FUNCIONES INVEXAS DIFERENCIABLES Y EL TEOREMA DE KARUSH-KUHN-TUCKER

CAROLINA VESGA MANTILLA

Universidad Industrial de Santander
Facultad de Ciencias
Escuela de Matemáticas
Bucaramanga
2007

FUNCIONES INVEXAS DIFERENCIABLES Y EL TEOREMA DE KARUSH-KUHN-TUCKER

CAROLINA VESGA MANTILLA

Trabajo de grado presentado como
requisito parcial para optar al título de

Licenciada en Matemáticas

Director
Élder Jesús Villamizar Roa

Universidad Industrial de Santander
Facultad de Ciencias
Escuela de Matemáticas
Licenciatura en Matemáticas
Bucaramanga
2007

A Dios,
a mis hijos,
a la felicidad de mi madre
y a todo aquel que, como el ave Fénix,
puede surgir de las cenizas.

Agradecimientos

Agradezco muy especialmente a:

Dios, que me ha dado la vida y con sus pruebas he podido crecer y madurar.

La Santísima Virgen María, por acompañarme e interceder por mí.

José Félix Andrés, Juan Pablo Angel y Francisco Javier, mis tres grandes retoños, que con su existencia cambiaron mi vida y filosofía.

Carmen Felisa Mantilla, mi madre, que gracias a su ejemplo, esfuerzo, apoyo e inmensos sacrificios estoy aquí.

José de Jesús Vesga, mi padre, quien ha sido un buen ejemplo para mí y espero con esto recompensar un poco su esfuerzo y responsabilidad, y además recuperar su confianza y cariño.

Maritza y Yamid, mis hermanos, por su apoyo y ayuda después de todo.

Karen Julieth y Stephany Dayanna, mis sobrinas, por su perdón, respeto y cariño.

La Universidad Industrial de Santander, por sus grandes oportunidades.

El profesor Élder Jesús Villamizar Roa, mi director, por su gran colaboración, paciencia y dedicación en la realización del presente trabajo.

Los profesores Edilberto Reyes, Rafael Isaac y Sonia Sabogal quienes durante el tiempo que compartimos supieron ganarse mi admiración, respeto y cariño.

Adriana's, Andrea, Angela, Belky, Fercho, Gis, Ivonne, Juan, Liliana, Lina, Martha, Melissa, Nico, Rosita, Sandra, Sergio y Walter, mis grandes amigos, por su inigualable amistad.

Demás compañeros y familiares que también han aportado algo a mi vida y en la realización de este proyecto.

A todos ellos, porque nunca dejaron de confiar en mí, GRACIAS.

TITLE: FUNTIONS INVEXAS DIFERENTIABLES AND THE THEOREM OF KARUSH-KUHN-TUCKER[1]
AUTHOR: CAROLINA VESGA MANTILLA[2]

KEY WORDS: Convexity, Invexidad, Karush-Kuhn-Tucker.

DESCRIPTION

In 1980 appeared the definition of *Invexa Function*, which generalizes the classic definition of Convex Function. After this discovery, several studies were made about the mathematical theory of this new class of functions in order to guarantee the existence of solutions in optimization problems.

The principal aim of this work is to analyze this new class of functions based in the papers: *M.A. Hanson. On sufficiency of the Kunh-Tucker conditions. Journal of Mathematical Analysis and Applications*, 80, (1981), 545-550 and *Differentiable invexa functions and the Karush-Kuhn-Tucker theorem. (Portuguese) TEMA Tend. Mat. Apl. Comput. 7 (2006), no.* 1, 53–61. 90C46 (26B25).

This work is organized as follows. In the first chapter, is given a review about some facts of differential calculus, topology in $\mathbb{R}^n$, and the principal aspects of the convex function theory, including, convex sets, convex functions, properties, optimization of convex functions, examples, etc. In Chapter 2, firstly is shown the definition of invexa function; moreover, it is proved several properties of this class of functions and finally, it is established a relationship between invexa function and convex functions with another classes o functions which are called as pseudo-convex functions and almost-convex funtion. In Third Chapter, it is analyzed the Karush-Kuhn-Tucker conditions applied in the context of invexa functions, generalizing the classical Karush-Kuhn-Tucker theorem for solve optimization problems whose minimizing function and their constrains are convex functions.

[1] Thesis

[2] FACULTY OF SCIENCES, LICENCIATURA EN MATEMÁTICAS.
DIRECTOR Élder Jesús Villamizar Roa.

TITULO: FUNCIONES INVEXAS DIFERENCIABLES Y EL TEOREMA DE KARUSH-KUHN-TUCKER[1]
AUTOR: CAROLINA VESGA MANTILLA[2]

PALABRAS CLAVES: Convexidad, Invexidad, Karush-Kuhn-Tucker.

DESCRIPCIÓN

En 1980 apareció la definición de *Función Invexa*, que generaliza la definición clásica de Función Convexa. Después de éste descubrimiento, varios estudios fueros hechos sobre la teoría matemática de ésta nueva clase de funciones a fin de garantizar la existencia de soluciones a problemas de optimización.

El fin principal de éste trabajo es analizar ésta nueva clase de funciones basada en los escritos: *M.A. Hanson. Sobre suficiencia de las condiciones de Kunh-Tucker. Revista de Análisis Matemático y aplicaciones*, 80, (1981), 545-550 and *Funciones invexas diferenciables y el teorema de Karush-Kuhn-Tucker. (Portugués) TEMA Tend. Mat. Apl. Comput. 7 (2006), no.* 1, 53–61. 90C46 (26B25).

Este trabajo es organizado como sigue. En el primer capítulo, se dá un revición sobre algunos datos del cálculo diferencial, topología en $\mathbb{R}^n$, y los principales aspectos de la teoría de función convexa, incluyendo, conjuntos convexos, funciones convexas, propiedades, optimización de funciones convexas, ejemplos, etc. En el Capítulo 2, primeramente se muestra la definición de función invexa; además, se muestran varias propiedades de ésta clase de funciones y finalmente, se establece una relación entre función invexa y funciones convexas con otra clase de funciones las cuales son llamadas como funciones seudoconvexas y función casiconvexa. En el tercer capítulo, se analizan las condiciones del Karush-Kuhn-Tucker aplicadas en el contexto de funciones invexas, generalizando el teorema clásico de Karush-Kuhn-Tucker para resolver problemas de optimización cuya función a minimizar y sus restricciones son funciones convexas.

[1] Tesis
[2] FACULTAD DE CIENCIAS, LICENCIATURA EN MATEMÁTICAS.
DIRECTOR Élder Jesús Villamizar Roa.

CONTENIDO

Introducción

La importancia que representan las funciones convexas en el análisis de criterios de optimalidad, es considerable, por tanto, definir una nueva clase de funciones que generalicen el concepto de función convexa, y con ello establecer una teoría matemática relativa a la resolución de un problema de optimización, es un aspecto escencial para la resolución de problemas de programación matemática: minimización de funciones definidas en $\mathbb{R}^n$ sujetas a un número finito de restricciones de igualdad y/o desigualdad.

En 1980 M.A. Hanson [7] introdujo una nueva clase de funciones mayor que la clase de funciones convexas, a saber, las funciones invexas. Explícitamente, dado un conjunto abierto Ω de $\mathbb{R}^n$, Hanson [7] consideró funciones diferenciables f para las cuales existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tales que para todo $x, \mu \in \Omega$,

$$f(x) - f(\mu) \geq \nabla f(\mu)\eta(x, \mu), \tag{0.1}$$

y las llamó funciones invexas diferenciables. Una propiedad importante de esta nueva clase de funciones es que éstas tienen la característica de que todo punto estacionario es un punto de mínimo global. A partir de este descubrimiento, varios estudios han sido realizados intentando utilizar ésta clase de funciones para garantizar la optimalidad en problemas de programación matemática, esto es, problemas como:

$$\begin{cases} \text{Minimizar } f_0(x) \\ \text{sujeto a } f_j(x) \leq 0, \text{ para } j = 1, \ldots, m. \end{cases} \tag{0.2}$$

donde f_0, $f_j : \Omega \to \mathbb{R}$ son funciones invexas diferenciables que satisfacen (0.1), en lugar de las condiciones usuales de convexidad.

En este trabajo presentamos una descripción de la clase de funciones invexas, estableciendo su utilidad en la teoría de la programación matemática, analizando cómo las condiciones del teorema clásico de Karush-Kuhn-Tucker garantizan la optimalidad global en el caso donde las funciones involucradas son invexas en lugar de convexas diferenciables.

En primer lugar presentaremos un resumen de ciertos aspectos del cálculo diferencial y de la topología del espacio $\mathbb{R}^n$ que consideramos importantes para desarrollar este trabajo; así mismo discutiremos los aspectos escenciales de la teoría de las funciones convexas, y la caracterización de máximos y mínimos, revisando las condiciones necesarias y suficientes para la existencia de solución de problemas de minimización.

En el segundo capítulo se estudian las nociones de invexidad, pseudoconvexidad, pseudoinvexidad, casiconvexidad, casinvexidad y se discuten algunas relaciones entre ellas.

Finalmente, en el tercer capítulo se analizan las condiciones de Karush-Kuhn-Tucker que garantizan optimilidad global cuando la función a ser minimizada y las funciones que definen las restricciones, son invexas.

El contenido de esta monografía se basa principalmente en la recopilación de resultados que se encuentran en la bibliografía citada al final del trabajo. En particular, los conceptos y resultados relativos a las nociones de convexidad generalizada son basados en el artículo de J. Cervelati y M.A. Rojas-Medar [5]. Agradecemos a los autores del material bibliográfico referenciado por los aportes respectivos.

CAPÍTULO 1

ASPECTOS BÁSICOS DE LA TEORÍA DE OPTIMIZACIÓN

El objetivo del presente capítulo es establecer la notación que será utilizada a lo largo de esta monografía; así mismo, presentaremos un resumen de los principales aspectos de la Teoría Básica de Optimización, comunmente abordados en cursos de cálculo de una y varias variables.

1.1. Fundamentos de la Topología de $\mathbb{R}^n$

Sea n un número natural. El espacio *euclidiano n-dimensional* $\mathbb{R}^n$ está definido como el conjunto de n-úplas $(x_1, \ldots, x_n)$ tales que $x_i \in \mathbb{R}$, para todo $i = 1, \ldots, n$. Como es habitual, $\mathbb{R}^1$ se denota simplemente por $\mathbb{R}$.

Dados dos puntos $x = (x_1, \ldots, x_n)$ y $y = (y_1, \ldots, y_n)$ en $\mathbb{R}^n$, consideramos la distancia entre x y y definida por:

$$|x - y| = \left(\sum_{j=1}^{n} (x_j - y_j)^2 \right)^{\frac{1}{2}}.$$

El conjunto de todos los puntos de $\mathbb{R}^n$ que están a una distancia menor que $r > 0$ de un punto fijo $x_0 \in \mathbb{R}^n$, es llamado *bola abierta con centro en x_0 y radio r* y será denotado

por $B(x_0;r)$, esto es,
$$B(x_0;r) = \{x \in \mathbb{R}^n : |x - x_0| < r\}.$$
Un subconjunto Ω de $\mathbb{R}^n$ es *abierto* si, dado cualquier $x_0 \in \Omega$, existe una bola abierta centrada en x_0 enteramente contenida en Ω. Dado un subconjunto Ω de $\mathbb{R}^n$, el conjunto de todos los puntos $x_0 \in \Omega$ tales que existe $r > 0$ donde $B(x_0, r) \subset \Omega$ se denomina interior de Ω.
Un subconjunto F de $\mathbb{R}^n$ es *cerrado* si su complemento
$$\mathbb{R}^n - F = \{x \in \mathbb{R}^n : x \notin F\}$$
es abierto.

Recordemos la definición de conjunto acotado, conjunto compacto en $\mathbb{R}^n$ y punto de acumulación.
Un conjunto Ω de $\mathbb{R}^n$ se dice acotado si existe $K > 0$ y $x_0 \in \mathbb{R}^n$ tal que $\Omega \subseteq B(x_0;K)$.
Un conjunto Ω de $\mathbb{R}^n$ cerrado y acotado se dice compacto en $\mathbb{R}^n$.
Dado un espacio métrico (M,d), $\Omega \subseteq M$ y $a \in M$, se dice que a es punto de acumulación de Ω, si para todo $r > 0$, $(B(a,r) - \{a\}) \cap \Omega \neq 0$.

Suponemos que el lector está familiarizado con las propiedades elementales de funciones $f : \Omega \to \mathbb{R}$, donde $\Omega \subseteq \mathbb{R}^n$, en particular con los conceptos de continuidad y diferenciabilidad.

