$$

Aplicando L' Hospital, obtenemos

$$\lim_{r\to\infty} \frac{1}{10^{r-1}\log 10} = 0$$

Lo cual es equivalente a

$$\lim_{r\to\infty} 10^{-\varphi(r)} = 0 \Rightarrow \lim_{r\to\infty} -\varphi = -\infty$$

Por lo tanto, $\lim_{r\to\infty} \varphi(r) = \infty$. De estos resultados, deducimos inmediatamente que d satisface las propiedades antes mencionadas. La iteración

$$x_k = \arg\min \left\{ f(x) + \lambda_k^{-1} d(x, x_{k-1}) \right\}$$

es el método proximal cuando $d(x, x_{k-1})$ es formalmente tomado como el usual termino cuadrático $\frac{1}{2}\|x - x_{k-1}\|^2$ es decir, con $\mu = 0$, $\nu > 0$. La motivación para estudiar los métodos de esta forma, con d definida como en (5.2) con $\nu = 0$ y $\mu > 0$, puede ser encontrada en recientes estudios; ver [4]. La diferencia fundamental de estos métodos, en comparación con el esquema del método clásico, es que el término d es usado para garantizar que las iteraciones $\{x_k\}$ se encuentren en el interior del octante no negativo R_{++}^n, por eso el algoritmo genera automaticamente una sucesión positiva $\{x_k\}$.
Nuestra motivación es el funcional logaritmo-cuadrático, motivado por el reciente trabajo de Auslender, Teboulle y Ben-Tiba [4], quienes estudiaron métodos pero con homogeneidad de segundo orden usando la siguiente función distancia

$$D(x,y) = \frac{\nu}{2}\|x-y\|^2 + \mu \sum_{j=1}^{n} y_j^2 \varphi\left(y_j^2 x_j\right).$$

Por otro lado, notemos que para minimizar una función convexa en R^n, el esquema clásico proximal cuadrático genera una sucesión $\{x_k\} \in R^n$ para solucionar la inclusión

$$0 \in \lambda_k^{-1}\left(x^k - x^{k-1}\right) + \partial f\left(x^k\right),\ \lambda_k > 0$$

donde ∂f denota la subdiferencial de la función convexa f, ver [25].
En efecto, desde que la función f es propia y $x_k \in \arg\min\left\{f(x) + \lambda_k^{-1} d\left(x, x^{k-1}\right)\right\}$, entonces

$$0 \in \partial\left\{f(x) + \lambda_k^{-1} d\left(x, x^{k-1}\right)\right\}(x_k)$$

Lo cual es equivalente a

$$0 \in \partial f(x_k) + \lambda_k^{-1}\left(x^k - x^{k-1}\right)$$

El esquema iterativo visto anteriormente puede ser visto también como la discretización implícita de Euler de la subdiferencial inclusión

$$0 \in x'(t) + \partial f(x(t))$$

con $x : [0, \infty) \to R^n$. Supongamos que tenemos que resolver (P) por el método proximal con d definida en (5.2).
Entonces, escribimos las condiciones de optimalidad con d definida en (5.2)

$$x_k \in \arg\min\left\{f(x) + \lambda_k^{-1}\left(\frac{\nu}{2}\|x - x_{k-1}\|^2 + \mu\sum_{j=1}^{n} x_{k-1_j}\varphi\left(x_{k-1_j}^{-1} x_j\right)\right)\right\},$$

A partir de esto obtenemos

$$0 \in \partial\left\{f(x) + \lambda_k^{-1}\left(\frac{\nu}{2}\|x - x_{k-1}\|^2 + \mu\sum_{j=1}^{n} x_{k-1_j}\varphi\left(x_{k-1_j}^{-1} x_j\right)\right)\right\}(x_k),$$

Lo cual es equivalente a

$$0 \in g_j^k + \lambda_k^{-1}\left[\nu\left(x_{k_j} - x_{k-1_j}\right) + \mu x_{k-1_j}\varphi'\left(\frac{x_{k_j}}{x_{k-1_j}}\right)\right],$$

luego,

$$0 \in g_j^k + \lambda_k^{-1}\left[\nu\left(x_{k_j} - x_{k-1_j}\right) + \mu x_{k-1_j}\left(\frac{1}{x_{k-1_j}} - \frac{x_{k-1_j}}{x_{k_j}}\right)\right],$$

de ahí,

$$0 \in g_j^k + \lambda_k^{-1}\left[\nu\left(x_j^k - x_j^{k-1}\right) + \mu\left(1 - \frac{x_j^{k-1}}{x_j^k}\right)\right],$$

donde $g^k = \left(g_1^k, ..., g_n^k\right) \in \partial f\left(x^k\right)$.
La inclusión anterior puede ser escrita equivalentemente como

$$0 \in g_j^k + \lambda_k^{-1}\left[-x_j^{k-1}\left(\nu + \frac{\mu}{x_j^k}\right) + x_j^k\left(\nu + \frac{\mu}{x_j^k}\right)\right] \quad j = 1, ..., n,$$

lo cual implica que,

$$0 \in g_j^k + \lambda_k^{-1}\left(\nu + \frac{\mu}{x_j^k}\right)\left(x_j^k - x_j^{k-1}\right) \quad j = 1, ..., n,$$

entonces

$$0 \in \lambda_k^{-1}\left(x_j^k - x_j^{k-1}\right) + \frac{x_j^k}{\mu + \nu x_j^k} g_j^k \quad j = 1, ..., n.$$

Esta expresión puede ser vista como la discretización implícita de Euler del Sistema Dinámico Continuo.

$$0 \in x_j'(t) + \left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right] g_j(x(t)) \quad j = 1, ..., n$$

la cual es base del estudio del Sistema Dinámico.

5.2 Existencia Global y Viabilidad de Resultados

Adaptamos aquí las mismas hipótesis generales de la sección anterior para la función f, es decir, la función $f : R^n \to R$ pertenece a $C^2(R^n)$. El caso convexo será tratado en la siguiente sección. Estamos interesados en estudiar el problema de Cauchy descrito por el siguiente sistema (S).

$$x_j'(t) + \left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right] \frac{\partial f(x(t))}{\partial x_j} = 0, \quad j = 1, ..., n \tag{5.5}$$

$$x_j(0) = x_{0_j} \in R_{++}^n, \quad j = 1, ..., n \tag{5.6}$$

Es importante destacar en el análisis del sistema, la propiedad de viabilidad de las trayectorias las cuales se deben encontrar en el octante positivo;

$$x(t) \in R_+^n, \quad \forall t \in [0, \infty).$$

Nuestro principal resultado en esta sección es la prueba de la existencia global de la trayectoria y la propiedad de viabilidad del conjunto R_+^n para el sistema dinámico (S).

Teorema 5.1 *Sea $f : R^n \to R$ una función en $C^2(R^n)$ la cual esta acotada inferiormente en R_+^n, existe una única solución global $x : [0, \infty) \to R^n$ del sistema (S). Además, $x(.)$ satisface la siguiente propiedad de viabilidad:*

$$x_j(t) > 0 \;\forall t \geq 0, \;\forall j = 1, ..., n.$$

Demostración. Consideremos $F : R^n_+ \to R^n$ una aplicación definida por

$$F(x) = \left(\frac{x_1}{\mu + \nu x_1} \frac{\partial f}{\partial x_1}, ..., \frac{x_n}{\mu + \nu x_n} \frac{\partial f}{\partial x_n} \right)$$

Desde que $x_0 \in R^n_{++}$ y $\mu > 0$, trabajando en el sistema (S) obtenemos

$$\frac{x_j(t)}{\nu x_j(t) + \mu} \frac{\partial f(x(t))}{\partial x_j} = -x_j'(t)$$

Entonces $F(x(t)) = (-x_1'(t), ..., -x_n'(t))$.
Probaremos que F es una función convexa, consideremos $x, y \in R^n_+$

$$\begin{aligned} F(\lambda x + (1-\lambda) y) &= (-\lambda x_1'(t) - (1-\lambda) y_1'(t), ..., -\lambda x_n'(t) - (1-\lambda) y_n'(t)) \\ &= \lambda(-x_1'(t), ..., -x_n'(t)) + (1-\lambda)(-y_1'(t), ..., -y_n'(t)) \\ &= \lambda F(x) + (1-\lambda) F(y), \quad \forall \lambda \in [0,1] \end{aligned}$$

entonces es localmente lipchitziana en x_0, es decir, F es lipschitz continua en una vecindad de x_0. Por lo tanto, por el teorema de Cauchy-Lipschitz, existe $T > 0$ tal que el sistema admite una única solución en $[0, T]$.
Sea $t \to x(t)$ la solución y se prueba que

$$x_j > 0 \quad \forall t \in [0, T], \ \forall j = 1, ..., n.$$

Desde que $x_0 \in R^n_{++}$ y $x(0) = x_0$ por la continuidad de $x(.)$, como $x(.)$ es diferenciable, existe $\tau_0 > 0$ tal que

$$x_j(t) > 0 \quad \forall t \in [0, \tau_0], \ \forall j = 1, ..., n. \tag{5.7}$$

En efecto, por hipótesis de continuidad de $x(.)$ sobre $[0, \tau_0]$ aplicamos el teorema de Weirstrass entonces

$$x_j(0) \leq x_j(t) \leq x_j(\tau_0)$$

Por tanto, $x_{0_j} \leq x_j(t)$ y desde que $x_0 > 0$ entonces $x_{0_j} > 0$, para todo $j = 1, ..., n$.
A partir de aquí obtenemos la afirmación anterior.
Definimos

$$\begin{aligned} U &= \{\tau \in [0, T] : x_j(t) > 0, \forall t \in [0, \tau], \forall j = 1, ..., n\} \\ &= \left\{\tau \in [0, T] : x(t) \in R^n_{++}, \forall t \in [0, \tau]\right\}. \end{aligned}$$

De la afirmación anterior (5.7) se concluye que U es no vacío. Desde que R^n_{++} es abierto , por la continuidad de $x(.)$, U es abierto en $[0, T]$.

$$x : U \to R^n_{++}$$

En efecto, por continuidad de $x(.)$ la imagen inversa de R^n_{++} es abierto en $[0, T]$.
Podemos concluir que $U = [0, T]$
Por la definición de U tenemos que $U \subset [0, T]$. Probaremos que U es cerrado en $[0, T]$.

Sea $\tau_n \in U$, sin pérdida de generalidad asumimos que $\tau_n \leq \tau$, es decir, $0 \leq \tau_n \leq \tau$ tal que $x_j(t) > 0, \ \ \forall t \in [0, \tau_n]$.
Desde que $\tau_n \to \tau$ se tiene

$$0 \leq \tau \leq \tau \wedge x_j(t) > 0, \ \forall t \in [0, \tau].$$

entonces $\tau \in U$.
Primero, notamos que $x(t)$ es acotado en $[0,T]$ desde que $x(.)$ es continua y $[0,T]$ es compacto. Además, desde que $f \in C^2(R^n)$ entonces sus derivadas parciales son continuas.
Como $\frac{\partial f}{\partial x_j}$ es continua aplicada en $x(t)$ (conjunto compacto) para todo $t \in [0,T]$ entonces $\frac{\partial f(x(t))}{\partial x_j}$ es compacto, es decir, existe una constante positiva $C > 0$ tal que

$$\left|\frac{\partial f(x(t))}{\partial x_j}\right| \leq C, \ \forall t \in [0,T], \ \forall j = 1, ..., n. \tag{5.8}$$

Ahora analizamos el hecho que $x(.)$ resuelve (S) y satisface $x(t) \in R^n_{++}$ para todo $t \in [0, \tau_n]$.
De la desigualdad probada en (5.8) y usando el siguiente resultado

$$\frac{x_j(t)}{\mu + \nu x_j(t)} \geq 0 \tag{5.9}$$

en $[0, \tau_n]$. En el sistema (S) tenemos

$$0 = x_j'(t) + \frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j} \leq x_j'(t) + C\frac{x_j(t)}{\mu + \nu x_j(t)}, \qquad \forall t \in [0, \tau_n].$$

Por lo tanto,

$$0 \leq x_j' + C\left[\frac{x_j(t)}{\mu + \nu x_j(t)}\right], \ \forall t \in [0, \tau_n].$$