Definición 1.1. *(Vector gradiente).*
Sea Ω un abierto de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función diferenciable con derivada continua. El gradiente de f en un punto $x = (x_1, \ldots, x_n)$ de Ω se define como el vector:
$$\nabla f(x_1, \ldots, x_n) = \Big(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\Big).$$

Definición 1.2. *(Matriz Hessiana).*
Sea Ω un abierto de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función dos veces derivable con derivadas continuas. La matriz Hessiana de f, $H_f(x)$, en un punto $x = (x_1, \ldots, x_n)$ de Ω se define como la matriz de las segundas derivadas parciales de f evaluadas en el punto x; es decir:
$$H_f(x) = \begin{pmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & & & \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}.$$

Note que sobre las hipótesis de f, se tiene que $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j} = \dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ y así, la matriz Hessiana es simétrica por definición.

1.2. Caracterización de Máximos y Mínimos

Como se estudia en el cálculo de una variable, toda función continua en un intervalo cerrado $[a, b]$ tiene máximos y mínimos globales; esto es, existen $x_m, x_M \in [a, b]$ tales que $f(x_M) \geq f(x) \geq f(x_m)$ para todo $x \in [a, b]$.

La conclusión de este resultado no necesariamente se cumple si el intervalo no es cerrado o si la función no es continua en él. El objetivo de esta subsección es recordar las condiciones necesarias y suficientes para que una función f de $\mathbb{R}^n$ en $\mathbb{R}$ posea un mínimo (o máximo) en un subconjunto Ω de $\mathbb{R}^n$.

Definición 1.3. *(Punto estacionario).*
Sea Ω un conjunto abierto no vacío de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función diferenciable. Se dice que $u \in \Omega$ es un punto estacionario de f si $\nabla f(u) = 0$.

Definición 1.4. *(Punto de mínimo global, Punto de mínimo relativo).*
Sea Ω un conjunto abierto no vacío de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función diferenciable. Se dice que $u \in \Omega$ es un punto de mínimo global de f (en relación a Ω) si $f(u) \leq f(x)$, para todo $x \in \Omega$. Si $f(u) \leq f(x)$ para todo x en una bola $B(u; r)$, se dice que $u \in \Omega$ es un punto de mínimo relativo de f.

Definición 1.5. *(Punto de máximo global, Punto de máximo relativo).*
Sea Ω un conjunto abierto no vacío de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función diferenciable. Se dice que $u \in \Omega$ es un punto de máximo global de f (en relación a Ω) si $f(u) \geq f(x)$, para todo $x \in \Omega$. Si $f(u) \geq f(x)$ para todo x en una bola $B(u; r)$, se dice que $u \in \Omega$ es un punto de máximo relativo de f.

La primera pregunta que surge al tratar un problema de optimización sin restricciones es saber si existe la solución óptima. Para el caso en el cual la función está definida sobre un compacto, el Teorema de Weierstrass nos da la respuesta.

Teorema 1.6. *(Teorema de Weierstrass)[1]. Si $f : \Omega \to \mathbb{R}$, donde Ω es un compacto de $\mathbb{R}^n$, es continua, entonces f alcanza su valor máximo y su valor mínimo en Ω, esto es, existen $x_m, x_M \in \Omega$ tales que $f(x_m) \leq f(x) \leq f(x_M)$, para todo $x \in \Omega$.*

1.2.1. Condiciones necesarias para la minimización sin restricciones

Las condiciones necesarias para la minimización sin restricciones vienen dadas en los siguientes resultados.

Definición 1.7. *Dado $x \in \Omega$ se dice que un punto $d \in \mathbb{R}^n$ es una dirección factible en x si existe algún $\bar{\alpha} \in \mathbb{R}^+$ tal que $x + \alpha d \in \Omega$, para todo $\alpha \in [0, \bar{\alpha}]$.*

La Definición 1.7 nos dice que si d es una dirección factible, podemos hacer un desplazamiento desde x en la dirección d y encontrar un nuevo punto que pertenece a Ω.

Proposición 1.8. *(Condiciones necesarias de primer orden).*
Sea $\Omega \subseteq \mathbb{R}^n$ y $f : \Omega \to \mathbb{R}$ una función de clase C^1, esto es, f tiene primeras derivadas continuas. Si x^* es un punto de mínimo relativo de f en Ω, entonces para todo $d \in \mathbb{R}^n$ que sea una dirección factible en x^*, se tiene que $\nabla^T f(x^*)d \geq 0$ (recordemos que $\nabla^T f(x^*)$ es la transpuesta del gradiente de f en x^*).

Demostración.
Para todo $\alpha \in [0, \bar{\alpha}]$, el punto $x(\alpha) = x^* + \alpha d \in \Omega$. Definamos la función $g(\alpha) = f(x(\alpha))$, para todo $\alpha \in [0, \bar{\alpha}]$. Entonces g tiene un mínimo relativo en $\alpha = 0$. Esto se sigue debido a que $g(0) = f(g(0)) = f(x^*) \leq f(x)$ para toda x en una cierta bola con centro x^* ya que f tiene un mínimo relativo en x^*; y así $g(0) \leq g(\alpha)$, $\forall \alpha \in [0, \tilde{\alpha}]$, $\tilde{\alpha} \leq \bar{\alpha}$. Haciendo una aproximación por Taylor, tenemos que $g(\alpha) - g(0) = g'(0)\alpha + 0(\alpha)$, donde $0(\alpha)$ es una magnitud tal que

$$\lim_{\alpha \to 0} 0(\alpha) = 0.$$

Si $g'(0) < 0$, entonces para valores suficientemente pequeños de α, el lado derecho de la ecuación será negativo y, por tanto $g(\alpha) - g(0) < 0$, lo cual contradice la naturaleza minimal de $g(0)$. Así, $g'(0) = \nabla^T f(x^*)d \geq 0$. □

Un caso particular bastante importante es aquel cuando x^* está en el interior de Ω. En este caso hay direcciones factibles que salen en todas direcciones desde x^*, y por tanto, $\nabla^T f(x^*)d \geq 0$, para todo $d \in \mathbb{R}^n$. Esto implica que $\nabla f(x^*) = 0$.

Corolario 1.9. *Sea $\Omega \subseteq \mathbb{R}^n$, y sea $f \in C^1$ una función en Ω. Si x^* es un punto de mínimo de f en Ω, y x^* es un punto interior de Ω, entonces $\nabla f(x^*) = 0$.*

Proposición 1.10. *(Condiciones necesarias de segundo orden).*
Sea $\Omega \subseteq \mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ una función de clase C^2, esto es, f tiene segundas derivadas continuas. Si x^* es un punto de mínimo relativo de f en Ω, entonces para todo $d \in \mathbb{R}^n$ que sea dirección factible en x^*, resulta:

i). $\nabla^T f(x^*)d \geq 0$.

ii). *Si* $\nabla^T f(x^*)d = 0$, *entonces* $d^T\nabla^2 f(x^*)d \geq 0$, *donde* $\nabla^2 f(x^*)$ *es la Hessiana* $H_f(x^*)$ *de* f *evaluada en* x^**.*

Demostración.
La primera condición es precisamente la proposición anterior, y la segunda solo se aplica si $\nabla^T f(x^*)d = 0$. En este caso, introduciendo $x(\alpha) = x^* + \alpha d$, y $g(\alpha) = f(x(\alpha))$, donde $\alpha \in [0, \bar{\alpha}]$ con $\bar{\alpha}$ un real positivo, realizamos una aproximación de Taylor de segundo orden y utilizando la hipótesis de que $g'(0) = 0$, resulta

$$g(\alpha) - g(0) = \frac{1}{2}g''(0)\alpha^2 + 0(\alpha^2), \tag{1.1}$$

donde $0(\alpha^2)$ es una magnitud tal que

$$\lim_{\alpha\to 0} 0(\alpha^2) = 0.$$

Si $g''(0) < 0$, el lado derecho de (1.1) es negativo para valores de α suficientemente pequeños, lo que contradice la naturaleza de mínimo relativo de $g(0)$, luego

$$g''(0) = d^T H_f(x^*)d \geq 0.$$

□

De la Proposición 1.10 se deriva un resultado clásico del cálculo, el cual establece las condiciones necesarias de segundo orden cuando x^* está en el interior del conjunto Ω.

Teorema 1.11. *Sea* $f \in C^2$, *sea* x^* *un punto en el interior de* Ω, *y supóngase que* x^* *es un punto de mínimo relativo en* Ω *de* f*. Entonces,*

i). $\nabla f(x^*) = 0$.

ii). *Para todo* $d \in \mathbb{R}^n$, $d^T H_f(x^*)d \geq 0$.

Demostración.
La demostración de la parte *i)*. se sigue de la proposición que establece las condiciones necesarias de primer orden.

La segunda parte nos dice que la matriz Hessiana de f es positivamente semidefinida.

Una aproximación de Taylor de segundo orden para la función f de n variables alrededor de x^* puede expresarse como:

$$f(x^* + \Delta x) = f(x^*) + \nabla^T f(x^*)\Delta x + \frac{1}{2}\Delta x^T H_f(x^*)\Delta x + 0(\Delta x^3),$$

donde $H_f(x^*)$ es la matriz Hessiana de f evaluada en el punto x^* y $0(\Delta x^3)$ representa la expansión de términos de tercer orden. Si ahora hacemos $f(x^* + \Delta x) - f(x^*) = \Delta f$, tendremos

$$\Delta f = \nabla^T f(x^*)\Delta x + \frac{1}{2}\Delta^T x H_f(x^*)\Delta x + 0(\Delta x^3).$$

Por la primera parte del teorema, $\nabla f(x^*) = 0$. Despreciando la expansión de términos de tercer orden, se tiene que

$$\Delta f = \frac{1}{2}\Delta x^T H_f(x^*)\Delta x.$$

Como x^* es punto de mínimo, se tiene que $\Delta f \geq 0$ para cualquier desviación Δx, por lo tanto la matriz Hessiana de $f(x^*)$ es positivamente semidefinida. $\square$

1.2.2. Condiciones suficientes para la existencia de un mínimo relativo

Las condiciones suficientes para la existencia de un mínimo relativo aparecen como resultado del siguiente teorema.

Teorema 1.12. *Sea $f \in C^2$ una función definida en un conjunto $\Omega \subseteq \mathbb{R}^n$. Sea x^* un punto en el interior de Ω tal que:*

i). $\nabla f(x^*) = 0$*;*

ii). La matriz Hessiana $H_f(x^*)$ es positivamente definida.

Entonces x^* es un punto de mínimo relativo estricto de f en Ω.

Demostración.
Como la matriz Hessiana es definida positiva, existe un $a > 0$ tal que para todo $d \in \mathbb{R}^n$ $d^T H_f d \geq a|d|^2$. Así, por el Teorema de Taylor,

$$\begin{aligned} f(x^* + d) - f(x^*) &= \frac{1}{2}d^T H_f(x^*)d + 0(|d|^2) \\ &\geq \left(\frac{a}{2}\right)|d|^2 + 0(|d|^2), \end{aligned}$$

donde $0(|d|^2)$ es una magnitud tal que

$$\lim_{d \to 0} 0(|d|^2) = 0.$$

Para $|d|$ pequeña, el primer término de la derecha domina al segundo, lo que implica que ambos lados son positivos para una d pequeña; en consecuencia x^* es punto de mínimo relativo de f. $\square$

Observación 1.13. *Los resultados anteriores sintetizan el método clásico para encontrar los extremos de una función, es decir:*

i). Encontramos las n derivadas parciales de la función objetivo.

ii). Resolvemos el sistema lineal o no-lineal de ecuaciones simultáneas

$$\frac{\partial f}{\partial x_i} = 0, i = 1, \ldots, n.$$

Las soluciones de dicho sistema conforman el conjunto de puntos llamados puntos estacionarios.

iii). Calculamos las $n \times n$ segundas derivadas parciales de $f(x)$, que conforman la Hessiana.

iv). Comprobamos para cada punto estacionario las condiciones de suficiencia.

Recordemos que los puntos estacionaros son aquellos que anulan el gradiente de la función, y que pueden catalogarse como puntos de máximo, puntos de mínimo y puntos de silla.

Una de las grandes limitaciones que se presentan en el método clásico es la dificultad para resolver sistemas altamente no lineales de ecuaciones simultáneas para encontrar los puntos estacionarios; de ahí la importancia de los métodos numéricos para la solución de problemas de optimización.
Para comprobar las condiciones de suficiencia, y en particular para determinar el comportamiento de la matriz Hessiana de f, existen varios criterios, de los cuales algunos de ellos se tienen como resultado de los siguientes teoremas.

Teorema 1.14. *La matriz Hessiana es positivamente (negativamente) definida si, y sólo si, todos sus valores propios son positivos (negativos).*

Teorema 1.15. *Sea $H_f(a)$ la matriz Hessiana de f evaluada en el punto a.*

i). Si todos los valores propios $H_f(a)$ son positivos, f tiene un mínimo relativo en a.

ii). Si todos los valores propios $H_f(a)$ son negativos, f tiene un máximo relativo en a.

iii). Si $H_f(a)$ posee valores propios positivos y negativos, f tiene un punto de silla en a.

La demostración de estos dos teoremas se puede consultar en [1].
El siguiente resultado puede resultar útil a la hora de determinar la positividad de una matriz cuadrada A, y su demostración se puede consultar en [13].

Teorema 1.16. *(Criterio de Silvester).*
Sea $A = [a_{ij}]$, $i, j = 1, \ldots, n$ una matriz cuadrada de orden n y definamos A_k así:

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{bmatrix}, \quad k = 1, 2, \cdots, n.$$

La matriz A es positivamente definida si $\det(A_k) > 0$, para todo $k = 1, 2, \ldots, n$.

Ejemplo 1.17. *Considerese la matriz*

$$M = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 4 \\ 2 & 1 & 3 \end{bmatrix}.$$

Entonces M es positivamente definida, porque, según el criterio de Silvester, $a_{11} = 3 > 0$, $(a_{11} * a_{22} - a_{21} * a_{12}) = 1 > 0$, y $\det(M) = 6 > 0$.

1.3. Convexidad

El objetivo de esta sección es recordar algunos aspectos fundamentales de la teoría de las funciones convexas.

Definición 1.18. *(Conjunto convexo).*
Sea $\Omega \subseteq \mathbb{R}^n$. Se dice que Ω es un conjunto convexo si, y sólo si, para todo $x, y \in \Omega$ y $\lambda \in [0, 1]$, se tiene que $\lambda x + (1 - \lambda)y \in \Omega$.

Definición 1.19. *(Función convexa).*
Sea Ω un conjunto convexo no vacio de $\mathbb{R}^n$. Una función f definida sobre Ω es convexa si, y sólo si, para todo $x, y \in \Omega$, $\lambda \in [0, 1]$, se cumple que:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Si para todo $x, y \in \Omega, \lambda \in (0, 1)$, con $x \neq y$, se tiene que:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y),$$

se dice que f es estrictamente convexa.

Ejemplo 1.20. *Consideremos la función* $f(x_1, x_2) = x_1^2 + x_2^2$, *con* $(x_1, x_2) \in \mathbb{R}^2$. f *es estrictamente convexa. De hecho, si* $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}^2$, *y* $\lambda \in (0, 1)$, *tenemos que*

$$(\lambda - \lambda^2)[(x_1 - y_1)^2 + (x_2 - y_2)^2] > 0,$$

luego

$$(\lambda - \lambda^2)[x_1^2 - 2x_1y_1 + y_1^2 + x_2^2 - 2x_2y_2 + y_2^2] > 0,$$

y así,

$$(\lambda - \lambda^2)(x_1^2 + x_2^2) - 2(\lambda - \lambda^2)(x_1y_1 + x_2y_2) + (\lambda - \lambda^2)(y_1^2 + y_2^2) > 0.$$

De la desigualdad anterior tenemos que

$$(\lambda - \lambda^2)(x_1^2 + x_2^2) - 2\lambda(1 - \lambda)(x_1y_1 + x_2y_2) + ((1 - \lambda) - (1 - \lambda)^2)(y_1^2 + y_2^2) > 0,$$

es decir,

$$\lambda(x_1^2 + x_2^2) + (1 - \lambda)(y_1^2 + y_2^2) > \lambda^2(x_1^2 + x_2^2) + 2\lambda(1 - \lambda)(x_1y_1 + x_2y_2) + (1 - \lambda)^2(y_1^2 + y_2^2),$$

luego podemos decir que:

$$\begin{aligned}
\lambda f(x_1, x_2) + (1 - \lambda) f(y_1, y_2) &> \lambda^2(x_1^2 + x_2^2) + 2\lambda(1 - \lambda)(x_1y_1 + x_2y_2) + (1 - \lambda)^2(y_1^2 + y_2^2) \\
&= \lambda^2 x_1^2 + \lambda^2 x_2^2 + 2\lambda(1 - \lambda)x_1y_1 + 2\lambda(1 - \lambda)x_2y_2 \\
&\quad +(1 - \lambda)^2 y_1^2 + (1 - \lambda)^2 y_2^2 \\
&= [(\lambda x_1)^2 + 2(1 - \lambda)x_1\lambda y_1 + ((1 - \lambda)y_1)^2] \\
&\quad +[(\lambda x_2)^2 + 2(1 - \lambda)x_2\lambda y_2 + ((1 - \lambda)y_2)^2] \\
&= [\lambda x_1 + (1 - \lambda)y_1]^2 + [\lambda x_2 + (1 - \lambda)y_2]^2 \\
&= f(\lambda x_1 + (1 - \lambda)y_1, \lambda x_2 + (1 - \lambda)y_2) \\
&= f((\lambda x_1, \lambda x_2) + ((1 - \lambda)y_1, (1 - \lambda)y_2)) \\
&= f(\lambda(x_1, x_2) + (1 - \lambda)(y_1, y_2)),
\end{aligned}$$

y así probamos que $f(x_1, x_2) = x_1^2 + x_2^2$ *es estrictamente convexa.*

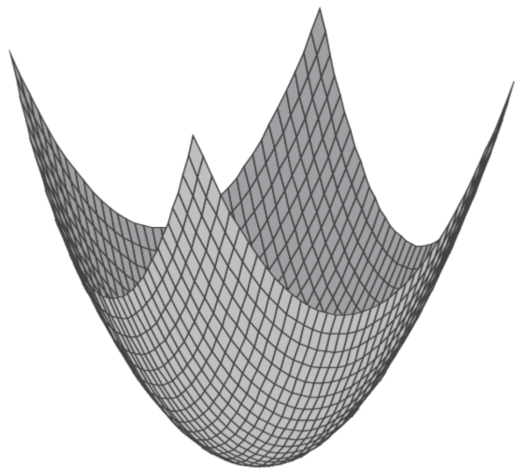


Figura 1.1: $f(x_1, x_2) = x_1^2 + x_2^2$.

Geométricamente una función de una variable es convexa, si al trazar el segmento de recta que une dos puntos de la gráfica, éste queda por encima de la parte de la gráfica que está entre los dos puntos.

Definición 1.21. *(Función cóncava).*
Sea Ω un conjunto convexo no vacio de $\mathbb{R}^n$. Una función f definida sobre Ω es cóncava si, y sólo si, para todo $x, y \in \Omega, \lambda \in [0,1]$, se cumple

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y).$$

Si para todo $x, y \in \Omega, \lambda \in (0,1)$, con $x \neq y$, se tiene que:

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y),$$

se dice que f es estrictamente cóncava.

Ejemplo 1.22. *Consideremos la función $f(x_1, x_2) = \sqrt{x_1 x_2}$, tenemos que f es cóncava en $\mathbb{R}_+^2$, donde $\mathbb{R}_+^2 = \{(x,y) \in \mathbb{R}_+^2 : x \geq 0, y \geq 0\}$. De hecho, si $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}_+^2$, y $\lambda \in [0,1]$, entonces tenemos que ver que:*

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y),$$

es decir,

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

o equivalentemente

$$\sqrt{(\lambda x_1 + (1-\lambda)y_1)(\lambda x_2 + (1-\lambda)y_2)} \geq \lambda\sqrt{x_1x_2} + (1-\lambda)\sqrt{y_1y_2}.$$

Elevando al cuadrado, obtenemos

$$(\lambda x_1 + (1-\lambda)y_1)(\lambda x_2 + (1-\lambda)y_2) \geq \lambda^2 x_1x_2 + (1-\lambda)^2 y_1y_2 + 2\lambda(1-\lambda)\sqrt{x_1x_2}\sqrt{y_1y_2}.$$

Por lo tanto,

$$\begin{aligned} & \lambda^2 x_1x_2 + \lambda(1-\lambda)x_1y_2 + \lambda(1-\lambda)y_1x_2 + (1-\lambda)^2 y_1y_2 \\ \geq \; & \lambda^2 x_1x_2 + (1-\lambda)^2 y_1y_2 + 2\lambda(1-\lambda)\sqrt{x_1x_2}\sqrt{y_1y_2}, \end{aligned}$$

luego,

$$\lambda(1-\lambda)x_1y_2 + \lambda(1-\lambda)y_1x_2 \geq 2\lambda(1-\lambda)\sqrt{x_1x_2}\sqrt{y_1y_2}.$$

Si $\lambda = 0$ ó $\lambda = 1$ esta desigualdad es cierta. Si $\lambda \in (0,1)$ entonces tenemos que

$$x_1y_2 + y_1x_2 \geq 2\sqrt{x_1x_2}\sqrt{y_1y_2}.$$

Como $(\sqrt{x_1y_2} - \sqrt{y_1x_2})^2 \geq 0$ es verdadero siempre que $x, y \in \mathbb{R}^2_+$ y $\lambda \in (0,1)$, tenemos que $f(x) = f(x_1, x_2) = \sqrt{x_1x_2}$ es cóncava en $\mathbb{R}^2_+$.

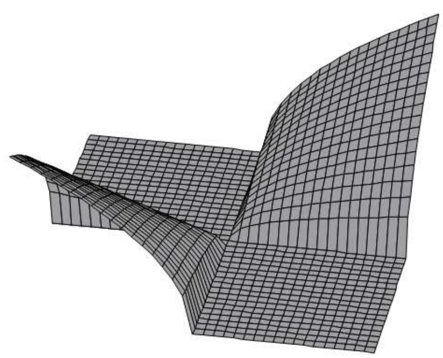


Figura 1.2: $f(x_1, x_2) = \sqrt{x_1x_2}$.

Observación 1.23. *Dadas las definiciones de concavidad y convexidad, es claro que una función f es convexa (estrictamente convexa) si, y sólo si, $-f$ es cóncava (estrictamente concava). Observamos también que la convexidad es una noción de conjunto; es decir, una función puede ser convexa en cierta región de su dominio y cóncava en otro.*

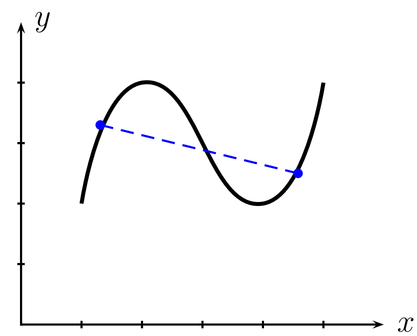


Figura 1.3: Función cóncava en una región y convexa en otra.

Teorema 1.24. *Sea $f : \Omega \to \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^n$, convexo no vacío. Si f es convexa en Ω, entonces f es continua en el interior de Ω.*

Demostración.
Sea $x \in \Omega$ un punto de acumulación, y $\{x_n\}_{n=1}^{\infty}$ una sucesión de puntos de Ω que tiende a x cuando n tiende a ∞. Sea $\epsilon > 0$ y k tal que para todo $n \geq k$, $|x_n - x| < \epsilon$. Tal k existe ya que $x_n \to x$. Sea $A = \{y \in \Omega : |y - x| = \epsilon\}$, entonces, para todo $n \geq k$ existen $y_n \in A$ y $\lambda_n \in [0, 1]$ tales que

$$x_n = \lambda_n x + (1 - \lambda_n) y_n.$$

Como $x_n \to x$ y $|y_n - x| = \epsilon$, la igualdad $x_n = \lambda_n x + (1 - \lambda_n) y_n$ implica que $\lambda_n \to 1$. Ahora, como f es convexa,

$$f(\lambda_n x + (1 - \lambda_n) y_n) \leq \lambda_n f(x) + (1 - \lambda_n) f(y_n)$$

y como $f(x_n) = f(\lambda_n x + (1 - \lambda_n) y_n)$, entonces

$$\limsup_{n \to \infty} f(x_n)^1 \leq f(x).$$

Análogamente, escogemos $z_n \in A$ y $\lambda_n \in [0, 1]$ tales que $x = \lambda_n x_n + (1 - \lambda_n) z_n$, luego

$$f(x) = f(\lambda_n x_n + (1 - \lambda_n) z_n) \leq \lambda_n f(x_n) + (1 - \lambda_n) f(z_n),$$

[1] (x_n) es una suceción de números reales, el límite superior y el límite inferior de esta suceción se define como:

$$\limsup x_n = \inf_m (\sup_{n \geq m} x_n),$$

$$\liminf x_n = \sup_m (\inf_{n \geq m} x_n).$$

Si $L = \limsup x_n = \liminf x_n$, entonces

$$L = \lim_{n \to \infty} x_n,$$

es el límite de la suceción.

por tanto

$$f(x) \leq \liminf_{n\to\infty} f(x_n).$$

De las desigualdades anteriores tenemos que

$$\limsup_{n\to\infty} f(x_n) \leq f(x) \leq \liminf_{n\to\infty} f(x_n),$$

es decir,

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

luego f es continua. $\square$

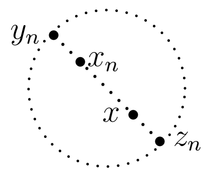


Figura 1.4: Representación gráfica de la bola.