Usando nuevamente la propiedad que $x(t) \in R^n_{++}$ para todo $t \in [0, \tau_n]$ y $\mu > 0$ entonces obtenemos

$$0 \leq x_j'(t) + C\frac{x_j(t)}{\mu + \nu x_j(t)} \leq x_j'(t) + \frac{C}{\mu} x_j(t), \ \forall t \in [0, \tau_n]$$

desde que $\mu \leq \mu + \nu x_j(t)$, se cumple que $\frac{1}{\mu + \nu x_j(t)} \leq \frac{1}{\mu}$.
Multiplicando por $e^{\frac{Ct}{\mu}}$ en ambos miembros

$$0 \leq e^{\frac{Ct}{\mu}} x_j'(t) + \frac{C}{\mu} e^{\frac{Ct}{\mu}} x_j(t), \ \ \forall t \in [0, \tau_n].$$

Esta desigualdad puede ser escrita como

$$0 \leq \frac{d}{dt} e^{\frac{Ct}{\mu}} x_j(t), \ \forall t \in [0, \tau_n], \ \forall j = 1, ..., n$$

Integrando ambos miembros de 0 a t

$$\int_0^t 0 \leq \int_0^t \frac{d}{dt} e^{\frac{Ct}{\mu}} x_j(t), \ \forall t \in [0, \tau_n].$$

Luego,

$$\begin{aligned} 0 &\leq e^{\frac{Ct}{\mu}} x_j(t) - e^{\frac{C0}{\mu}} x_j(0) \\ &\leq e^{\frac{Ct}{\mu}} x_j(t) - x_{0_j} \end{aligned}$$

La cual es equivalente a

$$0 < e^{\frac{-Ct}{\mu}} x_{0_j} \leq x_j(t), \ \forall t \in [0, \tau_n], \ \forall j = 1, ..., n.$$

En consecuencia $x(t) \in R^n_{++}$, $\forall t \in [0, \tau]$ y $\tau \in U$.
Probaremos, que la única solución máxima $x(.)$ de (S) es definida en el intervalo $[0, \infty]$. Haremos la prueba por contradicción.
En efecto, supongamos que la solución máxima está definida en $[0, T_{max})$, es decir,

$$x : [0, T_{max}) \to R^n \ \text{ con } \ T_{max} < \infty.$$

Usando el resultado anterior, tenemos

$$x(t) \in R^n_{++} \ \text{ para todo } \ t \in [0, T_{max}).$$

Primero probaremos que $\lim_{t \to T_{max},\ t < T_{max}} x(t)$ existe. Para demostrar esto, probaremos que

$$\int_0^{T_{max}} \|x'(t)\|^2 \, dt < \infty.$$

De esta manera, obtenemos como una consecuencia el hecho que el sistema (S) es un método de descenso para la función f. Partimos con la identidad

$$\frac{df(x(t))}{dt} = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(x(t))\, x_j'(t) \tag{5.10}$$

y usando (S) para $t < T_{max}$ (recalcando $x_j(t) > 0$, $\forall t > 0$) obtenemos

$$-\frac{\mu + \nu x_j(t)}{x_j(t)} x_j'(t) = \frac{\partial f(x(t))}{\partial x_j}$$

Reemplazando la expresión anterior en (5.10)

$$\frac{d}{dt} f(x(t)) + \sum_{j=1}^{p} \left(\nu + \frac{\mu}{x_j(t)} \right) x_j'^2 = 0 \tag{5.11}$$

y de aquí se concluye

$$\frac{d}{dt}f(x(t)) + \nu \sum_{j=1}^{n} x_j'^2(t) \leq 0$$

Integrando de 0 a t a ambos miembros de la desigualdad donde $t < T_{max}$,

$$\int_0^t \frac{d}{dt}f(x(t)) + \nu \int_0^t \sum_{j=1}^{n} x_j'^2(t) \leq 0$$

entonces,

$$f(x(t)) - f(x_0) + \nu \int_0^t \|x'(s)\|^2 ds \leq 0.$$

Desde que f es asumida en R_+^n, deducimos que

$$\nu \int_0^t \|x'(s)\|^2 ds \leq f(x_0) - \inf_{R_+^n} f < \infty$$

La siguiente desigualdad se cumple para todo $t < T_{max}$ entonces

$$\int_0^{T_{max}} \|x'(t)\|^2 dt < \infty$$

lo cual implica que el $\lim_{t \to T_{max}} x(t)$ existe.
Retornando al sistema (S) desde que $\mu > 0$ y $x(t) \in R_{++}^n$ para todo $t \in [0, T_{max}]$

$$x(t) > 0, \ \forall t \in [0, T_{max}].$$

Analizando el límite cuando $t \to T_{max}$, desde que

$$x_j'(t) = -\frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j}.$$

Entonces

$$\lim_{t \to T_{max}} x_j'(t) = \lim_{t \to T_{max}} -\frac{x_j(t)}{\mu + \nu x_j(t)} . \lim_{t \to T_{max}} \frac{\partial f(x(t))}{\partial x_j}$$

El $\lim_{t \to T_{max}} \frac{x_j(t)}{\mu + \nu x_j(t)}$ existe desde que el $\lim_{t \to T_{max}} x_j(t)$ existe y μ, ν son positivos.
El $\lim_{t \to T_{max}} \frac{\partial f(x(t))}{\partial x_j}$ existe desde que $\frac{\partial f}{\partial x_j}$ es continua y $\lim_{t \to T_{max}} x(t)$ existe, es decir,

$$\lim_{t \to T_{max}} \frac{\partial f(x(t))}{\partial x_j} = \frac{\partial f}{\partial x_j}\left(\lim_{t \to T_{max}} x(t)\right)$$

Por lo tanto, $\lim_{t \to T_{max}} x'(t)$ existe. Entonces, $x(.)$ es una solución de (S) en $[0, T_{max})$.
En particular, se tiene $x(T_{max}) > 0$. El final de la prueba termina aplicando el teorema de Cauchy-Lipschitz con el dato $x(T_{max})$ lo cual contradice la maximilidad

de $x(.)$. ■

Es interesante notar que cuando partimos con un punto arbitrario $x_0 \in R^n_{++}$, la correspondiente trayectoria $x : [0, \infty) \to R^n$ permanece en R^n_{++}, es decir,

$$x(t) > 0, \; \forall t \geq 0.$$

Esta importante propiedad es consecuencia del termino de barrera logarítmica en la función distancia definida anteriormente en (5.2). Tal propiedad puede ser formulada como la propiedad de viabilidad del conjunto R^n_{++} para el sistema dinámico (S). Equivalentemente, decimos que R^n_{++} es un conjunto invariante por el sistema dinámico (S).

5.3 Convergencia Asintótica

Trabajaremos con las mismas hipótesis sobre la función f, que en la sección anterior.

Teorema 5.2 *Sea $f : R^n \to R$ una función de clase $C^2(R^n)$ la cual es acotada inferiormente en el octante no negativo y consideremos el sistema asociado (S). Toda solución global $x(.)$ de S satisface*

1. *$f(x(t))$ es una función decreciente en t.*
2. *$\int_0^\infty \|x'(t)\|^2 dt < \infty$*
   Además, si asumimos que $x(.)$ es acotada, entonces
3. *$\lim_{t\to\infty} x'(t) = 0$*
4. *$\lim_{t\to\infty} x_j(t) \frac{\partial f(x(t))}{\partial x_j} = 0 \; j = 1, ..., n$*

Demostración. Todas las propiedades establecidas seran derivadas gracias a que $f(x(t))$ es una función decreciente de t.
Recalcamos la ecuación obtenida por el sistema (S) en (5.11)

$$\frac{d}{dt} f(x(t)) + \sum_{j=1}^{n} \left(\nu + \frac{\mu}{x_j(t)}\right) x_j'^2(t) = 0 \tag{5.12}$$

desde que $\mu > 0, \nu > 0$ y $x(t) \in R^n_{++}$, tenemos

$$\frac{df(x(t))}{dt} < 0$$

luego, $f(x(t))$ es decreciente en t.
Por lo tanto, queda demostrado el item (a).
Ahora, integramos la identidad (5.12) de 0 a t

$$\int_0^t \frac{df(x(t))}{dt} + \int_0^t \sum_{j=1}^{p} \left(\nu + \frac{\mu}{x_j(t)}\right) x_j'^2(t) = 0$$

de ahí,

$$f(x(t)) - f(x(0)) + \sum_{j=1}^{n} \int_0^t \left(\nu + \frac{\mu}{x_j(t)}\right) x_j'^2(s)\,ds = 0$$

desde que $x(0) = x_0$, tenemos

$$f(x(t)) - f(x_0) + \sum_{j=1}^{n} \int_0^t \left(\nu + \frac{\mu}{x_j(s)}\right) x_j'^2(s)\,ds = 0$$

lo cual es equivalente a,

$$f(x(t)) - f(x_0) + \nu \int_0^t \sum_{j=1}^{n} x_j'^2(s)\,ds + \sum_{j=1}^{n} \int_0^t \frac{\mu}{x_j(s)} x_j'^2(s)\,ds = 0$$

entonces,

$$f(x(t)) - f(x_0) + \nu \int_0^t \|x'(s)\|^2\,ds + \mu \sum_{j=1}^{n} \frac{x_j'^2(s)}{x_j(s)} ds = 0$$

donde

$$\|x'(s)\| = \sqrt{\sum_{j=1}^{n} x_j'^2(s)}.$$

Por tanto,

$$\nu \int_0^t \|x'(s)\|^2\,ds = f(x_0) - f(x(t)) - \mu \sum_{j=1}^{n} \int_0^t \frac{x_j'^2(s)}{x_j(s)} ds$$

luego,

$$\begin{aligned}\nu \int_0^t \|x'(s)\|^2\,ds &\leq f(x_0) - \inf f - \mu \sum_{j=1}^{n} \int_0^t \frac{x_j'^2(s)}{x_j(s)} ds \\ &\leq f(x_0) - \inf_{R_+^n} f.\end{aligned}$$

La siguiente desigualdad es verdadera para $t \geq 0$ obteniendo

$$\int_0^\infty \|x'(t)\|^2\,dt \leq \frac{1}{\nu}\left[f(x_0) - \inf_{R_+^n} f\right] < \infty \tag{5.13}$$

es decir $x \in L^2([0,\infty), R^n)$.
Supongamos que $x(.)$ es acotada, es decir,

$$\sup_{t\in(0,\infty]} \|x(t)\| < \infty.$$

Desde que f es suave, $\frac{\partial f}{\partial x_j}$ continua y $x(.)$ acotada entonces

$$\frac{\partial f(x(.))}{\partial x_j}$$

es acotada. Por tanto,

$$\sup_{t\in[0,\infty)} \left|\frac{\partial f(x(t))}{\partial x_j}\right| < \infty \tag{5.14}$$

Desde que $x(t) \in R^p_{++}$ en el sistema (S) obtenemos

$$x_j'(t) = -\frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j},$$

aplicando valor absoluto en ambos miembros

$$\left|x_j'(t)\right| = \left|\frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j}\right| \tag{5.15}$$

Además, como $\nu x_j(t) \leq \nu x_j(t) + \mu$ entonces

$$\frac{\nu x_j(t)}{\nu x_j(t) + \mu} \leq 1,$$

lo cual es equivalente a

$$\frac{x_j(t)}{\nu x_j(t) + \mu} \leq \frac{1}{\nu}.$$

En (5.15) tenemos

$$\left|x_j'(t)\right| = \left|\frac{x_j(t)}{\mu + \nu x_j(t)}\right| \left|\frac{\partial f(x(t))}{\partial x_j}\right| \leq \frac{1}{\nu} \left|\frac{\partial f(x(t))}{\partial x_j}\right| \tag{5.16}$$

De (5.16) y (5.14), concluimos $x'(.)$ es acotada en $[0, \infty)$, es decir,

$$\sup_{t\in[0,\infty)} \|x'(t)\| < \infty.$$

Aplicando la diferencial del sistema (S) con respecto a t, tenemos

$$x_j^{\cdot\cdot}(t) + \theta(x_j(t)) \sum_{j=1}^{n} \frac{\partial^2 f(x(t))}{\partial x_j \partial x_k} x(t)\, x_k'(t) + \theta'(x_j(t))\, x_j'(t) \frac{\partial f(x(t))}{\partial x_j} = 0$$

donde $\theta : R_+ \to R_+$ está dada por

$$\theta(r) = \frac{r}{\mu + \nu r}.$$

Afirmamos que θ y θ' son acotados en R_+.
En efecto, desde que $\nu r < \nu r + \mu$ se tiene

$$\frac{r}{\nu r + \mu} < \frac{1}{\nu},$$

entonces

$$\theta(r) < \frac{1}{\nu}.$$

Analizando la derivada de $\theta(r) = \frac{r}{\mu+\nu r}$ obtenemos

$$\theta'(r) = \frac{\mu}{(\mu+\nu r)^2} \leq \frac{1}{\mu}$$

entonces $\theta'(r) \leq \frac{1}{\mu}$.
Como $f \in C^2$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$ y $\frac{\partial f}{\partial x_i}$ son continuas

$$x_j^{\cdot\cdot}(t) = -\theta(x_j(t)) \sum_{k=1}^{p} \frac{\partial^2 f(x(t))}{\partial x_j \partial x_k} x_k'(t) - \theta'(x_j(t))\, x_j'(t) \frac{\partial f(x(t))}{\partial x_j}$$