Proposición 1.25. *Sea $f : \Omega \to \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^n$ convexo no vacío. Si f es convexa en Ω, entonces el conjunto de nivel inferior a α, $I_\alpha = \{x \in \Omega : f(x) \leq \alpha\}$, es convexo para todo $\alpha \in \mathbb{R}$. La recíproca no siempre es verdadera.*

Demostración.
Sea $x, y \in I_\alpha$. Entonces $f(x) \leq \alpha$, $f(y) \leq \alpha$. Como f es convexa, para $\lambda \in [0, 1]$ se tiene que

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

es decir, $\lambda x + (1-\lambda)y \in I_\alpha$, lo que significa que I_α es convexo. $\square$

Ahora mostraremos con un ejemplo que la recíproca de la proposición anterior no siempre es cierta.

Ejemplo 1.26. *Consideremos la función $f(x, y) = x^2y^2$. f no es cóncava en $\mathbb{R}^2_{++} = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$. Entonces la función $g(x, y) := -x^2y^2$ no es*

convexa en $\mathbb{R}^2_{++}$.
Ahora veamos el conjunto de nivel inferior a α, *para todo* $\alpha \in \mathbb{R}$,

$$I_\alpha = \{(x,y) \in \mathbb{R}^2 : g(x,y) \leq \alpha\} = \Big\{(x,y) \in \mathbb{R}^2 : y \leq \frac{(\alpha)^{\frac{1}{2}}}{x}\Big\}.$$

Si $\alpha \leq 0$ *el conjunto es convexo, si* $\alpha > 0$, $I_\alpha = \emptyset$ *que también es convexo, por tanto, tenemos una función no convexa con un conjunto de nivel inferior convexo.*

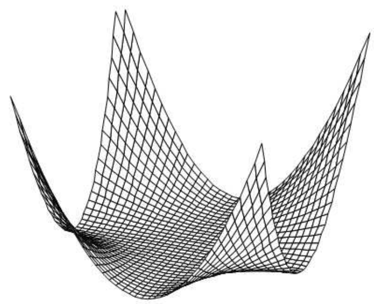


Figura 1.5: $f(x_1, x_2) = x_1^2 x_2^2$.

Ahora veamos el siguiente teorema que caracteriza las funciones convexas diferenciables.

Teorema 1.27. *Sea* Ω *un conjunto convexo no vacío de* $\mathbb{R}^n$ *y* $f : \Omega \to \mathbb{R}$ *continua en* Ω *y diferenciable con continuidad en el interior de* Ω, *entonces* f *es convexa en* Ω *si, y sólo si, para todo* x, y *en el interior de* Ω, *con* $x \neq y$, *se tiene que*

$$f(x) - f(y) \geq \nabla f(y) \cdot (x - y).$$

Demostración.
$\Rightarrow$) Consideremos el caso $n = 1$. En este caso $\nabla f(y) = f'(y)$. El caso general es análogo. Como f es convexa en Ω, entonces para todo $\lambda \in [0,1]$, con $x, y \in \Omega$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y);$$

lo cual implica,

$$\frac{f(\lambda(x-y) + y) - f(y)}{\lambda(x-y)}(x-y) \leq f(x) - f(y), x \neq y, \lambda \neq 0.$$

Puesto que f es diferenciable, tenemos que si $\lambda \to 0$,

$$f'(y)(x-y) \leq f(x) - f(y).$$

$\Leftarrow$) Ahora, supongamos que $f'(y)(x-y) \leq f(x)-f(y)$, y que f no es convexa. Entonces existen $x, y \in \Omega$, $\lambda \in [0,1]$ tal que

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y).$$

Por un argumento similar al usado anteriormente, tenemos que

$$f'(y)(x-y) > f(x) - f(y),$$

lo cual es una contradicción. Por lo tanto f es convexa. $\square$

Dada la definición de funciones convexas, fácilmente podemos mostrar la siguiente propiedad relativa al álgebra de funciones convexas.

Propiedad 1.28. ***i).*** *Si $f_1, f_2, \dots f_n$ son funciones convexas en $\Omega \subseteq \mathbb{R}^n$, y $\alpha_1, \alpha_2, \dots \alpha_n \in \mathbb{R}^+$, entonces*

$$\sum_{i=1}^{n} \alpha_i f_i$$

es convexa en Ω.
ii). *Si $a \geq 0$, y f es una función convexa en $\Omega \subseteq \mathbb{R}^n$, la función af también es convexa.*

Observación 1.29. *Es fácil ver que si f, g son funciones convexas, las funciones $f \cdot g$ y $\frac{f}{g}$ no siempre son convexas. De hecho, las funciones $f = x$ y $g = -\sqrt{x}$, $x > 0$ son funciones convexas, pero $(f \cdot g)(x) = -x^{\frac{3}{2}}$ no es convexa.*

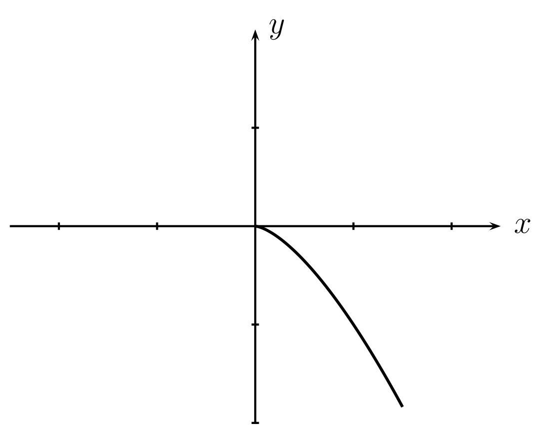


Figura 1.6: $(f \cdot g)(x) = -x^{\frac{3}{2}}$

Proposición 1.30. *Sea f una función convexa definida en el conjunto convexo $\Omega \subseteq \mathbb{R}^n$. Entonces el conjunto Γ de puntos de mínimos relativos donde f alcanza su mínimo global, es convexo, y cualquier mínimo relativo de f es un punto de mínimo global.*

Demostración.
Si f no posee mínimos relativos, el teorema se cumple por defecto. Supongamos que m_0 es el mínimo de f. Entonces, $\Gamma = \{x : f(x) \leq m_0, x \in \Omega\}$ es convexo.
Supóngase ahora que x^* es un punto de mínimo relativo de f, y que existe otro punto $y \in \Omega$ con $f(y) < f(x^*)$. Como f es convexa, para todos los puntos entre x^* y y, con $\theta \in (0, 1)$ resulta:

$$f\Big(\big((1-\theta)x^*\big) + \theta y\Big) \leq (1-\theta)f(x^*) + \theta f(y) < f(x^*),$$

lo que contradice el hecho de que x^* es un punto de mínimo relativo, luego cualquier mínimo relativo de f es un punto de mínimo global. $\square$

Teorema 1.31. *(Optimización de funciones convexas.)*
Sea Ω un conjunto convexo no vacío de $\mathbb{R}^n$ y $f : \Omega \to \mathbb{R}$ una función convexa y diferenciable con continuidad en el interior de Ω. Si x^* es un punto estacionario, entonces x^* es un punto de mínimo global.

Demostración.
Como f es convexa diferenciable con continuidad en el interior de Ω, entonces para $y \in \Omega$, por el Teorema 1.27 tenemos que

$$f(y) - f(x^*) \geq \nabla f(x^*) \cdot (y - x^*).$$

Ahora, como x^* es un punto estacionario, entonces $\nabla f(x^*) = 0$, y así $f(y) - f(x^*) \geq 0$, para todo $y \in \Omega$. Por lo tanto, x^* es un punto de mínimo global. $\square$

CAPÍTULO 2

CASICONVEXIDAD, SEUDOCONVEXIDAD E INVEXIDAD

El objetivo de este capítulo es introducir el concepto de seudoconvexidad, casiconvexidad e invexidad, como generalizaciones del concepto de convexidad; en particular, estableceremos propiedades y realizaremos un cuadro comparativo entre estos cuatro conceptos.

2.1. Funciones Casiconvexas y Casiconcavas

Existe una nueva clase de funciones, que generalizan las funciones convexas, cuya definición es motivada por el deseo de saber para qué clase de funciones la recíproca de la Proposición 1.25 es verdadera. La respuesta es encontrada en un conjunto de funciones llamadas funciones casiconvexas.
Realizar una descripción sobre este conjunto de funciones es el objetivo de esta sección.

Definición 2.1. *(Función casiconvexa).*
Sea Ω un conjunto convexo no vacío de $\mathbb{R}^n$. Una función $f : \Omega \to \mathbb{R}$ se dice casiconvexa si para todo $x, y \in \Omega$, $\lambda \in [0, 1]$, se cumple que:

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

Si para todo $x, y \in \Omega$, $x \neq y$, $\lambda \in [0,1]$, se tiene $f(\lambda x + (1-\lambda)y) < \max\{f(x), f(y)\}$, entonces f se dice estrictamente casiconvexa.

Ejemplo 2.2. *Sea $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) := x^2$. f es una función casiconvexa. De hecho, como $1 - \lambda \geq 0$ para $\lambda \in [0,1]$, entonces para todo $x, y \in \mathbb{R}$ y suponiendo que $\max\{x^2, y^2\} = x^2$, tenemos*

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= (\lambda x + (1-\lambda)y)^2 \\ &= \lambda^2 x^2 + 2\lambda(1-\lambda)xy + (1-\lambda)^2 y^2 \\ &\leq \lambda^2 x^2 + 2\lambda(1-\lambda)x^2 + (1-\lambda)^2 x^2 \\ &= (\lambda^2 + 2\lambda(1-\lambda) + (1-\lambda)^2)x^2 \\ &= (\lambda^2 + 2\lambda - 2\lambda^2 + 1 - 2\lambda + \lambda^2)x^2 \\ &= x^2 \end{aligned}$$

Análogamente, si $\max\{x^2, y^2\} = y^2$.

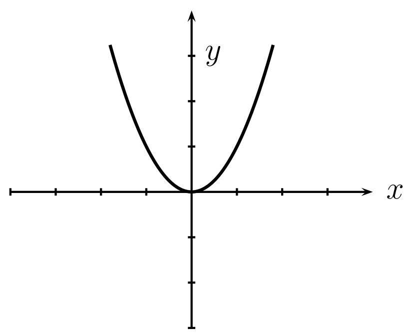


Figura 2.1: $f(x) = x^2$

Definición 2.3. *(Función casicóncava).*
Sea Ω un conjunto convexo no vacío de $\mathbb{R}^n$. Una función $f : \Omega \to \mathbb{R}$ se dice casicóncava si para todo $x, y \in \Omega$, $\lambda \in [0,1]$ se tiene que:

$$f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}.$$

Si para todo $x, y \in \Omega$, $x \neq y$, $\lambda \in [0,1]$, se tiene que

$$f(\lambda x + (1-\lambda)y) > \min\{f(x), f(y)\}$$

entonces se dice que f es estrictamente casicóncava.

Proposición 2.4. *Toda función convexa (estrictamente convexa) es casiconvexa (estrictamente casiconvexa). Además, toda función cóncava (estrictamente cóncava) es casicóncava (estrictamente casicóncava).*

Demostración.
Si f es convexa, entonces se tiene que

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Se sabe que $f(x) \leq \max\{f(x), f(y)\}$ y que $f(y) \leq \max\{f(x), f(y)\}$ luego

$$\begin{array}{rcl} f(\lambda x + (1-\lambda)y) & \leq & \lambda f(x) + (1-\lambda)f(y) \\ & \leq & \lambda \max\{f(x), f(y)\} + (1-\lambda)\max\{f(x), f(y)\} \\ & = & \max\{f(x), f(y)\}, \end{array}$$

por lo tanto

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

lo que significa que f es casiconvexa. La demostración en el caso donde f es cóncava, estrictamente cóncava y estrictamente convexa es análoga. □

Observación 2.5. *Notemos que la definición de casiconvexidad (casiconcavidad) es una definición más débil que la definición de convexidad (concavidad). Este debilitamiento se ve en el hecho de que no toda función casiconvexa (casicóncava) es convexa (cóncava). Veamos la siguiente figura:*

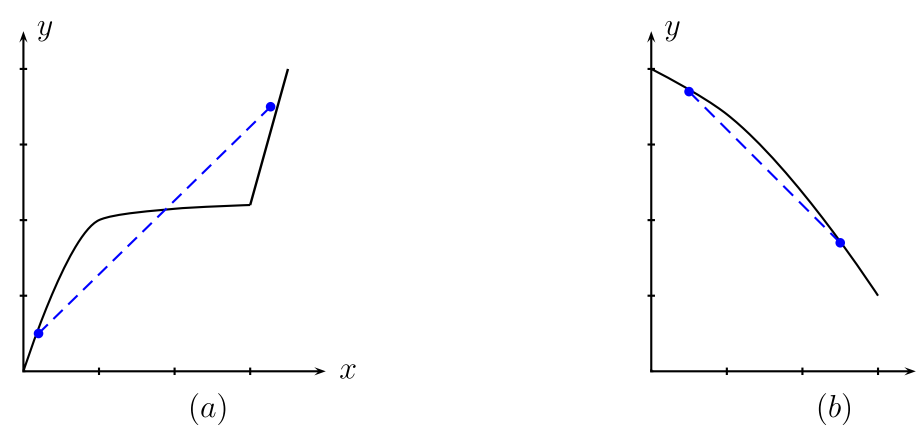


Figura 2.2: (a) Función casicóncava no cóncava y (b) Función casiconvexa no convexa

A diferencia de las funciones convexas, no toda función casiconvexa es continua. Por ejemplo, la función $f : (0, \infty) \to \mathbb{R}$ definida por

$$f(x) := \begin{cases} 1, & \text{si } x \in (0, 1) \\ 2, & \text{si } x \in [1, \infty) \end{cases}$$

es casiconvexa y no es continua en su dominio de definición. Sin embargo existe una relación entre monotonicidad y casiconvexidad para funciones de variable real.

Proposición 2.6. *Sea Ω un conjunto convexo no vacío de $\mathbb{R}$. Si $f : \Omega \to \mathbb{R}$ es monótona (creciente o decreciente) entonces es casiconvexa. Sin embargo no toda función casiconvexa es monótona.*

Demostración.
Como f es monótona entonces f es creciente o f es decreciente. Sin pérdida de generalidad, supongamos que f es creciente. Entonces $x < y$, implica $f(x) \leq f(y)$, es decir,

$$\text{máx}\{f(x), f(y)\} = f(y).$$

Además, como $\lambda \in [0, 1]$ tenemos $\lambda(x - y) < 0$, por lo tanto $\lambda x + (1 - \lambda)y < y$. Como f es creciente,

$$f(\lambda x + (1 - \lambda)y) \leq f(y).$$

Así concluimos que f es casiconvexa. $\square$

El siguiente ejemplo muestra que la recíproca de la Proposición 2.6 no siempre es verdadera.

Ejemplo 2.7. *La función $f(x) = x^2$ es una funciòn casiconvexa pero no monótona.*

El siguiente teorema nos da una caracterización de las funciones casiconvexas en términos de los conjuntos de nivel.

Teorema 2.8. *Sea $f : \Omega \to \mathbb{R}$ con $\Omega \subseteq \mathbb{R}^n$ convexo no vacío. $f(x)$ es casiconvexa si, y sólo si, el conjunto de nivel $I_\alpha = \{x \in \Omega : f(x) \leq \alpha\}$ es convexo para todo $\alpha \in \mathbb{R}$.*

Demostración.
$\Rightarrow$) Sean $x, y \in I_\alpha$. Entonces $f(x) \leq \alpha$, $f(y) \leq \alpha$. Como f es casiconvexa, entonces para $\lambda \in [0, 1]$,

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

luego, $\lambda x + (1 - \lambda)y \in I_\alpha$, lo que prueba que I_α es un conjunto convexo.

$\Leftarrow$) Supongamos ahora que el conjunto I_α es convexo. Sean x, y fijos y $\alpha = \text{máx}\{f(x), f(y)\}$. Como I_α es convexo entonces $\lambda x + (1 - \lambda)y \in I_\alpha$, luego $f(\lambda x + (1 - \lambda)y) \leq \alpha$, por tanto

$$f(\lambda x + (1 - \lambda)y) \leq \text{máx}\{f(x), f(y)\},$$

es decir, f es casiconvexa. $\square$

Proposición 2.9. *(Álgebra de funciones casiconvexas).*
Sean f, g funciones casiconvexas. Entonces se cumple que:

i). Si $a \in \mathbb{R}$, la función $h(x) = f(x) + a$ es casiconvexa.

ii). Si $a \geq 0$, la función $h(x) = af(x)$ es casiconvexa.

iii). Las funciones $h_1(x) = f(x) + g(x)$, $h_2(x) = f(x)g(x)$ y $h_3(x) = \dfrac{f(x)}{g(x)}$ no siempre son casiconvexas.

Demostración.

i). Como f es casiconvexa entonces $f(\lambda x + (1-\lambda)y) \leq \text{máx}\{f(x), f(y)\}$, luego, si sumamos a en ambos lados de la desigualdad,

$$f(\lambda x + (1-\lambda)y) + a \leq \text{máx}\{f(x), f(y)\} + a.$$

Así $f(\lambda x + (1-\lambda)y) + a \leq \text{máx}\{f(x) + a, f(y) + a\}$, es decir,

$$h(\lambda x + (1-\lambda)y) \leq \text{máx}\{h(x), h(y)\},$$

luego $h(x) = f(x) + a$ es casiconvexa.

ii). Si f es casiconvexa entonces $f(\lambda x + (1-\lambda)y) \leq \text{máx}\{f(x), f(y)\}$. Como $a \geq 0$, se tiene que

$$af(\lambda x + (1-\lambda)y) \leq a\,\text{máx}\{f(x), f(y)\},$$

por tanto

$$af(\lambda x + (1-\lambda)y) \leq \text{máx}\{af(x), af(y)\},$$

es decir,

$$h(\lambda x + (1-\lambda)y) \leq \text{máx}\{h(x), h(y)\},$$

por tanto $h(x) = af(x)$ es casiconvexa.

iii). Las funciones $f = x$, $g = -x$, y $h = x^3$ son funciones casiconvexas por ser monótonas, pero $f \cdot g = -x^2 = \dfrac{h}{g}$ no es casiconvexa. De hecho, si $\lambda = \dfrac{1}{2}$, $x = 1$, $y = -1$, entonces

$$\begin{aligned} -(\lambda x + (1-\lambda)y)^2 &= -\left(\frac{x}{2} + \frac{y}{2}\right)^2 \\ &= 0 \end{aligned}$$

Por otro lado, $\text{máx}\{-x^2, -y^2\} = -1 < 0$.

$\square$

El siguiente teorema nos da una caracterización diferencial de las funciones casiconvexas.

Teorema 2.10. *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en el interior de Ω. $f(x)$ es casiconvexa si, y sólo si, $f(x) \leq f(y)$ implica $\nabla f(y)(x-y) \leq 0$ para todo $x, y \in \Omega$.*

Demostración.
Presentamos la demostración del caso $n = 1$. La demostración del caso general es análoga. Supongamos que f es casiconvexa y que $f(x) \leq f(y)$. Entonces para todo $x, y \in \Omega$, $x \neq y$, con $\lambda \in [0, 1)$, se tiene que

$$f(\lambda x + (1-\lambda)y) \leq f(y),$$

luego

$$f(\lambda x + (1-\lambda)y) - f(y) \leq 0,$$

y así

$$\frac{f(\lambda x + (1-\lambda)y) - f(y)}{(1-\lambda)(x-y)}(x-y) \leq 0.$$

Por tanto, como $f(x)$ es diferenciable, tenemos que

$$\lim_{\lambda \to 1} \frac{f(\lambda x + (1-\lambda)y) - f(y)}{(1-\lambda)(x-y)}(x-y) = f'(y)(x-y) \leq 0.$$

Ahora supongamos que $f'(y)(x-y) \leq 0$, $f(x) \leq f(y)$ y que f no es casiconvexa. Entonces existe $\lambda \in [0, 1]$ tal que

$$f(\lambda x + (1-\lambda)y) > f(y),$$

luego $f(\lambda x + (1-\lambda)y) - f(y) > 0$, y así

$$\frac{f(\lambda x + (1-\lambda)y) - f(y)}{(1-\lambda)(x-y)}(x-y) > 0, x \neq y, \lambda \neq 1.$$

Como f es diferenciable tenemos que

$$\lim_{\lambda \to 1} \frac{f(\lambda x + (1-\lambda)y) - f(y)}{(1-\lambda)(x-y)}(x-y) = f'(y)(x-y) > 0,$$

lo que contradice la hipótesis de que $f'(y)(x-y) \leq 0$. $\square$

Ejemplo 2.11. *La función x^3 es casiconvexa en $\mathbb{R}$. De hecho, sean $x, y \in \mathbb{R}$ entonces tenemos que si $x^3 \leq y^3$, $3y^2(x-y) \leq 0$. Concluimos por el Teorema 2.10 que x^3 es casiconvexa.*

2.2. Funciones Seudoconvexas y Seudocóncavas

Definición 2.12. *(Función seudoconvexa).*
Sea $\Omega \subseteq \mathbb{R}^n$ un conjunto abierto convexo no vacío, $f : \Omega \to \mathbb{R}$ diferenciable en Ω. f se dice seudoconvexa en $y \in \Omega$ si para toda $x \in \Omega$, $\nabla f(y)(x-y) \geq 0$ implica $f(x) \geq f(y)$. Si f es seudoconvexa en cada $y \in \Omega$, entonces f es seudoconvexa en Ω.

Ejemplo 2.13. *La función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x + x^3$ es seudoconvexa en $\mathbb{R}$. De hecho, si $\nabla f(y)(x-y) \geq 0$ entonces*

$$(1 + 3y^2)(x - y) \geq 0,$$

y como $1 + 3y^2 \geq 0$ para todo $y \in \mathbb{R}$, entonces $x - y \geq 0$, luego $x \geq y$. Por lo tanto $x^3 \geq y^3$. De aquí concluimos que $x + x^3 \geq y + y^3$, es decir, $f(x) \geq f(y)$.

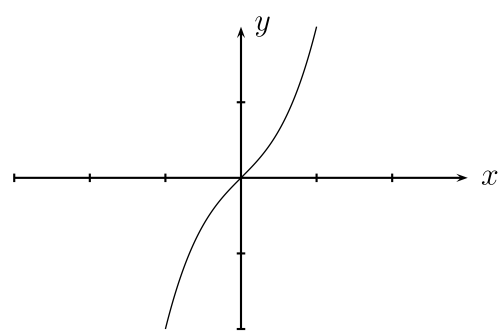


Figura 2.3: $f(x) = x^3 + x$

Proposición 2.14. *Si f es una función convexa diferenciable, entonces f es seudoconvexa.*

Demostración.
Sea f una función convexa definida en un conjunto $\Omega \subseteq \mathbb{R}^n$, entonces para todo $x, y \in \Omega$,

$$f(x) - f(y) \geq \nabla f(y)(x - y);$$

si tenemos que $\nabla f(y)(x-y) \geq 0$, entonces $f(x) - f(y) \geq 0$ luego $f(x) \geq f(y)$ y así f es seudoconvexa.

$\square$

La reciproca no siempre se cumple. Veamos el siguiente ejemplo.

Ejemplo 2.15. *La función $f(x) = x + x^3$ del ejemplo anterior es seudoconvexa pero no es convexa.*

2.3. Invexidad

En 1981, Hanson [7] introdujo una nueva clase de funciones mayor que la clase de funciones convexas diferenciables las cuales tienen la propiedad de que todo punto estacionario es un punto de mínimo global.
Esta nueva clase de funciones es llamada funciones invexas las cuales son presentadas en esta sección.

Definición 2.16. *(Función invexa).*
Sea Ω un abierto no vacío de $\mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ una función diferenciable. Se dice que f es invexa en $y \in \Omega$ si existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que

$$f(x) - f(y) \geq \nabla f(y)\eta(x, y)$$

para todo $x \in \Omega$. Si f es invexa en todo punto de Ω entonces f es invexa en Ω.