Aplicando $|.|$ en ambos miembros

$$\begin{aligned}
\left|x_j^{\cdot\cdot}(t)\right| &= \left|\theta(x_j(t)) \sum_{k=1}^{p} \frac{\partial^2 f(x(t))}{\partial x_j \partial x_k} x_k'(t) + \theta'(x_j(t))\, x_j'(t) \frac{\partial f(x(t))}{\partial x_j}\right| \\
&\leq |\theta(x_j(t))| \left|\sum_{k=1}^{n} \frac{\partial^2 f(x(t))}{\partial x_j \partial x_k} x_k'(t)\right| + \left|\theta'(x_j(t))\, x_j'(t)\right| \left|\frac{\partial f(x(t))}{\partial x_j}\right| \\
&\leq \frac{1}{\nu} \sum_{k=1}^{n} \left|\frac{\partial^2 f(x(t))}{\partial x_j \partial x_k}\right| |x_k'(t)| + \frac{1}{\mu} \left|x_j'(t)\right| \left|\frac{\partial f(x(t))}{\partial x_j}\right| < \infty
\end{aligned}$$

es decir,

$$\sup_{t \in [0,\infty)} \left|x_j^{\cdot\cdot}(t)\right| < \infty \tag{5.17}$$

desde que $\frac{\partial^2 f(x(t))}{\partial x_j \partial x_k}$ y $\frac{\partial f(x(t))}{\partial x_j}$ son acotadas.
De (5.13) y (5.17) obtenemos

$$\int_0^\infty \|x'(t)\|^2 dt < \infty \ \wedge \ \sup_{t \in [0,\infty)} \|x^{\cdot\cdot}(t)\| < \infty$$

entonces

$$\lim_{t\to\infty} \left\|x_j'(t)\right\| = 0. \tag{5.18}$$

Además,

$$\left|x_j'(t)\right| = \left|\frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j}\right|$$

aplicando límite cuando $t \to \infty$

$$\lim_{t\to\infty} \left|x_j'(t)\right| = \lim_{t\to\infty} \left|\frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j}\right|$$

de (5.18)

$$\lim_{t\to\infty} \left|x_j'(t)\right| = 0$$

entonces

$$\lim_{t\to\infty} \frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j} = 0.$$

Lo cual demuestra el item 4. ■

5.4 Convergencia Asintótica: Caso Convexo

Estamos interesados en analizar la convergencia asintótica cuando $t \to \infty$ de las trayectorias $\{x(t) : t \in [0,\infty)\}$ generada por el sistema dinámico (S) donde se asume que la función f es convexa.
Nuestro principal resultado en esta sección es probar que toda trayectoria generada por (S) converge a una solución óptima del problema de minimización convexa.

$$(P) \quad \inf\left\{f(x) : x \in R^n_+\right\}.$$

Denotemos por

$$X := \arg \min_{R^n_+} f$$

Consideremos las siguientes hipótesis

(H1) $f : R^n \to R$ es una función convexa en C^1.

(H2) Para cada $x_0 \in R^n_{++}$, existe una solución global $x : [0,\infty) \to R^n_{++}$ del sistema dinámico (S)

$$\begin{aligned} x_j(t) + \frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j} &= 0, \ \forall j = 1, ..., n \\ x_j(0) = x_{0_j} &\in R^n_{++} \ \forall j = 1, ..., n. \end{aligned}$$

Proposición 5.1 *Supongamos que se satisfacen las hipótesis anteriores y sea $\{x(t) : t \geq 0\}$ una trayectoria generada por (S).*
Para todo punto fijo $z \in R^n_+$ y todo $t \geq 0$ definimos

$$\begin{aligned} E: \ & R && \to R^n_+ \times R^n_+ \\ & t && \to E(t) = \epsilon(z, x(t)) \end{aligned}$$

Entonces

1. $f(x(t)) - f(z) \leq t^{-1}\left[\epsilon(z, x_0) - \epsilon(z, x(t))\right], \ \forall z \in R^n_+, \ \forall t > 0.$
2. *La función $E(t)$ es decreciente para todo $z \in R^n_+$ tal que $f(z) \leq \inf_{s\geq 0} f(x(s))$.*

Demostración. Para todo $z \in R^n_{++}$ consideremos

$$\begin{aligned} E(t) &= \epsilon(z, x(t)) \\ &= \frac{\nu}{2}\|z - x(t)\|^2 + \mu \sum_{j=1}^{n} x_j(t)\, \varphi\left(\frac{z_j}{x_j(t)}\right) \end{aligned}$$

donde $\varphi\left(\frac{z_j}{x_j(t)}\right) = \frac{z_j}{x_j(t)} \log \frac{z_j}{x_j(t)} - \frac{z_j}{x_j(t)} + 1$ con $x(t)$ solución del sistema (S). Entonces para todo $t \geq 0$

$$E(t) = \frac{\nu}{2}\|z - x(t)\|^2 + \mu \sum_{j=1}^{n}\left(z_j \log \frac{z_j}{x_j(t)} - z_j + x_j(t)\right)$$

Derivando $E(t)$ respecto a t obtenemos

$$\frac{dE(t)}{dt} = 2\frac{\nu}{2}\|z - x(t)\| \|z - x(t)\|' (-x'(t)) + \mu \sum_{j=1}^{n} -z_j \frac{x_j(t)}{z_j} \frac{z_j}{x_j^2(t)} x_j'(t) + x_j'(t)$$

Lo cual es equivalente a la siguiente expresión

$$\frac{dE(t)}{dt} = \nu \|z - x(t)\| \frac{\langle x(t) - z, x'(t) \rangle}{\|z - x(t)\|} + \mu \sum_{j=1}^{n} -\frac{z_j}{x_j(t)} x_j'(t) + x_j'(t)$$

Reduciendo la expresión anterior tenemos

$$\begin{aligned}\frac{dE(t)}{dt} = E'(t) &= \nu \langle x(t) - z, x'(t) \rangle + \sum_{j=1}^{n} \mu \frac{x_j'(t)}{x_j(t)} (x_j(t) - z_j) \\ &= \nu \sum_{j=1}^{n} x_j'(t)(x_j(t) - z_j) + \sum_{j=1}^{n} \mu \frac{x_j'(t)}{x_j(t)} (x_j(t) - z_j)\end{aligned}$$

Por lo tanto, la derivada de $E(t)$ con respecto a t se reduce a

$$\sum_{j=1}^{n} (x_j(t) - z_j) \left(\frac{x_j'(t)}{x_j(t)} \right) (\nu x_j(t) + \mu) \tag{5.19}$$

Del sistema (S) obtenemos

$$\frac{\partial f(x(t))}{\partial x_j} = -x_j'(t) \left[\frac{\mu + \nu x_j(t)}{x_j(t)} \right]$$

entonces reemplazando en (5.19)

$$\frac{dE(t)}{dt} = \langle z - x(t), \nabla f(x(t)) \rangle \tag{5.20}$$

desde que f es una función convexa

$$\frac{dE(t)}{dt} \leq f(z) - f(x(t)) \tag{5.21}$$

Por lo tanto, la función $t \to E(z, x(t))$ es decreciente para todo $z \in R_+^n$ tal que

$$f(z) \leq \inf_{s \geq 0} f(x(s)).$$

En consecuencia queda demostrado el item (b).
Ahora probaremos el primer item, consideremos la desigualdad probada en (5.21), donde $t > 0$, para todo $s \in [0, t]$

$$f(x(s)) - f(z) \leq -E'(s) \tag{5.22}$$

Notemos que $E(t)$ es una función continua en $[0, t]$, aplicando el teorema de valor medio obtenemos

$$E(t) - E(0) = E'(s)(t - 0)$$

Reemplazando este resultado en (5.22)

$$f(x(s)) - f(z) \leq \frac{E(0) - E(t)}{t}$$

lo cual es equivalente a

$$t(f(x(s)) - f(s)) \leq E(0) - E(t).$$

Por la definición de $E(t) = \epsilon(z, x(t))$

$$t(f(x(s)) - f(z)) \leq \epsilon(z, x_0) - \epsilon(z, x(t))$$

entonces,

$$f(x(s)) - f(z) \leq t^{-1}[\epsilon(z, x_0) - \epsilon(z, x(t))]$$

Lo cual demuestra el item (a). Por lo tanto, queda demostrada la proposición. ∎

Antes de probar el teorema, es iteresante recalcar la prueba de convergencia de Bruck.en el caso de descenso continuo , el cual corresponde formalmente para $\mu = 0$. En este caso , existe una familia de funciones de Lyapunov definida como sigue. Supongamos que $X \neq \phi$, para todo punto fijo $z \in X$ definimos

$$\epsilon(x, y) = \frac{\nu}{2}\|x - y\|^2 + \mu \sum_{j=1}^{n}\left(x_j log\frac{x_j}{y_j} + y_j - x_j\right)$$

Teorema 5.3 *Supongamos que se satisfacen **(H1)** y **(H2)**. Si $x : [0, \infty) \to R^n_{++}$ es una solución global de (S), entonces*

1. $\lim_{t\to\infty} f(x(t)) = \inf_{R^n_+} f$
2. *Cuando $X \neq \emptyset$, existe $x_\infty \in X$ tal que $\lim_{t\to\infty} x(t) = x_\infty$*

Demostración. Primero, notemos que

$$f(x(t)) \geq \inf_{R^n_+} f \tag{5.23}$$

aplicando límite cuando $t \longrightarrow \infty$ tenemos

$$\lim_{t\to\infty} f(x(t)) \geq \inf_{R^n_+} f.$$

Por otro lado, por su definición

$$\epsilon(z, x(t)) \geq 0, \ \forall z \in R^n_+$$

de la proposición anterior obtenemos

$$f(x(t)) - f(z) \leq t^{-1}[\epsilon(z, x_0) - \epsilon(z, x(t))]$$

desde que $t^{-1}\epsilon(z, x(t)) \geq 0$

$$f(x(t)) - f(z) \leq t^{-1}\epsilon(z, x_0), \ \forall z \in R_+^n$$

aplicando límite cuando $t \to \infty$

$$\lim_{t\to\infty} f(x(t)) \leq f(z) \leq \inf_{s\geq 0} f(x(s)) \tag{5.24}$$

Por tanto, de (5.23) y (5.24) obtenemos

$$\lim_{t\to\infty} f(x(t)) = \inf_{R_+^n} f$$

Supongamos que $X = \arg\min_{R_+^n} f \neq \emptyset$. Por la definición de ϵ, el segundo término es siempre no negativo

$$\begin{aligned}
\epsilon(z, x(t)) &= \frac{\nu}{2}\|z - x(t)\|^2 + \mu \sum_{j=1}^{n} x_j(t)\varphi\left(\frac{z_j}{x_j(t)}\right) \\
&\geq \frac{\nu}{2}\|z - x(t)\|^2, \quad \forall z \in R_+^n.
\end{aligned}$$

Desde que $t \to \epsilon(z, x(t))$ es decreciente, la desigualdad anterior implica

$$\|z - x(t)\|^2 \leq \frac{2}{\nu}\epsilon(z, x(t)) \leq \frac{2}{\nu}\epsilon(z, x_0), \quad \forall z \in R_+^n. \tag{5.25}$$