Observación 2.17. *Es claro que las funciones convexas diferenciables satisfacen la desigualdad $f(x) - f(y) \geq \nabla f(y)\eta(x, y)$ con $\eta(x, y) = x - y$, por lo tanto toda función convexa diferenciable en un conjunto Ω, es una función invexa en Ω.*

Ejemplo 2.18. *La función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = 1 - e^{-x^2}$ es invexa. De hecho, si definimos:*

$$\eta(x, y) := \begin{cases} \dfrac{-e^{-x^2} + e^{-y^2}}{2ye^{-y^2}}, & si\ y \neq 0, \\ 0, & si\ y = 0, \end{cases}$$

entonces, si $y \neq 0$, tenemos que

$$\begin{aligned} f(x) - f(y) &= (1 - e^{-x^2}) - (1 - e^{-y^2}) \\ &= 1 - e^{-x^2} - 1 + e^{-y^2} \\ &= e^{-y^2} - e^{-x^2} \\ &= \frac{(-e^{-x^2} + e^{-y^2})(2ye^{-y^2})}{2ye^{-y^2}} \\ &= \eta(x, y)(2ye^{-y^2}) \\ &= \eta(x, y)\nabla f(y). \end{aligned}$$

Ahora, si $y = 0$ entonces $\eta(x, y) = 0$ y

$$\begin{aligned} f(x) - f(y) &= e^{-y^2} - e^{-x^2} \\ &= \frac{1}{e^{y^2}} - \frac{1}{e^{x^2}} \end{aligned}$$

$$
\begin{aligned}
&= 1 - \frac{1}{e^{x^2}} \\
&\geq \nabla f(y)\eta(x,y).
\end{aligned}
$$

Así, concluimos que f es invexa.

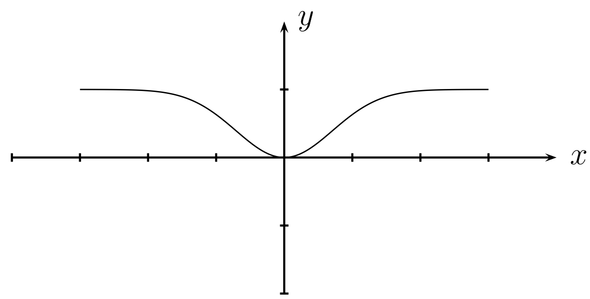


Figura 2.4: $f(x) = 1 - e^{-x^2}$

Observación 2.19. *El Ejemplo 2.18 nos muestra que la clase de funciones convexas diferenciables está contenida estrictamente en el conjunto de las funciones invexas.*

Teorema 2.20. *Sea Ω un abierto no vacío de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ diferenciable. Entonces f es invexa en Ω si, y sólo si, todo punto estacionario de f es un punto de mínimo global en Ω.*

Demostración.
$\Rightarrow$) Sea y^* un punto estacionario de f. Si f es invexa en Ω entonces existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que:

$$f(x) - f(y) \geq \nabla f(y)\eta(x,y), \text{ para todo } x, y \in \Omega.$$

Ahora, como y^* es un punto estacionario, entonces $\nabla f(y^*) = 0$. Así, $f(x) - f(y) \geq 0$, es decir, $f(x) \geq f(y)$ para todo $x \in \Omega$; luego y^* es un punto de mínimo global.

$\Leftarrow$) Supongamos ahora que todo punto estacionario de f es un punto de mínimo global en Ω. Si definimos:

$$\eta(x,y) := \begin{cases} \dfrac{f(x) - f(y)}{\nabla f(y)}, & \text{si } \nabla f(y) \neq 0, \\ 0, & \text{si } \nabla f(y) = 0, \end{cases}$$

tenemos que:

i). Si $\nabla f(y) \neq 0$, entonces

$$\begin{aligned}\nabla f(y)\eta(x,y) &= \nabla f(y)\frac{f(x)-f(y)}{\nabla f(y)} \\ &= f(x)-f(y),\end{aligned}$$

luego, $\nabla f(y)\eta(x,y) = f(x) - f(y)$.

ii). Si $\nabla f(y) = 0$, entonces $\nabla f(y)\eta(x,y) = 0$. Como y es un punto estacionario y mínimo global, $f(x) \geq f(y)$ para todo $x \in \Omega$, luego:

$$f(x) - f(y) \geq 0,$$

por lo tanto, $\nabla f(y)\eta(x,y) \leq f(x) - f(y)$.

De *i*). y *ii*). concluimos que f es invexa en Ω. □

Corolario 2.21. *Sea Ω un abierto no vacío de $\mathbb{R}^n$ y sea $f : \Omega \to \mathbb{R}$ diferenciable. Si f no tiene puntos estacionarios, entonces f es invexa.*

Demostración.
Si f no tiene puntos estacionarios entonces $\nabla f(y) \neq 0$, para todo $y \in \Omega$. Tomando

$$\eta(x,y) = \frac{f(x)-f(y)}{\nabla f(y)}$$

tenemos que $\nabla f(y)\eta(x,y) = f(x) - f(y)$, y así concluimos que

$$\nabla f(y)\eta(x,y) \leq f(x) - f(y),$$

es decir, f es invexa. □

Ejemplo 2.22. *Del Ejemplo 2.18 se puede ver, sin necesidad de definir la función $\eta(x,y)$, que la función es invexa, ya que la derivada de f es $\nabla f(x) = 2xe^{-x^2}$ y el único punto estacionario $x = 0$ es un punto de mínimo global de f sobre $\mathbb{R}$; así por el Teorema 2.20, f es invexa.*

Ejemplo 2.23. *Sea $f : (0,\infty) \to \mathbb{R}$ definida por $f(x) = \sqrt{x}$. Como $f'(x) = \dfrac{1}{2\sqrt{x}}$ para todo $x \in (0,\infty)$, entonces no tiene puntos estacionarios pues no existe x tal que $f'(x) = 0$; por tanto f es invexa en $(0,\infty)$.*

Teorema 2.24. *Sean f_i, $i = 1,\ldots,p$ funciones invexas para una misma función $\eta(x,y)$, y λ_i, $i = 1,\ldots,p$ constantes no negativas. Entonces la función*

$$\sum_{i=1}^{p} \lambda_i f_i$$

es invexa.

Demostración.
Como cada f_i es invexa entonces para $i = 1, 2, \ldots, p$, tenemos

$$f_i(x) - f_i(y) \geq \eta(x, y)\nabla f_i(y).$$

Como $\lambda_i \geq 0$ para todo $i = 1, 2, \ldots, p$, tenemos que

$$\lambda_i(f_i(x) - f_i(y)) \geq \lambda_i \eta(x, y)\nabla f_i(y),$$

luego

$$\sum_{i=1}^{p} \lambda_i f(x) - \lambda_i f_i(y) \geq \eta(x, y)\Big[\sum_{i=1}^{p} \lambda_i \nabla f_i(y)\Big]$$

y así

$$\sum_{i=1}^{p} \lambda_i f(x) - \lambda_i f_i(y) \geq \nabla\Big[\sum_{i=1}^{p} \lambda_i f_i(y)\Big]\eta(x, y);$$

luego la función

$$\sum_{i=1}^{p} \lambda_i f_i$$

es invexa. $\square$

Ahora presentamos algunos conceptos y resultados de invexidad generalizada para funciones diferenciables.

2.3.1. Casinvexidad

Definición 2.25. *(Función Casinvexa).*
Sea $\Omega \subseteq \mathbb{R}^n$ un conjunto abierto no vacio y sea $f : \Omega \to \mathbb{R}$. f se dice casinvexa en $y \in \Omega$ si existe una función $\eta(x, y) : \Omega \times \Omega \to \mathbb{R}^n$ tal que si $f(x) \leq f(y)$ entonces

$$\nabla f(y)\eta(x, y) \leq 0.$$

La función f se dice casinvexa en Ω si lo es en cada punto de Ω.

Ejemplo 2.26. *Sea $f : \mathbb{R} - \{1\} \to \mathbb{R}$, definida por $f(x) = \dfrac{2}{x-1}$. f es casinvexa. De hecho, si $f(x) \leq f(y)$ entonces tenemos que $x \geq y$; además vemos que*

$$\nabla f(y) = \frac{-2}{(x-1)^2} < 0.$$

Si tomamos $\eta(x, y) =\mid x - y \mid\geq 0$, entonces

$$\nabla f(y)\eta(x, y) \leq 0.$$

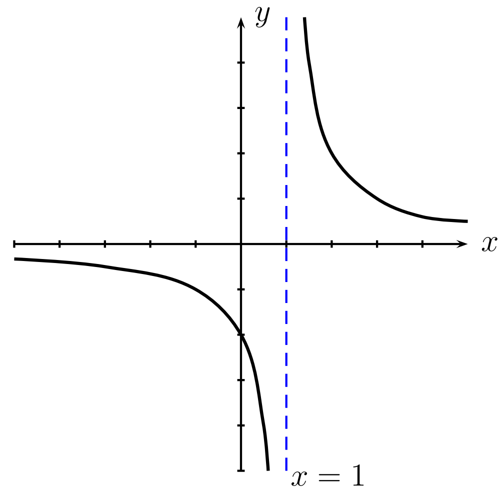


Figura 2.5: $f(x) = \dfrac{2}{x-1}$

Proposición 2.27. *Si f es invexa, entonces f es casinvexa.*

Demostración.
Si f es invexa en $y \in \Omega$, entonces existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que para todo $x \in \Omega$,

$$f(x) - f(y) \geq \nabla f(y)\eta(x, y).$$

Supongamos que$f(x) - f(y) \leq 0$, por lo tanto, para todo $x \in \Omega$

$$\nabla f(y)\eta(x, y) \leq f(x) - f(y) \leq 0,$$

es decir, $f(x) - f(y) \leq 0$ implica $\nabla f(y)\eta(x, y) \leq 0$ para todo $x \in \Omega$, luego f es casinvexa. $\square$

La recíproca no siempre es verdadera, veamos:

Ejemplo 2.28. *Tomemos la función $f(x) = x^3$. Esta función es casinvexa con $\eta(x, y) = x - y$, pero no es invexa pues en $x = 0$ tiene un punto estacionario que no es mínimo global.*

2.3.2. Seudoinvexidad

Definición 2.29. *(Función seudoinvexa).*
Sea $f : \Omega \to \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ abierto no vacío. f se dice seudoinvexa en $y \in \Omega$ si existe $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que $\nabla f(y)\eta(x, y) \geq 0$ implica $f(x) \geq f(y)$. Si f es seudoinvexa para todo $y \in \Omega$ entonces f es seudoinvexa en Ω.

Ejemplo 2.30. *Sea $f : (1, \infty) \to \mathbb{R}$ definida por $f(x) = \dfrac{x}{x^3+2}$, y sea $\eta(x, y) = (x - y)$. f es seudoinvexa. De hecho, si $\nabla f(y)\eta(x, y) \geq 0$ entonces*

$$\begin{array}{ccc} \nabla f(y) \geq 0 & y & (x-y) \geq 0, \\ & o & \\ \nabla f(y) \leq 0 & y & (x-y) \leq 0. \end{array}$$

Como la función es decreciente en el intervalo definido, $f'(y) \leq 0$ para todo $y \in (1, \infty)$; entonces $(x - y) \leq 0$, por tanto $x \leq y$, es decir, $f(x) \geq f(y)$ puesto que f es decreciente.

Teorema 2.31. *La clase de las funciones seudoinvexas coincide con la clase de las funciones invexas.*

Demostración.
$\Rightarrow$) Sea $f : \Omega \to \mathbb{R}$ una función diferenciable seudoinvexa en el abierto Ω. Entonces existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$, $\eta(x, y)$ tal que $\nabla f(y)\eta(x, y) \geq 0$ implica $f(x) \geq f(y)$. Para que f sea invexa veamos si sus puntos estacionarios son mínimos globales. Sea $y^* \in \Omega$ un punto estacionario cualquiera, entonces $\nabla f(y^*) = 0$, luego

$$\nabla f(y^*)\eta(x, y) = 0,$$

por lo tanto $f(x) \geq f(y^*)$, es decir, y^* es un punto de mínimo global y como y^* es un punto estacionario arbitrario, entonces f es invexa.

$\Leftarrow$) Ahora, si f es invexa, existe una función $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que para todo $x, y \in \Omega$

$$f(x) - f(y) \geq \nabla f(y)\eta(x, y).$$

Si $\nabla f(y)\eta(x, y) \geq 0$, entonces $f(x) - f(y) \geq 0$, luego $f(x) \geq f(y)$, por lo tanto f es seudoinvexa. □

2.4. Invexidad Vs Convexidad

A continuación mostraremos un paralelo entre los conceptos y resultados ya presentados de invexidad, invexidad generalizada, convexidad y convexidad generalizada.

El siguiente teorema establece una relación entre la noción de seudoconvexidad y casiconvexidad.

Teorema 2.32. *Sea $\Omega \subseteq \mathbb{R}^n$ abierto no vacío, y $f : \Omega \to \mathbb{R}$ diferenciable. Si f es seudoconvexa entonces f es casiconvexa. La recíproca no siempre es verdadera.*

Demostración.
Si f es seudoconvexa se tiene que, si $\nabla f(x)(x-y) \geq 0$ entonces $f(x) \geq f(y)$, es decir, si $f(x) < f(y)$ entonces
$$\nabla f(x)(x-y) < 0.$$
Por otro lado, si $f(x) = f(y)$ para todo $x, y \in \Omega$, entonces f es constante, luego
$$\nabla f(y) = 0,$$
es decir,
$$\nabla f(y)(x-y) = 0.$$
De lo anterior podemos concluir que f es casiconvexa, es decir, $f(x) \leq f(y)$ implica $\nabla f(x)(x-y) \leq 0$. $\square$

Ejemplo 2.33. *Ya sabemos por el Ejemplo 2.13 que la función $f(x) = x + x^3$ es seudoconvexa. Así por el Teorema 2.32, $f(x) = x + x^3$ es casiconvexa.*

Teorema 2.34. *Si f es seudoconvexa entonces f es invexa.*

Demostración.
Si f es seudoconvexa, entonces $\nabla f(y)\eta(x,y) \geq 0$ implica $f(x) - f(y) \geq 0$. Sea $y^* \in \Omega$ un punto estacionario cualquiera, es decir, $\nabla f(y^*) = 0$, luego
$$\nabla f(y^*)\eta(x,y) = 0$$
lo que implica $f(x) - f(y^*) \geq 0$, por lo tanto y^* es un punto de mínimo global y así f es invexa. $\square$

Existen funciones invexas que no son casiconvexas y funciones casiconvexas que no son invexas. Veamos los siguientes ejemplos.

Ejemplo 2.35. *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x_1, x_2) = x_1 + x_1^3 - 10x_2^3 - x_2$. Como*
$$\nabla f(x_1, x_2) = (1 + 3x_1^2, -30x_2^2 + 1),$$
no existe $(x_1, x_2) \in \mathbb{R}^2$ tal que $\nabla f(x_1, x_2) = 0$, por lo tanto f no tiene puntos estacionarios, luego, por el Corolario 2.21 la función f es invexa.
Si tomamos los puntos $y = (0,0)$ y $x = (2,1)$ entonces $f(x) = -1$ y $f(y) = 0$, es decir,
$$f(x) < f(y),$$
pero $\nabla f(y)(x-y) = (1,-1) \cdot (2,1) > 0$, lo que significa que f no es casiconvexa y por el Teorema 2.32 tampoco es seudoconvexa.

Ejemplo 2.36. *Consideremos la función $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$. f es casiconvexa pero no es seudoconvexa ni invexa, pues su único punto estacionario no es un punto de mínimo global.*

Teorema 2.37. *Si f es casiconvexa entonces f es casinvexa.*

Demostración.
Si f casiconvexa en $y \in \Omega$, entonces $f(x) - f(y) \leq 0$ implica $\nabla f(y)(x-y) \leq 0$ para todo $x \in \Omega$.
Sea $\eta(x,y) = x - y$, por lo tanto, existe $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que $f(x) - f(y) \leq 0$ implica $\nabla f(y)\eta(x,y) \leq 0$ para todo $x \in \Omega$, podemos concluir que f es casinvexa en $y \in \Omega$. $\square$

La recíproca en el Teorema 2.37 no es verdadera. Veamos el siguiente ejemplo.

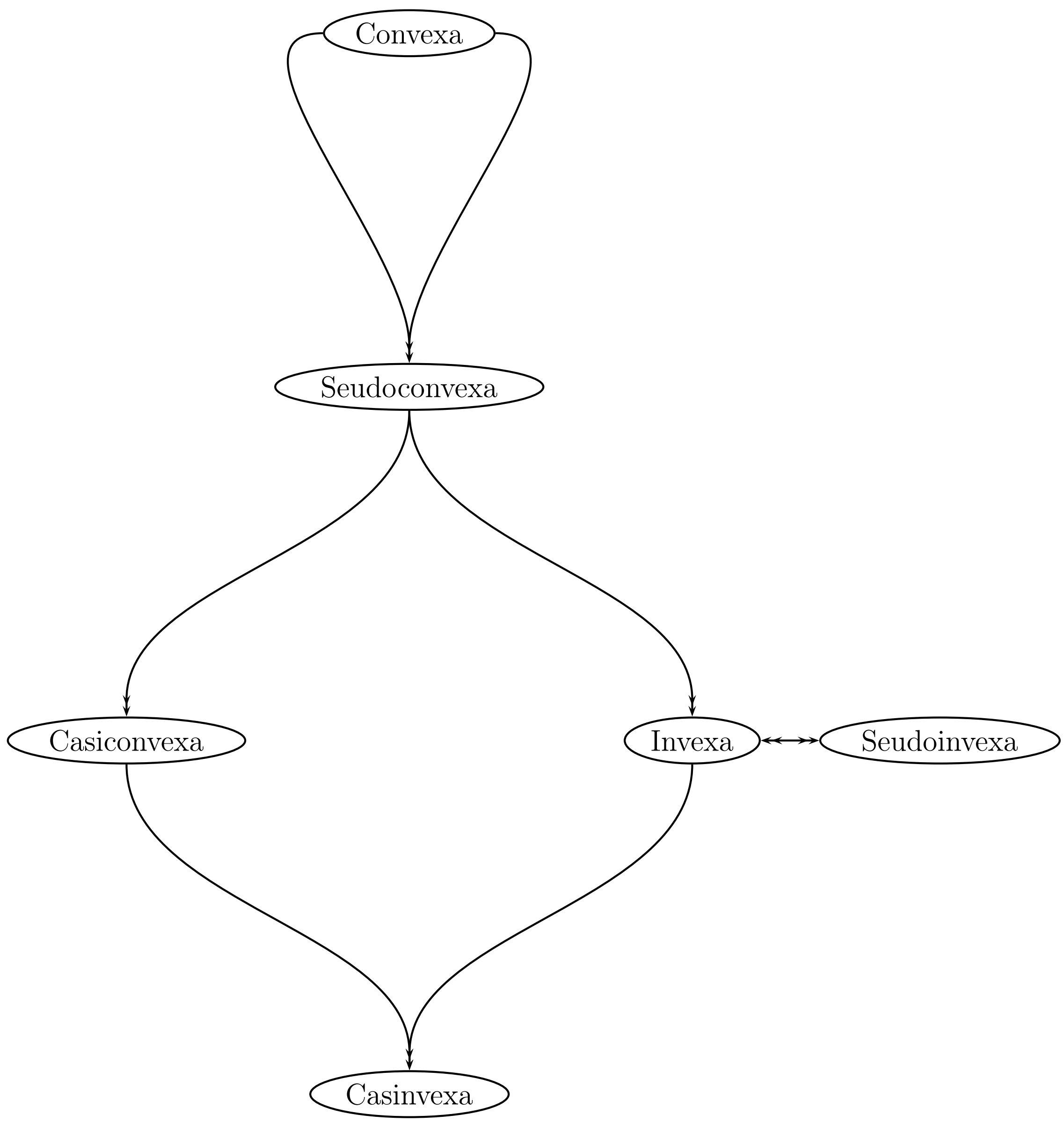
Ejemplo 2.38. *Consideremos la función $f : (0,\pi) \to \mathbb{R}$ dada por $f(x) = sen^3(x)$. f es casinvexa con $\eta(x,y) = \cos(y)(sen(x) - sen(y))$ pues, si $sen^3(x) \leq sen^3(y)$ y $\nabla f(y) = 3sen^2(y)\cos(y)$; entonces*

$$\nabla f(y)\eta(x,y) = 3sen^2(y)\cos^2(y)(sen(x) - sen(y)) \leq 0.$$

Sin embargo f no es casiconvexa, pues, para $x = \dfrac{3\pi}{4}$ y $y = \dfrac{\pi}{4}$, $f(x) = f(y)$, pero $\nabla f(y)(x-y) > 0$.

En [6], Giorgi muestra que la clase de funciones seudoconvexas es la intersección de las funciones casiconvexas y las funciones invexas. A partir de este resultado podemos concluir que la mayor clase de funciones es la clase de funciones casinvexas, dentro de las cuales están las funciones casiconvexas e invexas, como mostraremos en el siguiente diagrama.

Figura 2.6: Diagrama Comparativo.



CAPÍTULO 3

APLICACIONES DE LA INVEXIDAD EN PROGRAMACIÓN MATEMÁTICA

En este capítulo se mostrará cómo los resultados de invexidad vistos en el Capítulo 2 son aplicados en problemas de programación matemática.

3.1. Un Problema de Minimización

Sea Ω un subconjunto no vacío de $\mathbb{R}^n$ y sean $g_0, g_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, k$ funciones diferenciables. Consideremos el problema

$$\begin{cases} \text{Minimizar } g_0(x) \\ \text{sujeto a } g_i(x) \leq 0, \text{ para } i = 1, \ldots, k. \end{cases} \tag{3.1}$$

Sea $G := \{x \in \Omega : g_i(x) \leq 0,\ i = 1, \ldots, k\}$ el conjunto de los puntos que satisfacen (3.1). Suponemos que el conjunto G es diferente de vacío, y es llamado el conjunto de restricciones. Los puntos de G se denominan puntos factibles.
El objetivo aquí es encontrar candidatos óptimos a la solución del problema (3.1).

A pesar de que el problema de minimizar una función g_0 sujeta únicamente a restricciones de igualdad, como un caso especial del problema general de programación no lineal (3.1), merece un estudio particular, aquí no se enfatizará en ello, puesto que en este caso es común usar el método de multiplicadores de Lagrange bastante trabajado en los cursos regulares de cálculo de varias variables. Haremos énfasis en el caso general.

3.2. Condiciones necesarias y suficientes

Dado un problema de minimización sujeto a restricciones, una condición necesaria que debe satisfacerse en un punto de mínimo, es que el gradiente de la función objetivo sea ortogonal al plano tangente de la superficie de restricción. Entonces, si el punto es regular, el plano tangente tiene una representación que depende de los gradientes de las funciones de restricción, y la condición se puede expresar en función de los multiplicadores de Lagrange.
Si las funciones tienen segundas derivadas parciales continuas y hay multiplicadores de Lagrange, entonces, la matriz Hessiana del Lagrangiano restringido al plano tangente debe ser positivamente semidefinida en un punto mínimo, algo análogo al desempeño del Hessiano en problemas sin restricciones. De igual manera, si la matriz del Langrangiano restringido es positivamente definida en un punto que satisface las condiciones necesarias de primer orden, ese punto es de mínimo local estricto.

3.2.1. Condiciones necesarias de primer orden (Condiciones de Kuhn-Tucker)

Iniciamos esta subsección presentando la definición de punto regular de las restricciones

$$g_i(x^*) \leq 0, i = 1, \ldots, k, \qquad i \in J, \tag{3.2}$$

donde J es el conjunto de índices para los cuales $g_i(x^*) = 0$.

Definición 3.1. *(Punto regular). Sea x^* un punto que satisface las restricciones (3.2). Entonces se dice que x^* es un punto regular de las restricciones (3.2) si los vectores gradientes $\nabla g_i(x^*)$, $1 \leq i \leq k$, $i \in J$ son linealmente independientes.*

Las condiciones de primer orden están dadas en el siguiente teorema.

Teorema 3.2. *Sea x^* un punto de mínimo relativo para el problema (3.1) y supóngase que x^* es un punto regular para las restricciones. Entonces, existe un vector $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$, con $\lambda_i \geq 0$, $i = 1, \ldots, k$, tal que*

$$\begin{aligned}\nabla g_0(x^*) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x^*) &= \mathbf{0}, & & (3.3)\\ \lambda_i g_i(x^*) &= 0, \qquad i = 1, \ldots, k, & & (3.4)\\ \lambda_i &\geq 0, \qquad i = 1, \ldots, k. & & (3.5)\end{aligned}$$

Demostración.
Notemos que si $\lambda_i \geq 0$ y $g_i(x^*) \leq 0$, la última ecuación es una formulación equivalente a decir que λ_i puede ser distinta de cero si, y sólo si, la restricción correspondiente es activa (una restricción de desigualdad $g_i(x) \leq 0$ se dice que es *activa* en un punto factible x, si $g_i(x) = 0$, y es *inactiva* en x si $g_i < 0$. Por convención, cualquier restricción de igualdad $g_i(x) = 0$ se considera activa en cualquier punto factible).