La expresión (5.25) es válida para todo $z \in R_+^n$, en particular para $z = 0$

$$\|x(t)\|^2 \leq \frac{2}{\nu}\epsilon(0, x_0)$$

A partir de esta expresión obtenemos

$$\sup_{t\in[0,\infty)} \|x(t)\| < \infty$$

es decir, $\|x(t)\|$ es acotada. Como, $x(t)$ es acotada posee una subsucesión $x(t_k)$ convergente, es decir, $\lim_{t_k\to\infty} x(t_k) = x_\infty$, donde x_∞ es un punto límite de la tayectoria $x(t)$.
Por el primer item, se tiene el siguiente resultado

$$\lim_{t\to\infty} f(x(t)) = \inf_{R_+^n} f \tag{5.26}$$

desde que f es continua

$$f\left(\lim_{t\to\infty} x(t)\right) = f(x_\infty) = \inf_{R_+^n} f$$

y de esta manera $x_\infty \in X$.
Para completar la prueba que la trayectoria converge a $x_\infty \in X$, mostraremos que

el punto límite x_∞ es único.
En efecto, asumiremos que existen dos puntos límites

$$\lim x(t_k) = x_\infty \in X \ \wedge \ \lim x(s_k) = y_\infty \in X$$

donde $\{t_k\}$ y $\{s_k\}$ son sucesiones en el octante no negativo.
Por la proposición anterior $E(t) = \epsilon(z, x(t))$ es decreciente

$$E(t) \leq E(0), \ \forall t \in [0, \infty)$$

aplicando límite cuando $t \longrightarrow \infty$

$$\lim_{t \longrightarrow \infty} E(t) \leq E(0), \ \forall t \in [0, \infty)$$

tal que $f(z) \leq \inf_{s \geq 0} f(x(s))$ entonces

$$\lim_{t \longrightarrow \infty} \epsilon(z, x(t)) \leq \epsilon(z, x_0).$$

Consideremos $z = x_\infty$, desde que $x_\infty \in X$, entonces

$$\lim_{t \longrightarrow \infty} \epsilon(x_\infty, x(t)) \leq \epsilon(x_\infty, x_0)$$

Por lo tanto, $\lim_{t \longrightarrow \infty} \epsilon(x_\infty, x(t))$ existe y

$$\lim_{t_k \longrightarrow \infty} \epsilon(x_\infty, x(t_k)) = \lim_{s_k \longrightarrow \infty} \epsilon(x_\infty, x(s_k))$$

Además, es fácil verificar que

$$\begin{aligned} \lim_{t_k \longrightarrow \infty} \epsilon(x_\infty, x(t_k)) &= \frac{\nu}{2} \lim_{t_k \longrightarrow \infty} \|x_\infty - x(t_k)\| + \mu \sum_{j=1}^{n} \lim_{t_k \longrightarrow \infty} \left(x_{\infty_j} \log \frac{x_{\infty_j}}{x_j(t_k)} - x_{\infty_j} + x_j(t_k) \right) \\ &= 0 \end{aligned}$$

a partir de la expresión anterior obtenemos

$$\lim_{s_k \longrightarrow \infty} \epsilon(x_\infty, x(s_k)) = 0$$

Por

$$\epsilon(x_\infty, x(s_k)) \geq \frac{\nu}{2} \|x_\infty - x(s_k)\|^2$$

aplicando límite cuando $s_k \longrightarrow \infty$

$$\lim_{s_k \longrightarrow \infty} \epsilon(x_\infty, x(s_k)) \geq \lim_{s_k \longrightarrow \infty} \frac{\nu}{2} \|x_\infty - x(s_k)\|^2$$

A partir de obtenemos

$$0 \geq \lim_{s_k \longrightarrow \infty} \frac{\nu}{2} \|x_\infty - x(s_k)\|^2 \geq 0$$

entonces

$$\lim_{s_k \longrightarrow \infty} \frac{\nu}{2} \|x_\infty - x(s_k)\|^2 = 0$$

Por tanto,

$$\lim_{s_k \longrightarrow \infty} x(s_k) = x_\infty$$

En consecuencia, obtenemos $x_\infty = y_\infty$. ■

5.4.1 Minimización Cuasi-convexa

Los problemas de minimización cuasi-convexa tienen diversas aplicaciones, ver el reciente trabajo de Crouziex. Recalcamos que una función es cuasi-convexa si y solo si

$$L(f,\alpha) = \{x \in R^n : f(x) \leq \alpha\}$$

es convexa para todo $\alpha \in R$ y que la función cuasi-convexa el mínimo local no es un mínimo global.

Teorema 5.4 *Supongamos que (H1') y (H2) se satisfacen. Si la trayectoria* $x : [0, +\infty) \to R^n_{++}$ *es una solución global del sistema* (S) *y si el conjunto*

$$Z = \left\{ z \in R^n_+ : f(z) \leq \inf_{t\geq 0} f(x(t)) \right\} \neq \emptyset,$$

entonces existe $x_\infty \in Z$ *tal que* $\lim_{t\to\infty} x(t) = x_\infty$.

Demostración. Utilizaremos la identidad (5.20) probada para el caso convexo

$$E(t) = \langle z - x(t), \nabla f(x(t)) \rangle, \quad \forall z \in R^n_+$$

para la trayectoria $x(t)$ tenemos

$$z \in Z = \left\{ z \in R^n_+ : f(z) \leq \inf_{t\geq 0} f(x(t)) \right\}$$

desde que Z es no vacío, notamos por la cuasi-convexidad de f el conjunto

$$C(t) = \{x \in R^n : f(x) \leq f(x(t))\}$$

asumiendo que es no vacío, es convexo. Desde que $z \in Z$

$$f(z) \leq \inf_{t\geq 0} f(x(t)) \leq f(x(t))$$

entonces obtenemos

$$\langle z - x(t), \nabla f(x(t)) \rangle \leq 0, \ \forall z \in Z$$

es decir

$$E'(t) \leq 0$$

entonces $E(t)$ es decreciente cuando $t \to \infty$

$$E(t) \leq E(0), \ \forall t \geq 0$$

Aplicando límite cuando $t \to \infty$

$$\lim_{t\to\infty} E(t) \leq E(0).$$

Por tanto, el $\lim_{t\to\infty} E(t)$ existe. Usando el mismo argumento que en el caso convexo concluimos que $x(t)$ es acotado, entonces $x(t)$ posee una subsucesión convergente, es decir, sin pérdida de generalidad $x(t_k) \longrightarrow x_\infty$.
Por otro lado, desde que $f(x(t))$ es decreciente cuando $t \longrightarrow \infty$ podemos asumir sin pérdida de generalidad $t_k \geq t$ para todo t arbitrario pero fijo entonces,

$$f(x(t_k)) \leq f(x(t))$$

aplicando límite cuando $t_k \longrightarrow \infty$

$$\lim_{t_k\to\infty} f(x(t_k)) \leq f(x(t))$$

por la continuidad de f se tiene

$$f\left(\lim_{t_k\to\infty} x(t_k)\right) \leq f(x(t))$$

entonces

$$f(x_\infty) \leq \inf f(x(t))$$

es decir, $x_\infty \in Z$. Para completar la prueba probaremos que el punto límite x_∞ es único. Supongamos que existen dos puntos límites x_∞, y_∞ tal que

$$\lim_{t_k\to\infty} x(t_k) = x_\infty \ \wedge \ \lim_{s_k\to\infty} y(s_k) = y_\infty,$$

entonces desde que existe el límite de $E(t) = \epsilon(z, x(t))$ cuando $t \longrightarrow \infty$ obtenemos

$$\lim_{t_k\to\infty} \epsilon(x_\infty, x(t_k)) = \lim_{s_k\to\infty} \epsilon(x_\infty, x(s_k)).$$

Además, es fácil verificar que

$$\begin{aligned}\lim_{t_k\to\infty} \epsilon(x_\infty, x(t_k)) &= \frac{\nu}{2}\lim_{t_k\to\infty} \|x_\infty - x(t_k)\| + \mu\sum_{j=1}^{p}\lim_{t_k\to\infty}\left(x_{\infty_j}\log\frac{x_{\infty_j}}{x_j(t_k)} - x_{\infty_j} + x_j(t_k)\right)\\ &= 0\end{aligned}$$

a partir de la expresión anterior obtenemos

$$\lim_{s_k\to\infty} \epsilon(x_\infty, x(s_k)) = 0$$

Por la definición de ϵ tenemos

$$\epsilon(x_\infty, x(s_k)) \geq \frac{\nu}{2}\|x_\infty - x(s_k)\|^2$$

aplicando límite cuando $s_k \longrightarrow \infty$

$$\lim_{s_k\to\infty} \epsilon(x_\infty, x(s_k)) \geq \lim_{s_k\to\infty} \frac{\nu}{2}\|x_\infty - x(s_k)\|^2$$

A partir de obtenemos

$$0 \geq \lim_{s_k \to \infty} \frac{\nu}{2} \|x_\infty - x(s_k)\|^2 \geq 0$$

entonces

$$\lim_{s_k \to \infty} \frac{\nu}{2} \|x_\infty - x(s_k)\|^2 = 0$$

Por tanto

$$\lim_{s_k \to \infty} x(s_k) = x_\infty$$

En consecuencia, obtenemos $x_\infty = y_\infty$. Por lo tanto, $\lim_{t \to \infty} x(t) = x_\infty$ tal que $x_\infty \in Z$. ■

Observación 5.1 *Notemos que si f es estrictamente cuasi-convexa entonces un mínimo local del problema es también un mínimo global.*

5.4.2 Localización del Punto Límite

En el análisis asintótico es interesante el localizar el punto límite de la trayectoria. Consideremos

$$X = \arg\min_{R_+^n} f \ \wedge \ \pi_X : R_{++}^n \longrightarrow R_+^n$$

el operador proyección en el conjunto cerrado convexo X definido por

$$\pi_X = \arg\min_{x \in X} \epsilon(x, x_0), \quad x_0 \in R_{++}^n$$

denotamos

$$\rho(x_0, X) = \inf\{\epsilon(x, x_0) : x \in X\}$$

Proposición 5.2 *Si $x(t)$ es una solución del sistema (S) con $x(0) = x_0$ y asumimos que $X \neq 0$. Consideremos $x_\infty \in X$ tal que $x(t) \longrightarrow x_\infty$ cuando $t \longrightarrow \infty$. Entonces,*

$$\|x_0 - x_\infty\| \leq \theta\sqrt{\rho(x_0, X)}$$

donde la constante $\theta = 2\sqrt{2\nu^{-1}} > 0$.

Demostración. Usando la proposición con $z \in X$, desde que ϵ es decreciente obtenemos

$$\|z - x(t)\|^2 \leq \frac{2}{\nu}\epsilon(z, x(t)) \leq \frac{2}{\nu}\epsilon(z, x(0))$$

Pasando el límite cuando $t \longrightarrow \infty$

$$\|z - x_\infty\|^2 \leq \frac{2}{\nu}\epsilon(z, x_0) \tag{5.27}$$

considerando $z = \pi_X(x_0)$ entonces

$$\|\pi_X(x_0) - x_\infty\| \leq \frac{2}{\nu}\epsilon(\pi_X(x_0), x_0) \leq \frac{2}{\nu}\rho(x_0, X)$$

Por la regla del paralelogramo para $u = x_0 - \pi_X(x_0)$ y $v = \pi_X(x_0) - x_\infty$

$$2\left(\|x_0 - \pi_X(x_0)\|^2 + \|\pi_X(x_0) - x_\infty\|^2\right) = \|x_0 - x_\infty\|^2 + \|x_0 - 2\pi_X(x_0) + x_\infty\|$$

entonces

$$\begin{aligned}
\|x_0 - x_\infty\|^2 &\leq 2\left(\|x_0 - \pi_X(x_0)\|^2 + \|\pi_X(x_0) - x_0\|^2\right) \\
&\leq 2\left(\frac{2}{\nu}\epsilon(\pi_X(x_0), x_0) + \frac{2}{\nu}\rho(x_0, X)\right) \\
&\leq 2\left(\frac{2}{\nu}\rho(x_0, X) + \frac{2}{\nu}\rho(x_0, X)\right) = \frac{8}{\nu}\rho(x_0, X)
\end{aligned}$$

Por lo tanto, $\|x_0 - x_\infty\| \leq 2\sqrt{2\nu^{-1}}\sqrt{\rho(x_0, X)}$ donde $\theta = 2\sqrt{2\nu^{-1}}$. ∎

5.5 Esquema Discreto de Aproximación Implícita

En esta sección, mostramos el esquema de discretización implícita el cual es el origen del sistema dinámico. De esta manera, para solucionar el problema de minimización

$$(P) \quad \begin{cases} \min f(x) \\ x \in R_+^n \end{cases}$$

Consideremos las siguientes hipótesis para el problema (P).

A0 $f : R_{++}^n \to R \cup \{+\infty\}$ es una función propia, convexa y cerrada con

$$\inf\{f(x) : x \in R_+^n\} = f^* > \infty.$$

A1 *$Dom f \cap R_{++}^n \neq \emptyset$.*

Teorema 5.5 *Supongamos que las hipótesis $(A0)$ y $(A1)$ se satisfacen. Sea $\{x^k\} \subset R_{++}^n$ es una sucesión generada por el método proximal. Entonces se satisfacen los siguientes resultados*

1. $f(x^k) - f(x) \leq \sigma_k^{-1}\epsilon(x, x^0)$, $\forall x \in R_+^n$.
2. *Si $\sigma_k \to \infty$, cuando $k \to +\infty$, $f(x^k) \to f^* = \inf\{f(x) : x \in R_+^n\}$*
3. *Si $X \neq \emptyset, \sigma_k \to \infty$, la sucesión $\{x^k\}$ converge a una solución óptima del problema (P).*
4. *Además, si $\lambda_k \geq \lambda > 0$, se tiene $\lim_{k\to\infty} x_j^k g_j^k = 0$, $\forall j = 1, ..., n$.*

Demostración. Usando la definición de la subdiferencial de la función convexa f tenemos

$$f(x) \geq f(x) + \langle g^k, x - x^k \rangle$$

donde $g^k \in \partial f(x^k)$. Entonces, reemplazamos g^k en la siguiente desigualdad

$$\lambda_k \left(f\left(x^k\right) - f(x) \right) \leq -\lambda \left\langle g^k, x - x^k \right\rangle$$

$$\lambda_k \left(f\left(x^k\right) - f(x) \right) \leq -\left\langle \nu\left(x^k - x^{k-1}\right) + \mu \left(\varphi'\left(\frac{x_1^k}{x_1^{k-1}}\right), ..., \varphi'\left(\frac{x_p^k}{x_p^{k-1}}\right) \right), x - x^k \right\rangle$$

$$= \nu \left\langle x^k - x^{k-1}, x - x^k \right\rangle + \mu \left\{ \sum_{j=1}^{n} \left(1 - \frac{x_j^{k-1}}{x_j^k}\right) \left(x_j - x_j^k\right) \right\}$$

desde que $\varphi(r) = r - logr - 1$ entonces $\varphi'(r) = 1 - \frac{1}{r}$ Usando el mismo argumento que en la Proposición 5.1

$$\begin{aligned} \epsilon\left(x, x^{k-1}\right) - \epsilon\left(x, x^k\right) &= \left(\tfrac{\nu}{2} \left\|x - x^{k-1}\right\|^2 + \mu \textstyle\sum_{j=1}^n x_j \log \frac{x_j}{x_j^{k-1}} + x_j^{k-1} - x_j \right) - \\ &\quad \left(\tfrac{\nu}{2} \left\|x - x^k\right\|^2 + \mu \textstyle\sum_{j=1}^n x_j \log \frac{x_j}{x_j^k} + x_j^k - x_j \right) \end{aligned}$$

Desde que $\left\|x - x^{k-1}\right\|^2 = \left\langle x - x^{k-1}, x - x^{k-1} \right\rangle$ y $\left\|x - x^k\right\|^2 = \left\langle x - x^k, x - x^k \right\rangle$. Reemplazando en la igualdad anterior obtenemos

$$\nu \left\langle x - x^k, x^k - x^{k-1} \right\rangle + \frac{\nu}{2} \left\|x^k - x^{k-1}\right\|^2 + \nu \sum_{j=1}^{n} \left(x_j \log \frac{x_j^k}{x_j^{k-1}} + x_j^{k-1} - x_j^k \right) \quad (5.28)$$

Usando el hecho que $\|.\|^2 \geq 0$ y de la desigualdad

$$\log r \geq 1 - \frac{1}{r}$$

en (5.28) y recalcando que $x_j \geq 0$ obtenemos que

$$\epsilon\left(x, x^{k-1}\right) - \epsilon\left(x, x^k\right) \geq \nu \left\langle x - x^k, x^k - x^{k-1} \right\rangle + \mu \left\{ \sum_{j=1}^{n} x_j \left(\frac{x_j^k - x_j^{k-1}}{x_j^k} \right) + x_j^{k-1} + x_j^k \right\} (5.29)$$

Combinando (5.28) y (5.29) obtenemos

$$\lambda_k \left(f\left(x^k - f(x)\right) \right) \leq \epsilon\left(x, x^{k-1}\right) - \epsilon\left(x, x^k\right), \ \forall x \in R_+^n$$

Sumando sobre $k = 1, ..., n$ la siguiente desigualdad y recalcando que $\sigma_n = \sum_{k=1}^n \lambda_k$ entonces

$$-\sigma_n f(x) + \sum_{k=1}^{n} \lambda_k f\left(x^k\right) \leq \epsilon\left(x, x^0\right) - \epsilon\left(x, x^n\right)$$

Por otro lado, desde que

$$x^k = \arg\min\left\{f(x) + \lambda_k^{-1} d\left(x, x^{k-1}\right)\right\}$$

$$f(x) + \lambda_k^{-1} d\left(x, x^{k-1}\right) \geq f\left(x^k\right) + \lambda_k^{-1} d\left(x^k, x^{k-1}\right)$$

En particular para $x = x^{k-1} \in R_{++}^n$, recalcando que $d(.,.) \geq 0$ y $d\left(x^{k-1}, x^{k-1}\right) = 0$

$$f\left(x^{k-1}\right) \geq f\left(x^k\right), \ \forall k$$

de esta manera se prueba que la sucesión $\left\{f\left(x^k\right)\right\}$ es decreciente. A partir de la definición de $\sigma_n = \sum_{k=1}^n \lambda_k$, entonces para $k \geq 1$ se puede escribir la siguiente relación $\sigma_k = \lambda_k + \sigma_{k-1}$ con $\sigma_0 = 0$, multiplicando la anterior desigualdad por σ_{k-1}

$$\sigma_{k-1} f\left(x^k\right) \leq \sigma_{k-1} f\left(x^{k-1}\right)$$

lo cual es equivalente a

$$\left(\sigma_k - \lambda_k\right) f\left(x^k\right) \leq \sigma_{k-1} f\left(x^{k-1}\right)$$

y sumando sobre $k = 1, ..., n$ obtenemos

$$\sigma_n f\left(x^n\right) - \sum_{k=1}^{n} \lambda_k f\left(x^k\right) \leq 0$$

De (5.30)

$$f\left(x^n\right) - f(x) \leq \sigma_n^{-1}\left(\epsilon\left(x, x_0\right) - \epsilon\left(x, x^n\right)\right) \leq \sigma_n^{-1} \epsilon\left(x, x_0\right)$$

Lo cual demuestra 1.
Tomando límite en el resultado de 1 cuando $n \to \infty$ y desde que $\sigma_n \to \infty$

$$\limsup_{n \to \infty} f\left(x^n\right) \leq \inf\left\{f(x) : x \in R_+^n\right\}$$

Además,

$$\liminf_{n \to \infty} f\left(x^n\right) \geq \inf\left\{f(x) : x \in R_+^n\right\}$$

y desde que la sucesión $\{f(x^n)\}$ es decreciente . De esta manera, obtenemos que

$$\lim_{n \to \infty} f\left(x^n\right) = f^*$$

Lo cual demuestra 2.
Sea $x^* \in \arg\min_{R_+^n} f \neq \phi$, desde que

$$f\left(x^k\right) \geq f\left(x^*\right), \forall k$$

obtenemos de

$$\epsilon\left(x^*, x^k\right) \leq \epsilon\left(x^*, x^{k-1}\right).$$

Pero, para todo $(x, y) \in R^n_{++} \times R^n_{++}$

$$\epsilon(x, y) \geq \frac{\nu}{2}\|x-y\|^2$$

Por lo tanto, obtenemos

$$\left\|x^* - x^k\right\|^2 \leq \frac{2}{\nu}\epsilon\left(x^*, x_0\right) < \infty$$

En consecuencia, la sucesión $\{x_k\}$ es acotada. La convergencia de x^k a una solución x^* satisface las mismas condiciones que en la prueba del caso continuo.
Usando tenemos que para todo $j = 1, ..., n$

$$\begin{aligned} x_j^k g_j^k = & \ \lambda_k^{-1}\left[\nu\left(x_j^{k-1} - x_j^k\right)_j^k + \mu\left(x_j^{k-1} - x_j^k\right)\right] \\ = & \ \lambda_k^{-1}\left(x_j^{k-1} - x_j^k\right)\left(\mu + \nu x_j^k\right) \end{aligned}$$

Desde que $\lambda_k > \lambda > 0$ y x_k es acotado, queda demostrado. ∎

5.6 Sistema Dinámico Lotka Volterra cuando $\nu \longrightarrow 0$

Fijamos $\mu > 0$ y consideramos la familia de sistemas dinámicos (S_ν) parametrizado por $\nu > 0$ con las correspondientes trayectorias x^ν definida por

$$\begin{aligned} (S_\nu) \qquad & x_j^{\prime\nu}(t) + \tfrac{x_j^\nu}{\mu + \nu x_j^\nu}\tfrac{\partial f(x^\nu(t))}{\partial x_j} = 0 \ j = 1, ..., n \\ & x_j^\nu(0) = x_{0j} \in R^n_{++} \quad j = 1, ..., n \end{aligned}$$

Una natural pregunta cuando $\nu \longrightarrow 0$, la sucesión de trayectorias $\{x^\nu\}$ converge al correspondiente sistema dinámico Lotka Volterra (S_0) dado por

$$\begin{aligned} (S_0) \qquad & x_j^{\prime}(t) + \mu^{-1} x_j(t)\tfrac{\partial f(x(t))}{\partial x_j} = 0 \ j = 1, ..., n \\ & x_j(0) = x_{0j} \in R^n_{++} \ j = 1, ..., n. \end{aligned}$$

En esta sección damos una respuesta positiva a esta pregunta y por eso hacemos una justificación de la terminología "sistema dinámico Lotka Volterra" para el sistema (S) introducido en este trabajo. Para probar este resultado, usamos un argumento de compacidad primero probaremos que la sucesión $\{x^\nu\}$ es relativamente compacto por la topología de la convergencia uniforme en el conjunto acotado de $[0, \infty)$.

Teorema 5.6 *Sea $f : R^n \longrightarrow R$ es una función de $C^2(R^n)$ acotada definida en el octante no negativo, consideremos para cada $\nu > 0$ la familia asociada del sistema (S_ν) con solución global $x^\nu(.) \in R^n_{++}$. Sea $x^\nu \in R^n_{++}$, una única solución del sistema Lotka Volterra (S_ν). Entonces*

1. La sucesión $\{x^{\nu}(.)\}_{\nu>0}$ es relativamente compacto para la topología de la convergencia uniforme en subconjuntos acotados de $[0,+\infty)$.

2. La trayectoria $x^{\nu}(.)$ generada por (S_{ν}) converge a $x(.)$ solución de (S_0) cuando $\nu \rightarrow 0$.

Demostración. Primero recalcamos que existe una única solución global del sistema (S_{ν}), $x^{\nu}(t) \in R^{n}_{++}$, $\forall t \geq 0$. Usando el mismo argumento pero ahora para el sistema obtenemos, para todo $\nu > 0$

$$f(x^{\nu}(t)) + \sum_{j=1}^{n} \int_{0}^{t} \left(\nu + \frac{\nu}{x_j^{\nu}}(s)\right) x_j^{\prime\nu}(s)^2 \, ds = f(x_0)$$

desde que f es acotada inferiormente y $\nu > 0$ obtenemos que

$$\sup_{\nu>0} \int_{0}^{+\infty} \left(\frac{x_j^{\prime\nu}(t)^2}{x_t^{\nu}(t)}\right) dt < \infty, \ \forall j = 1, ..., n \tag{5.30}$$

Definiendo

$$y_j^{\nu}(t) = \sqrt{x_j^{\nu}(t)} \ \forall j = 1, ..., p$$

y notando que

$$\frac{dy_j^{\nu}(t)}{dt} = \frac{1}{2} \frac{x_j^{\prime\nu}}{\sqrt{x_j^{\nu}(t)}} \quad \forall j = 1, ..., n$$

De (5.30) obtenemos que

$$M = \sup\left(\int_{0}^{+\infty} \|y^{\prime\nu}\|^2 \, dt\right)^{\frac{1}{2}} < \infty$$

Ahora usando la siguiente desigualdad y desde que

$$y^{\nu}(0) = \left(\sqrt{x_j^{\nu}(t)}\right)_{j=1,...,n}$$

se concluye, para todo $0 \leq s < t < \infty$,

$$\|y^{\nu}(t) - y^{\nu}(s)\| = \left\|\int_{0}^{t} y^{\prime\nu}(t) \, d\tau\right\| \leq \sqrt{t-s}\left(\int_{s} \|y^{\prime\nu}(\tau)\|^2 \, d\tau\right)^{\frac{1}{2}} \leq M\sqrt{t-s} \tag{5.31}$$

Desde que la desigualdad es válida para todo s, en particular para $s = 0$

$$\|y^{\nu}\| - \|y_0\| \leq \|y^{\nu}(t) - y^{\nu}(0)\| \leq M\sqrt{T}, \quad \forall t \geq 0$$

de ahí,

$$\|y^{\nu}\| \leq \|y_0\| + M\sqrt{t} \tag{5.32}$$

Elevando al cuadrado en ambos miembros de la desigualdad, para todo $T < \infty$

$$0 \leq x_j^\nu(t) \leq \left(\|y_0\| + M\sqrt{T}\right)^2, \quad \forall j = 1, ..., p \ , \ \forall t \in [0, T] .$$

Además, usando (5.32) para todo $j = 1, ..., n$ y todo $0 \leq s < t \leq T$,

$$\left|y_j^\nu(t) + y_j^\nu(s)\right| \leq \left|y_j^\nu(t)\right| + \left|y_j^\nu(s)\right| \leq 2\left(M\sqrt{T} + \|y_0\|\right) \tag{5.33}$$

Entonces usando (5.31) y (5.33) , para todo $j = 1, ..., n$ y todo $0 \leq s < t \leq T$, obtenemos

$$\begin{aligned} \left|x_j^\nu(t) - x_j^\nu(s)\right| &= \left|\left(\sqrt{x_j^\nu(t)} - \sqrt{x_j^\nu(s)}\right)\left(\sqrt{x_j^\nu(t)} + \sqrt{x_j^\nu(s)}\right)\right| \\ &\leq \left|\left(y_j^\nu(t) - y_j^\nu(s)\right)\left(y_j^\nu(t) + y_j^\nu(s)\right)\right| \\ &\leq 2\left(\|y_0\| + M\sqrt{T}\right) M\sqrt{t-s} \end{aligned}$$

Por lo tanto, $\{x^\nu(.)\}_{\nu>0}$ son uniformemente equicontinuas desde que la sucesión $\{x^\nu(.)\}_{\nu>0}$ es acotada en $[0, T]$.
Existe una subsucesión, la cual por simplicidad la denotamos por $\{x^\nu\}$, y una función continua $x(.)$ tal que, para todo $T < \infty$
$x_j^\nu(.) \rightarrow x_j(.)$ uniformemente en $[0, T]$, $\forall j = 1, ..., n$
Pasando el límite en el sistema cuando $\nu \rightarrow 0$ en el sentido de las distribuciones $x_j'^\nu \rightarrow x_j'(.)$ en $D'(0, \infty)$.
De otro lado, cuando $\nu \rightarrow 0$

$$\frac{x_j^\nu(.)}{\mu + \nu x_j^\nu(.)} \rightarrow \frac{x_j(.)}{\mu}$$

uniformemente en $[0, T]$.
Para probar esta afirmación, notemos que para todo $t \in [0, T]$

$$\begin{aligned} \left|\frac{x_j^\nu}{\mu + \nu x_j^\nu} - \frac{x_j(t)}{\mu}\right| &= \frac{\left|\mu x_j^\nu(t) - \mu x_j(t) - \nu x_j(t) x_j^\nu(t)\right|}{\mu\left(\mu + \nu x_j^\nu(t)\right)} \\ &\leq \frac{\left|x_j^\nu(t) - x_j(t)\right|}{\mu\nu x_j^\nu(t)} + \frac{\nu}{\mu}\frac{x_j(t) x_j^\nu(t)}{\mu + \nu x_j^\nu(t)} \\ &\leq \frac{1}{\mu}\left|x_j^\nu(t) - x_j(t)\right| + C\nu \end{aligned}$$

donde la constante C es obtenida acotando superiormente .
Por lo tanto, desde que $x^\nu(.) \rightarrow x(.)$ uniformemente en $[0, T]$, queda demostrado.
Además, desde que $f \in C^1$, tenemos también , para todo $j = 1, ..., n$ cuando $\nu \rightarrow 0$,

$$\frac{\partial f(x^\nu(t))}{\partial x_j} \text{ converge uniformemente a } \frac{\partial f(x(t))}{\partial x_j} \text{ en } [0, T]$$

Por lo tanto, cuando $\nu \rightarrow 0$

$$\frac{x_j^\nu(.)}{\mu + \nu x_j^\nu(t)} \frac{\partial f(x^\nu(t))}{\partial x_j} \rightarrow \frac{x_j(t)}{\mu} \frac{\partial f(x(t))}{\partial x_j}$$

uniformemente en $[0, T]$.
Por lo tanto, pasando el límite en $D'(0,\infty)$, obtenemos finalmente

$$x_j'(t) + \frac{x_j(t)}{\mu}\frac{\partial f(x(t))}{\partial x_j} = 0$$

Desde que $x_j(.)\frac{\partial f(x(.))}{\partial x_j}$ es continuo, esto implica que $x(.)$ es una clásica solución del sistema (S_0). Por el teorema de Cauchy Lipschitz, esta es la solución única de (S_0). Para terminar la prueba usamos argumentos de compacidad, entonces obtenemos que la sucesión $x^{\nu}(.) \to x(.)$ cuando $\nu \to 0$. ■

Concluimos en este trabajo notando el interesante hecho que el análisis del sistema dinámico Lotka Volterra (S_0) puede ser derivado en el sistema dinámico propuesto. Para el caso convexo, las funciones de Lyapunov deben elegirse simplemente como la entropía relativa Kullback-Liebler

$$K(\mu,\nu) = \mu\sum_{j=1}^{n}\left(\mu_j \log\left(\frac{\mu_j}{\nu_j}\right) + \nu_j - \mu_j\right)$$

Siguiendo la misma técnica usada en la Proposición , para todo $z \in R_+^n$ y todo $t \geq 0$ definimos

$$K(t) = K(z, x(t))$$

Entonces, es fácil verificar que para una función convexa f, la función es decreciente para todo $z \in R_+^p$ tal que

$$f(z) \leq \inf_{s\leq 0} f(x(s))$$

y esto sostiene que

$$f(x(t)) - f(z) \leq t^{-1}[K(0) - K(t)] \;\; \forall z \in R_+^n, \; \forall t > 0$$

de lo cual concluimos

$$\lim_{t\to\infty} f(x(t)) = \inf_{R_+^p} f$$

La parte mas delicada de probar la convergencia de las trayectorias para un minimizador de f sobre el octante no negativo es obtenido gracias a la simple función entropía K, es fácil verificar los siguientes afirmaciones

F1 Los conjuntos de nivel de $K(\mu, .)$ son acotados para todo $\mu \in R_+^n$

F2 Dada dos sucesiones $\{\mu^k\} \subset R_+^p$ acotada, $\{\nu^k\} \subset R_{++}^n$, tal que $\lim_{k\to\infty}\nu^k = \nu \in R_+^n$ y $\lim_{k\to\infty} K(\mu^k, \nu^k) = 0$ se obtiene $\lim_{k\to\infty}\mu^k = \nu$

asumiendo que el conjunto de soluciones

$$X = \arg\min_{R_+^n} f \neq \phi$$

entonces, para todo $z \in X$

$$K(z, x(t)) \leq K(z, x_0) < \infty, \ \forall t \geq 0$$

usando la propiedad (F1), obtenemos

$$\sup_{t \in [0,\infty)} \|x(t)\| < \infty$$

Es decir, $\{x(t)\}_{t \geq 0}$ es acotada. De esto, podemos deducir que el punto límite x_∞ de la trayectoria $x(t)$ satisface $x_\infty \in X$ y que el $\lim_{t\to\infty} K(x_\infty, x(t))$ existe.
Para probar la convergencia de la solución para una solución óptima, vemos conveniente proceder como en el final de la prueba del teorema. Sin embargo, no tenemos la cota inferior cuadrática, la cual concluye la prueba.
Gracias a la propiedad (F2) para K, se puede proceder también facilmente con las mismas técnicas de Opial para establecer la unicidad del punto límite x_∞.

5.7 Nuevos Resultados

A continuación presentamos un resultado importante en esta sección, probar que toda trayectoria generada por el sistema (S) converge a un punto KKT (punto candidato a solución) en el problema de minimización cuasi-convexa. Consideremos las siguientes hipótesis:

(H1') $f : R^n \to R$ es una función cuasi-convexa en C^2.

(H2) Para cada $x_0 \in R^n_{++}$, existe una solución global $x : [0, \infty) \to R^n_{++}$ del sistema dinámico (S).

$$\begin{aligned} x_j(t) + \frac{x_j(t)}{\mu + \nu x_j(t)} \frac{\partial f(x(t))}{\partial x_j} &= 0, \ \forall j = 1, ..., n \\ x_j(0) = x_{0_j} &\in R^n_{++} \ j = 1, ..., n \end{aligned}$$

Teorema 5.7 *Supongamos que las hipótesis (H1') y (H2) se satisfacen. Si la trayectoria $x : [0, +\infty) \to R^n_+$ es una solución global del sistema (S) y si el conjunto*

$$Z = \left\{ z \in R^n_+ : f(z) \leq \inf_{t \geq 0} f(x(t)) \right\} \neq \phi,$$

entonces existe $x_\infty \in Z$, tal que es un punto KKT del problema de minimización.

Demostración. En la subsección (5.4.1) se demostró,

$$\lim_{t\to\infty} x(t) = x_\infty,$$

donde $x_\infty \in Z$. Probaremos que x_∞ es un punto KKT del problema de minimización, esto es

$$x_{\infty_i} \geq 0, \ \ (\nabla f(x_\infty))_i \geq 0 \ \wedge \ \ x_{\infty_i} (\nabla f(x_\infty))_i = 0 \tag{5.34}$$

A partir de la viabilidad de la solución del sistema $x_{\infty_i} \geq 0$.
Para probar las otras condiciones de (5.34) consideremos los siguientes conjuntos:

$$I(x_\infty) = \{i \in \{1, ..., n\} : x_{\infty_i} = 0\}$$
$$J(x_\infty) = \{i \in \{1, ..., n\} : x_{\infty_i} > 0\}$$

Consideremos separadamente los casos cuando $i \in I(x_\infty)$ y $i \in J(x_\infty)$.

Primer Caso: Sea $i \in I(x_\infty)$ mostraremos que $\nabla f(x_\infty)_i \geq 0$.
Procederemos por contradicción, supongamos que $\frac{\partial f(x_\infty)}{\partial x_i} < 0$.
Desde que f es continuamente diferenciable

$$\lim_{t\to\infty} \frac{\partial f}{\partial x_i}(x(t)) = \frac{\partial f}{\partial x_i}\left(\lim_{t\to\infty} x(t)\right) = \frac{\partial f}{\partial x_i}(x_\infty) < 0$$

Por la definición de límite tenemos

$$\forall \epsilon > 0,\ \exists A > 0 :\ t \in R_+,\ t > A \Rightarrow \left|\frac{\partial f}{\partial x_i}(x(t)) - \frac{\partial f}{\partial x_i}(x_\infty)\right| < \epsilon$$

Es decir, para $t > A$ tenemos

$$\frac{\partial f}{\partial x_i}(x(t)) < \epsilon + \frac{\partial f}{\partial x_i}(x_\infty)$$

considerando

$$\epsilon = -\frac{\partial f}{\partial x_i}(x_\infty) > 0$$

entonces

$$\frac{\partial f}{\partial x_i}(x(t)) < 0, \quad \forall\, t > A.$$

En el sistema dinámico (S)

$$\frac{\partial f}{\partial x_i}(x(t)) = -x_i'(t)\left(\frac{\mu + \nu x_i(t)}{x_i(t)}\right) < 0$$

Desde que

$$\frac{x_i(t)}{\mu + \nu x_i(t)} > 0$$

obtenemos $-x_i(t) < 0$, esto implica que $x_i(t) > 0$. Por tanto, $x_i(t)$ es creciente

$$x_i(0) \leq x_i(t), \quad \forall\, t \geq 0$$

aplicando límite cuando $t \to \infty$

$$\lim_{t\to\infty} x_i(0) \leq \lim_{t\to\infty} x_i(t)$$

$$x_{0_i} \leq x_{\infty_i}$$

a partir de que $x_0 \in R^n_{++}$ concluimos

$$0 < x_{0_i} \leq x_{\infty_i} = 0$$

entonces $0 < 0$, lo cual es una contradicción. Por lo tanto, $\nabla f\,(x_\infty)_i \geq 0$.

Segundo Caso: Sea $i \in J\,(x_\infty)$
Mostraremos que $(\nabla f\,(x_\infty))_i = 0$.
En efecto, desde que

$$\lim_{t\to\infty} x_i\,(t) = x_{\infty_i} > 0$$

en el Sistema Dinámico (S) tenemos

$$\frac{\partial f}{\partial x_i}\,(x_\infty) = \lim_{t\to\infty} \frac{\partial f}{\partial x_i}\,(x\,(t)) = \lim_{t\to\infty} -x_i'\,(t)\left(\frac{\mu + \nu x_i\,(t)}{x_i\,(t)}\right)$$

usando las propiedades de límites

$$\lim_{t\to\infty} x_i'\,(t) \lim_{t\to\infty} \frac{\mu + \nu x_i\,(t)}{x_i\,(t)} = \lim_{t\to\infty} x_i'\,(t)\left(\frac{\mu}{x_{\infty_i}} + \nu\right)$$

desde que $x\,(.)$ es acotado entonces, $\lim_{t\to\infty} x_i'\,(t) = 0$.
Por lo tanto,

$$\frac{\partial f}{\partial x_i}\,(x_\infty) = 0.$$

Lo cual demuestra el teorema. ∎

5.7.1 Algunos Ejemplos

En esta sección, daremos algunos ejemplos.

Ejemplo 5.1 *La función*

$$f\,(x_1, x_2) = (\ln x_1)^2 + (\ln x_2)^2$$

es cuasi-convexa.

En efecto, sin pérdida de generalidad supongamos $1 \leq y_i \leq x_i$ para $i = 1, 2$ y desde que $0 \leq \lambda \leq 1$

$$y_i \leq y_i + \lambda\,(x_i - y_i) \leq x_i$$

desde que $\ln x$ es creciente

$$0 \leq (\ln y_1)^2 \leq (\ln\,(y_1 + \lambda\,(x_1 - y_1)))^2 \leq (\ln x_1)^2$$
$$0 \leq (\ln y_2)^2 \leq (\ln\,(y_2 + \lambda\,(x_2 - y_2)))^2 \leq (\ln x_2)^2$$

Sumando ambas desigualdades obtenemos

$$(\ln y_1)^2 + (\ln y_2)^2 \leq (\ln\,(y_1 + \lambda\,(x_1 - y_1)))^2 + (\ln\,(y_2 + \lambda\,(x_2 - y_2)))^2 \leq (\ln x_1)^2 + (\ln x_2)^2$$

entonces

$$f(y) \leq f(\lambda x + (1-\lambda) y) \leq f(x)$$

Por tanto,

$$f(\lambda x + (1-\lambda) y) \leq \max\{f(x), f(y)\}.$$

Por otro lado, si $0 < y_i \leq x_i < 1$ desde que $0 \leq \lambda \leq 1$

$$y_i \leq y_i + \lambda (x_i - y_i) \leq x_i$$

desde que $\ln x$ es creciente

$$\ln y_1 \leq \ln (y_1 + \lambda (x_1 - y_1)) \leq \ln x_1$$

como el $\ln x$ es negativo para $0 < x < 1$

$$0 < (-\ln x_1)^2 \leq (-\ln (y_1 + \lambda (x_1 - y_1)))^2 \leq (-\ln y_1)^2$$

entonces,

$$\begin{aligned} (\ln x_1)^2 &\leq (\ln (y_1 + \lambda (x_1 - y_1)))^2 \leq (\ln y_1)^2 \\ (\ln x_2)^2 &\leq (\ln (y_2 + \lambda (x_2 - y_2)))^2 \leq (\ln y_2)^2 \end{aligned}$$

por tanto,

$$f(\lambda x + (1-\lambda) y) \leq \max\{f(x), f(y)\}, \ \forall \lambda \in [0,1].$$

Si $1 \leq x_1 \leq y_1$ y $1 \leq y_2 \leq x_2$ sin pérdida de generalidad supongamos que $y_1 \leq y_2$ entonces

$$x_1 \leq y_1 \leq y_2 \leq x_2$$

luego

$$\begin{aligned} x_1 &\leq x_1 + \lambda (y_1 - x_1) \leq y_1 \\ y_2 &\leq x_2 + \lambda (y_2 - x_2) \leq x_2 \end{aligned}$$

desde que ln es creciente

$$\begin{aligned} (\ln x_1)^2 &\leq (\ln (x_1 + \lambda (y_1 - x_1)))^2 \leq (\ln y_1)^2 \\ (\ln y_2)^2 &\leq (\ln (x_2 + \lambda (y_2 - x_2)))^2 \leq (\ln x_2)^2 \end{aligned}$$

de ahí como $y_1 \leq y_2$

$$\begin{aligned} (\ln x_1)^2 \leq (\ln (x_1 + \lambda (y_1 - x_1)))^2 \leq (\ln y_1)^2 \leq (\ln y_2)^2 \\ (\ln y_1)^2 \leq (\ln y_2)^2 \leq (\ln (x_2 + \lambda (y_2 - x_2)))^2 \leq (\ln x_2)^2. \end{aligned}$$

Sumando ambos miembros tenemos

$$(\ln x_1)^2 + (\ln y_1)^2 \leq (\ln (x_1 + \lambda (y_1 - x_1)))^2 + (\ln (x_2 + \lambda (y_2 - x_2)))^2 \leq (\ln x_2)^2 + (\ln y_2)^2$$

lo cual implica que

$$f(\lambda x + (1-\lambda) y) \leq \max\{f(x), f(y)\}, \ \forall \lambda \in [0,1].$$

Si $0 < x_1 < y_1 < 1$ y $0 < y_2 < x_2 < 1$ entonces

$$\begin{aligned} y_2 &\leq \lambda x_2 + (1-\lambda) y_2 \leq x_2 \\ x_1 &\leq \lambda x_1 + (1-\lambda) y_1 \leq y_1 \end{aligned}$$

desde que $\ln x$ es creciente y $\ln x_2 < 0$, $\ln y_1 < 0$ tenemos

$$\begin{gathered} (\ln x_2)^2 \leq (\ln(\lambda x_2 + (1-\lambda) y_2))^2 \leq (\ln y_2)^2 \\ (\ln y_1)^2 \leq (\ln(\lambda x_1 + (1-\lambda) y_1))^2 \leq (\ln x_1)^2 \end{gathered}$$

sumando las desigualdades anteriores y considerando $(\ln y_2)^2 < (\ln x_2)^2$ se obtiene

$$(\ln y_1)^2 + (\ln y_2)^2 \leq (\ln(\lambda x_1 + (1-\lambda) y_1))^2 + (\ln(\lambda x_2 + (1-\lambda) y_2)^2 \leq (\ln x_2)^2 + (\ln y_2)^2$$

lo cual implica que

$$f\left(\lambda x + (1-\lambda) y\right) \leq \max\left\{f\left(x\right), f\left(y\right)\right\}, \ \forall \lambda \in [0,1].$$

En el sistema (S) tenemos

$$\frac{\partial f}{\partial x_1} = 2\frac{\ln x_1}{x_1} \ \wedge \ \frac{\partial f}{\partial x_2} = 2\frac{\ln x_2}{x_2}$$

En el sistema (S) tenemos

$$x_1'\left(t\right) + 2\frac{x_1\left(t\right)}{\mu + \nu x_1\left(t\right)} \frac{\ln x_1\left(t\right)}{x_1\left(t\right)} = 0$$

y

$$x_2'\left(t\right) + 2\frac{x_2\left(t\right)}{\mu + \nu x_2\left(t\right)} \frac{\ln x_2\left(t\right)}{x_2\left(t\right)} = 0$$

Sabemos que

$$\lim_{t \to \infty} x'\left(t\right) = 0$$

luego

$$\lim_{t \to \infty} 2\frac{x_j\left(t\right)}{\mu + \nu x_j\left(t\right)} \frac{\ln x_j\left(t\right)}{x_j\left(t\right)} = 0, \quad j = 1, 2$$

desde que $x_j\left(t\right) > 0$

$$\lim_{t \to \infty} \ln x_j\left(t\right) = 0, \quad j = 1, 2.$$

Por tanto,

$$\lim_{t \to \infty} x_j\left(t\right) = 1, \quad j = 1, 2.$$

Del teorema 5.7 concluimos que el punto $(1,1)$ es un punto candidato a solución del problema

$$\begin{gathered} \min\left(\ln x_1\right)^2 + \left(\ln x_2\right)^2 \\ x \in R_{++}^n \end{gathered}$$

Por tanto, la función f alcanza su mínimo en $(1,1)$.

Ejemplo 5.2 *Consideremos la función f definida en el conjunto convexo*

$$C = \left\{x : x \in R^2, 0 < x_1 < 1, 0 < x_2 < 1\right\}$$

es dada por $f(x) = -(x_1 - 1)^2 - (x_1 - 1)$.

La función es cuasi-convexa. En efecto, sin pérdida de generalidad supongamos que $x_1 \leq y_1$ entonces

$$x_1 \leq \lambda x_1 + (1-\lambda) y_1 \leq y_1, \quad \forall\lambda \in [0,1]$$

luego,

$$(x_1 - 1)^2 \leq (\lambda x_1 + (1-\lambda) x_2 - 1)^2 \leq (y_1 - 1)^2, \quad \forall\lambda \in [0,1]$$

de aquí,

$$-(y_1 - 1)^2 \leq -(\lambda x_1 + (1-\lambda) y_1)^2 \leq -(x_1 - 1)^2, \; \forall\lambda \in [0,1]. \tag{5.35}$$

Además,

$$-y_1 \leq -(\lambda x_1 + (1-\lambda) y_1) \leq -x_1. \tag{5.36}$$

Sumando (5.35) y (5.36) tenemos

$$f(y) \leq f(\lambda x + (1-\lambda) y) \leq f(x), \quad \forall\lambda \in [0,1].$$

Por tanto,

$$f(\lambda x + (1-\lambda) y) \leq \max\{f(x), f(y)\}, \; \forall\lambda \in [0,1].$$

Lo cual demuestra que f es cuasi-convexa. Notemos que,

$$\frac{\partial f}{\partial x_1} = -2x_1 + 1 \;\wedge\; \frac{\partial f}{\partial x_2} = 0.$$

En el sistema (S) tenemos

$$x_1'(t) + \frac{x_1(t)}{\mu + \nu x_1(t)}(-2x_1(t) + 1) = 0$$
$$x_2'(t) + \frac{x_2(t)}{\mu + \nu x_2(t)}\, 0 = 0$$

desde que la trayectoria del sistema (S_ν) converge a la trayectoria de (S). En el sistema (S_ν)

$$x_1'(t) + \frac{x_1(t)}{\mu}(-2x_1(t) + 1) = 0$$
$$x_2'(t) + \frac{x_2(t)}{\mu}\, 0 = 0$$

el cual es equivalente a

$$\begin{aligned} x_1'(t) &= \frac{x_1(t)}{\mu}(2x_1(t) - 1) \\ x_2'(t) &= 0 \end{aligned}$$

de ahí,

$$\frac{x_1'(t)}{x_1(t)(2x_1(t)-1)} = \frac{1}{\mu}$$
$$x_2'(t) = 0$$

primero hallaremos $x_1(t)$

$$-\frac{x_1'(t)}{x_1(t)} + \frac{2x_1'(t)}{2x_1(t)-1} = \frac{1}{\mu}$$

integrando con respecto a t en ambos miembros

$$-\ln x_1(t) + \ln(2x_1(t)-1) = \frac{t}{\mu}$$

usando propiedades de ln

$$\ln\left(\frac{2x_1(t)-1}{x_1(t)}\right) = \frac{t}{\mu}$$

por definición de ln tenemos

$$\epsilon^{\frac{t}{\mu}} = \frac{2x_1(t)-1}{x_1(t)}$$

luego,

$$2 + \frac{1}{x_1(t)} = e^{\frac{t}{\mu}}$$

entonces,

$$\frac{1}{x_1(t)} = \left(e^{\frac{t}{\mu}} - 2\right)$$

Por tanto,

$$x_1(t) = \frac{1}{e^{\frac{t}{\mu}} - 2}$$

Por otro lado, $x_2(t) = c$ tal que $0 < c < 1$.
Analizando la trayectoria cuando $t \to \infty$ para hallar un punto KKT del problema de minimización.

$$\lim_{t\to\infty} \frac{1}{e^{\frac{t}{\mu}} - 2} = 0 \quad \wedge \quad \lim_{t\to\infty} c = c$$

concluimos que $\bar{x} = (0, c)$ es un punto KKT del problema. Desde que,

$$\frac{\partial f(0)}{\partial x_1} = 1 > 0$$

Usando el resultado del Teorema 4.1 del capítulo de Optimización Cuasi-convexa, concluimos que el punto $\bar{x} = (0, c)$ minimiza la función en el octante no negativo.

Materiales y Métodos

Los materiales que se emplearon en el desarrollo de la tesis, son de ejecución y de impresión:

Entre los materiales de ejecución tenemos:

- Hojas bulki y bond
- Cuadernos, cuadernillos y lapiceros.
- Copias
- Diskette, USB y discos compactos para computadora
- Revistas especializadas
- Fotocopias de artículos y textos bibliográficos
- Internet.

  Los materiales de impresión utilizados para presentar los resultados de la tesis son los siguientes:

- Material humano para la digitación
- Procesador de texto Latex
- Programa CorelDRAW X3
- Hojas bond
- Cartuchos de impresora
- Thoner
- Material de anillado

En el desarrollo de la tesis, se utilizaron los métodos que acontinuación indicamos:

- El método comparativo, para contrastar los resultados del método proximal continuo en funciones convexas con las funciones cuasi-convexas.
- El método inductivo - deductivo con el propósito de establecer conclusiones y generalizar los resultados de funciones convexas para funciones cuasi-convexas.
- El método analítico para analizar el estudio del metodo proximal continuo y su incidencia en la teoría económica de decisión.

Resultados

En este trabajo se ha presentado un estudio sobre el Método Proximal Continuo para Minimizar Funciones Cuasi-convexas y además se ha dado una aplicación en la Teoría Económica de Decisión, los resultados más importantes son:

1. Las funciones cuasi-convexas tienen aplicaciones en la Teoría Económica de Decisión, como por ejemplo, el problema de decisión del consumidor que puede ser representado como un problema de minimización cuasi-convexa.

2. El Sistema Dinámico Lotka- Volterra es un método proximal continuo para minimizar funciones convexas. Además, permite el estudio de la minimizacion cuasi-convexa.

3. La solución del Sistema Dinámico converge a un punto candidato a solución del problema de minimización cuasi-convexa.

Discusiones

1. La bibliografía existente sobre funciones cuasi-convexas, muestran definiciones, proposiciones, teoremas y corolarios, pero no dan un ejemplo claro de minimización cuasi-convexa, en la Tesis hemos presentado un ejemplo de aplicación de minimización cuasi-convexa en teoría económica de decisión.

2. La mayoría de los métodos para minimizar funciones cuasi-convexas son discretos o iterativos, por ejemplo, el método del punto proximal, este se encuentra desarrollado en el artículo de Souza, S.S., Oliveira, P.R., da Cruz Neto, J.X., y Soubeyran, A. [28]. Sin embargo, en esta investigación, presentamos un Sistema Dinámico como un método continuo para encontrar un punto candidato a solución del problema planteado.

3. Las múltiples aplicaciones y las escasas investigaciones existentes, hacen necesario el estudio de la minimización de funciones cuasi-convexas.

Conclusiones

En este trabajo hemos recopilado resultados básicos de la teoría de funciones cuasi-convexas que no se encontraban de manera unificada en ningún libro o manuscrito. Además, presentamos un Sistema Dinámico Lotka-Volterra como un método proximal continuo para minimizar funciones convexas y cuasi-convexas en el octante no negativo. A manera de conclusión podemos decir que:

1. Al ser la teoría de convexidad y cuasi-convexidad un instrumento importante en muchos campos de la matemática, es indispensable hacer un estudio, de toda la teoría realizada en este trabajo, sobre espacios mas generales como: los espacios métricos y topológicos, y comparar los resultados obtenidos y las aplicaciones que se pueden dar en cada uno de los casos.

2. El Sistema Dinámico Lotka-Volterra aparece como una nueva propuesta para minimizar funciones convexas y cuasi-convexas, ya que este sistema puede ser visto como un método proximal continuo y de esta manera puede mejorar a los métodos iterativos ya conocidos.

3. La trayectoria del Sistema Dinámico cuando la función es convexa, converge a la solución del problema de minimización convexa.

4. Los resultados del caso convexo se extienden para el caso cuasi-convexo, por tanto, obtenemos un nuevo resultado, la convergencia de la trayectoria del Sistema Dinámico a un punto candidato a solución del problema de minimización cuasi-convexa.

5. El Sistema Dinámico aparece como un nuevo método para solucionar problemas de minimización convexa y cuasi-convexa en el octante no negativo, se debe realizar un estudio detallado para saber si este Sistema también podrá solucionar problemas de minimización con restricciones de igualdad y desigualdad.

6. Finalmente, dejamos un problema abierto para un futuro trabajo de tesis, el estudio de la convergencia de la solucion del Sistema Dinámico a un punto solución del problema de minimización cuasi-convexa.

Bibliografía

[1] Aliprantes, C.D., Brown, D.J., and Birkinshaw, O., *Existence and Optimality of Competitive Equilibrium.* Berlín, 1990.

[2] Arrow, K., and Debreu, G., *Existence of an Equilibrium for a Competitive Economy,* Econometría, 27, pp. 265-290, 1954.

[3] Arrow, K., and Enthoven , A., *Quasi-concave programming,* Econometría, 29, pp. 779-800, 1961.

[4] Attouch, H., and Teboulle, M., *Regularized Lotka-Volterra Dynamical System as Continuous proximal - Like Method in Optimization,* Journal of Optimization Theory and Applications. V. 121, No 3, pp. 541-570, 2004.

[5] Auslender, A., Teboulle, M., and Ben-Tiba, S., *Interior Proximal and Multiplier Methods Based on Second-Order Homogeneous Kernels,* Mathematics of Operations Research, Vol. 24, pp. 645-668, 1999.

[6] Avriel, M., *Nonlinear Programming: Analysis and Methods,* Prentice Hall, Englewood Cliffs, New Yersey, 1976.

[7] Avriel, M., and Schaible, S., *Second order characterization of pseudoconvex functions,* Math Programming 14, pp. 170-185, 1978.

[8] Barron, N., and Lui. W., *Calculus of Variation in* l^{∞}, Applied Math.Optim.V.35, pp. 237-263, 1997.

[9] Barten, A.P., and Bohm, Y., *Consumer Theory,* K.J. Arrrow and M.D. Intriligator (Eds.), Vol II, 1982.

[10] Bartle, R., *A Elementos de Análisis,* John Wiley & Sons, New York, 1964.

[11] Bazaraa, M.S., Sherald, H.D., and Shetty, C.M., *Nonlinear Programming: Theory and Algorithms, Second Edition,* , John Wiley & Sons, 1993.

[12] Berkovitz, L., *Convexity and Optimization in* R^n, John Wiley & Sons, Inc. New York, 2002.

[13] Canales, P., *Convexidad y Aplicaciones,* Sociedad de Matemática Peruana, XXII Coloquio, 19 - 23 de Julio del 2004.

[14] Crouzeix, J., and Ferland, J., *Criteria for quasiconvexity and pseudo-comvexity and their relationships, in Generalized Concavity in Optimization and Economics,* Edited by S. Schaible and W. Ziemba, Academic Press, New York, 1981.

[15] Cunha, F.G., da Cruz Neto, X.J., and Oliveira, P.R., *A Proximal Point Algorithm with ϕ-Divergente to Quasiconvex Programming,* Artículo sometido para publicación, 2007.

[16] Debreu, G., *Representation of Preference Ordering by a Numerical Function,* In .Tharall, R, Coombs, C. and Davies, R. (Eds.) Decision Processes New York, John Wiley, 1954.

[17] Diewert, W., Avriel, M., and Zang, I., *Nine kinds of quasiconcavity and concavity,* Department of Economics, University of British Columbia, Vancouver, N., Canada, 1977.

[18] Diewert, W., Avriel, M., and Zang, I., *Nine kinds of quasiconcavity and concavity,* J, Econ. Theory 25, pp. 397-420, 1981.

[19] Elkin, R., *Convergence theorems for Gauss-Seidel and other minimization algorithms,* University of Maryland, College Park, 1968.

[20] Fenchel, W., *Conex cones, sets and functions, Mimeographed lecture notes,* Princeton University, Princeton, New Jersey, 1951.

[21] Ferland, J., *Quasi-convex and pseudo-convex functions on sold convex sets,* Ph.D. dissertation, Standford University, Standforf, California, 1971.

[22] De Finetti, B., *Sulle stratificazioni convesse,* Ann. Math. Pura Appl.30, pp. 173-183,1949.

[23] Gromicho, J., *Quasiconvex Optimization and Location Theory,* Kluwer Academic Publishers Dordrecht, The Netherlands, 1998.

[24] Halmos, P., *Teoría Intuitiva de Conjuntos,* C.E.C.SA, 1973.

[25] Herden, G., *On the Existence of Utility Functions,* Mathematical Social Sciences 17, pp. 297-313, 1989.

[26] Izmailov, A., and Solodov, M., *Otimização -volume 1,* IMPA, Rio de Janeiro, Brazil, 2005.

[27] Katzner, D., Static Demand Theory, Macmillan, New York, 1970.

[28] Kiwiel, K., *Convergence and Efficiency of Subgradient Methods for Quasiconvex Minimizator,* Mathematical Programming, 90, pp. 1-25, 2001.

[29] Kreyszig, E., *Introductory Functional Analysis with Applications,* John Wiley & Sons.Inc.Canadá, 1978.

[30] Lima, E., *Curso de Análise, volumen I,* , Instituto de Matemática Pura y Aplicada, 1991.

[31] Lima, E. *Curso de Análise, volumen II,* Instituto de Matemática Pura y Aplicada, 1989.

[32] Martos, B., *Subdefinite matrices and quadratic forms,* Siam J. Appl. Math, 17, pp. 1215-1223, 1969.

[33] Martos, B., *Nonlinear Programming Theory and Methods,* North-Holland, Amsterdam, 1975.

[34] Mas-Collel A., Whinston, M.D., and Green, J.R., *Microeconomic Theory,* New York, 1995.

[35] Minoux, M., *Mathematical Programing: Theory and Algorithms,* Wiley- Interscience Publication, 1986.

[36] Newman, D., *Location of the maximum on unimodal surfaces,* J, Assoc. Comp. Mach. 12, pp. 395-398, 1965.

[37] Papa Quiroz, E.A., Quispe Cardenas, E.M., and Oliveira, P.R., *Steepest Descent Method with a Generalized Armijo Search for Quasiconvex Functions on Riemannian Manifolds,* por aparecer en Journal of Mathematical Analysis and Applications, 2007.

[38] Papa Quiroz, E.A., and Oliveira, P.R., *Proximal Point Methods for Quasiconvex Functions With Bregman Distances on Hadamard Manifolds,* publicado en Journal of Convex Analysis, 2009.

[39] Papa Quiroz, E.A., and Oliveira, P.R., *Clasiccal and Logarithmic- Quadratic Proximal Methods for Quasiconvex Minimization,* artículo sometido para publicación, 2006.

[40] Ponstein, J., *Seven kinds of convexity,* SIAM Rev 9, pp. 115-119, 1967.

[41] Rockafellar, R.T., *Convex Analysis,* Princeton University Press, Princeton, New Yersey, 1970.

[42] Rubinsten, A., *Lectures on Microeconomy Theory,* Oxford University Press. New York, 2005.

[43] Rudin,W., *Principios de Análisis Matemático,* McGraw-Hill, New York, 1976.

[44] Schaible, S., *Beitrage zur quasikonvexen Programming,* Doctoral Dissertation, Koln, Germany, 1971.

[45] Souza, S.S., Oliveira,P.R., da Cruz Neto, J.X., and Soubeyran, A., *A Proximal Point Algorithm with Bregman Distances for Quasiconvex Optimization over the Positive Orthant,* aceptado para publicación en EJOR 2010.

[46] Suppes, Patrick., *Teoría Axiomática de Conjuntos,* Editorial Norma, Colombia, 1968.

[47] Takayama, A., *Mathematical Economics, 2nd Edition,* Cambridge, Univ. Press, 1995.

[48] Thompson, W., and Parke, D., *Some properties of generalized concave functions,* Oper. Res 21, pp. 305-313, 1973.

[49] Villar, A., *Curso de Microeconomía Avanzada,* Bosch, Barcelona, 1996.

[50] Wilde, D., *Optimum Seeking Methods,* Prentice Hall, Englewood Cliffs, New Jersey, 1964.