Como x^* es un punto de mínimo relativo en el conjunto de restricciones, también lo es en el subconjunto definido igualando las restricciones activas a cero. Para el problema resultante con solo restricciones de igualdad, definido en un entorno de x^*, existen multiplicadores de Lagrange[1].
Por lo tanto podemos concluir que la ecuación (3.4) se cumple con $\lambda_i = 0$ si $g_i(x^*) \neq 0$, y en consecuencia se cumple también que $\lambda_i g_i(x^*) = 0$.
Para mostrar que $\lambda_i \geq 0$, para todo $i = 1, \ldots, k$, supongamos que una de sus componentes es negativa. Sea $\lambda_i < 0$ para algún $i \in J$. Sean S y M la superficie y el plano tangente, respectivamente, definidos por las restantes restricciones activas en el punto x^*. Como por hipótesis x^* es un punto regular, existe y tal que $y \in M$ y $\nabla g_i(x^*)y < 0$. Sea $x(t)$ una curva en S que pasa por x^* en $t = 0$ con $x(0) = y$. Entonces, para $t \geq 0$ pequeña, $x(t)$ es un punto factible y en consecuencia

$$\frac{dg_0}{dt}(x(t))\bigg|_{t=0} = \nabla g_0(x^*)y < 0,$$

lo que contradice el hecho de que x^* es un punto de mínimo, luego $\lambda_i \geq 0$, para todo $i = 1, \ldots, k$. □

Ejemplo 3.3. *Consideremos el siguiente problema de minimización:*

$$\begin{aligned}\textit{minimizar} \quad g_0(x) &= x_1^2 + x_2,\\ \textit{sujeta a} \quad g_1(x) &= x_1^2 + x_2^2 - 9 \leq 0,\\ g_2(x) &= x_1 + x_2 - 1 \leq 0.\end{aligned}$$

Primero que todo, veamos la gráfica que representa el problema.

[1] Ver Luenberger David [8], Parte II, pp. 304-305.

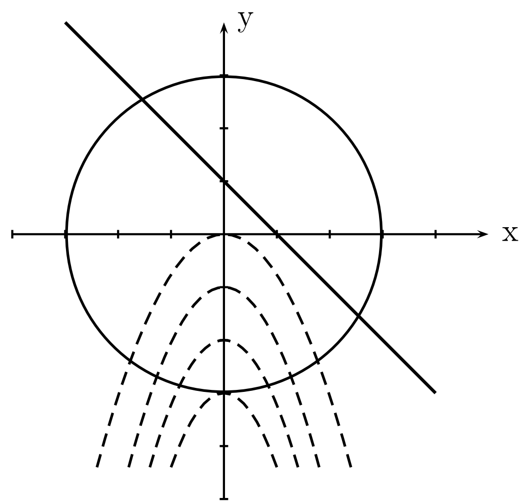


Figura 3.1: Gráfica del ejemplo 3.3

Según el Teorema 3.2, un punto $x^* = (x_1^*, x_2^*)$ *que sea de mínimo relativo satisface*

$$\begin{aligned} 2x_1^* + \lambda_1(2x_1^*) + \lambda_2 &= 0, \\ 1 + \lambda_1(2x_2^*) + \lambda_2 &= 0, \\ \lambda_1(x_1^* + x_2^* - 9) &= 0, \quad \lambda_i \geq 0, \\ \lambda_2(x_1^* + x_2^* - 1) &= 0, \quad \lambda_2 \geq 0. \end{aligned}$$

De hecho,

$$\nabla g_0(x^*) = \nabla g_0(x_1^*, x_2^*) = (2x_1^*, 1),$$

y además,

$$\begin{aligned} \sum_{i=1}^{2} \lambda_i \nabla g_i(x^*) &= \lambda_1(2x_1^*, 2x_2^*) + \lambda_2(1,1) \\ &= (2\lambda_1 x_1^*, 2\lambda_1 x_2^*) + (\lambda_1 \lambda_2) \\ &= (2\lambda_1 x_1^* + \lambda_2, 2\lambda_1 x_2^* + \lambda_2), \end{aligned}$$

por lo tanto,

$$\begin{aligned} \nabla g_0(x^*) + \sum_{i=1}^{2} \lambda_i \nabla g_i(x^*) &= (2x_1^* + 2\lambda_1 x_1^* + \lambda_2, 1 + 2\lambda_1 x_2^* + \lambda_2) \\ &= (0,0), \end{aligned}$$

es decir,

$$\begin{aligned} 2x_1^* + 2\lambda_1 x_1^* + \lambda_2 &= 0, \\ 1 + 2\lambda_1 x_2^* + \lambda_2 &= 0. \end{aligned}$$

Los escalares λ_1, λ_2 no pueden ser ambos cero, porque se produce una contradicción en la segunda ecuación (esto es, $1 = 0$). Supongamos que solo uno de estos valores es cero, por ejemplo que $\lambda_2 = 0$. Entonces, de la primera ecuación se tiene que $x_1^* = 0$ o $(1+\lambda_1) = 0$. Notemos que $(1+\lambda_1)$ no puede ser cero porque λ_1 debe ser no negativo, así que $x_1^* = 0$. Como $\lambda_1 \neq 0$, entonces de la tercera ecuación se tiene que $x_2^* = 3$ o $x_2^* = -3$; sin embargo, si $x_2^* = 3$, se viola la segunda restricción, por tanto $x^* = (0, -3)$, y en consecuencia $\lambda = (\lambda_1, \lambda_2) = \left(\frac{1}{6}, 0\right)$, que satisfacen las condiciones necesarias de primer orden, así que el punto x^* es un candidato para ser punto de mínimo relativo del problema.

3.2.2. Condiciones necesarias de segundo orden

Las condiciones necesarias de segundo orden están dadas en el siguiente teorema [8, 14].

Teorema 3.4. *Supóngase que en el Problema (3.1), $g_0, g_i(x)$ son funciones de clase C^2, esto es, funciones 2 veces derivables con segundas derivadas continuas, y que x^* es un punto regular de las restricciones*

$$g_i(x^*) \leq 0, \quad i = 1, \ldots, k. \tag{3.6}$$

Si x^* es un punto de mínimo relativo para el Problema (3.1), entonces existe $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$, $\lambda_i \geq 0$, tal que cumple las condiciones necesarias de primer orden y tal que la matriz

$$H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*)$$

es positivamente semidefinida en el subespacio tangente de las restricciones activas en x^*.
Notemos que las matrices $H_{g_0}(x^*)$, $H_{g_i}(x^*)$, $i = 1, \ldots, k$, corresponden a las Hessianas de g_0, g_i, en el punto x^*, respectivamente.

Demostración.
Si el punto x^* es el punto de mínimo relativo para las restricciones (3.6), también es un punto de mínimo relativo para el conjunto de restricciones activas tomadas como restricciones de igualdad, que son las condiciones necesarias de segundo orden para el caso en el cual se minimiza una función sujeta únicamente a restricciones de igualdad [8]; por lo tanto se tiene que el Langrangiano es positivamente semidefinido en el subespacio tangente de las restricciones activas en x^*. □

Ejemplo 3.5. *Tomemos el problema del Ejemplo 3.3 y veamos si el punto $x^* = (0, -3)$ satisface estas condiciones. Primero que todo hay que ver que el subespacio tangente de las restricciones activas está dado por los vectores $v = (v_1, v_2)$ tales que*

$$(v_1, v_2)\nabla g_1(x^*) = (v_1, v_2)(2x_1^*, 2x_2^*) = 0. \tag{3.7}$$

De la ecuación (3.7) resulta que v_1 es cualquier real y $v_2 = 0$. Ahora, si hallamos el valor de

$$H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*)$$

encontramos que

$$H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*) = \begin{bmatrix} \frac{7}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

y si aplicamos la definición de matríz positivamente semidefinida resulta

$$(v_1, 0)H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*)(v_1, 0)^T = (v_1)^2\left(\frac{7}{3}\right) \geq 0, \quad \forall\, v_1 \in \mathbb{R},$$

lo que hace que, efectivamente, el punto $x^* = (0, -3)$ cumpla las condiciones necesarias de segundo orden.

3.3. Condiciones suficientes

En esta sección presentamos un Teorema que nos da condiciones suficientes para la existencia de un mínimo relativo en un problema de minimización.

Teorema 3.6. *Sean g_0, g_i funciones de clase C^2. Las condiciones de suficiencia para que un punto x^* que satisfaga $g_i(x^*) \leq 0$ sea un punto de mínimo relativo estricto del Problema (3.1), son que exista $\lambda = (\lambda_1, \ldots, \lambda_k)$, tal que*

$$\lambda_i \geq 0,$$

$$\lambda_i g_i(x^*) = 0,$$

$$\nabla g_0(x^*) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x^*) = 0,$$

y que la matriz hessiana

$$H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*)$$

sea positivamente definida en el subespacio

$$M = \{y : \nabla g_i(x^*)y = 0, \forall i \in J\},$$

donde

$$J = \{i : g_i(x^*) = 0, \lambda_i > 0\}.$$

Demostración.
Supongamos que x^* no es un punto de mínimo relativo estricto del problema (3.1). Definamos $\{y_k\}$, una sucesión de puntos factibles que converge a x^* tal que $f(y_k) \leq f(x^*)$, y escribamos cada vector y_k en la forma $y_k = x^* + \delta_k s_k$ con $\|s_k\| = 1$, $\delta > 0$. Supongamos que $\delta_k \to 0$ y $s_k \to s^*$. Se tiene $0 \geq \nabla g_0(x^*)s^*$, y además, para cada una de las restricciones activas g_i se tiene $g_i(y_k) - g_i(x^*) \leq 0$, y por lo tanto,

$$\nabla g_i(x^*)s^* \leq 0.$$

Si $\nabla g_i(x^*)s^* = 0$ para toda $i \in J$, entonces este es el resultado para el caso donde las restricciones son de igualdad, (la demostración se puede consultar en Luenberger [8]). Si $\nabla g_i(x^*)s^* < 0$ para al menos una $i \in J$, entonces

$$0 \geq \nabla g_0(x^*)s^* = \sum_{i=1}^{k} \lambda_i \nabla g_i(x^*)s^* > 0,$$

lo cual es una contradicción.
Si retomamos los dos últimos ejemplos, nos podemos dar cuenta que

$$(v_1, 0)H_{g_0}(x^*) + \sum_{i=1}^{k} \lambda_i H_{g_i}(x^*)(v_1, 0)^T = (v_1)^2\left(\frac{7}{3}\right) > 0, \forall v_1 \neq 0$$

luego efectivamente el punto $(0, -3)$, es un punto de mínimo relativo del problema. Esto se puede observar en la gráfica que representa el problema. $\square$

3.4. Condiciones de Karush-Kuhn-Tucker (KKT) en el contexto de las funciones invexas

Es conocido que cuando las funciones del problema (3.1) son convexas, entonces las condiciones de Karush-Kuhn-Tucker también son suficientes [9].

Es posible encontrar en la literatura otras condiciones que garantizan la condición de regularidad exigida en el Teorema 3.7. (Ver por ejemplo [2] y [9]), de las cuales, las más conocidas son la condición de Slater y la condición de Mangasarian-Fromovitz.

Se han realizado estudios para ampliar la clase de funciones para las cuales las condiciones de KKT sean suficientes. Con el surgimiento de las funciones invexas fue posible concluir que esto puede ser usado cuando las funciones son invexas para una misma función η, como en [12].

El objetivo de esta sección es presentar un resultado utilizado, conocido como el Teorema de Karush-Kuhn-Tucker (KKT) en el contexto de las funciones invexas que establece condiciones necesarias (en forma de multiplicadores) para resolver el problema (3.1)

Teorema 3.7. *(Condiciones de Karush-Kuhn-Tucker).[5]*
Sea $x^* \in G$ una solución óptima (local o global) del problema (3.1), y sean g_0, g_i, $i = 1, \ldots, k$ que satisfacen alguna condición de regularidad. Entonces existen escalares λ_i, con $i = 1, \ldots, k$, tales que:

$$\begin{aligned} \nabla g_0(x^*) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x^*) &= 0; \qquad (3.8) \\ \lambda_i &\geq 0; \qquad (3.9) \\ \lambda_i g_i(x^*) &= 0; \qquad (3.10) \end{aligned}$$

Teorema 3.8. *Supóngase que x^* es un punto factible para (3.1) y que las funciones g_0, g_i sean invexas en x^* para una misma función η y que existen multiplicadores $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, k$, tales que las condiciones de KKT (3.8)-(3.10) se satisfacen en x^*, entonces x^* es un punto de mínimo global para (3.1).*

Demostración.
Por el Teorema 2.24, se sabe que la función

$$g_0 + \sum_{i=1}^{k} \lambda_i g_i \qquad (3.11)$$

es invexa en x^*. De (3.8) se ve que x^* es un punto estacionario de la función (3.11). Entonces por el Teorema 2.20, x^* es un punto de mínimo global para (3.11). De este modo,

$$g_0(x) + \sum_{i=1}^{k} \lambda_i g_i(x) \geq g_0(x^*) + \sum_{i=1}^{k} \lambda_i g_i(x^*)$$

para todo $x \in G$, siendo G el conjunto de puntos factibles. Como $\lambda_i g_i(x^*) = 0$ para todo i, entonces

$$g_0(x) + \sum_{i=1}^{k} \lambda_i g_i(x) \geq g_0(x^*).$$

Ya que $\lambda_i g_i(x) \leq 0$ para todo $x \in G$, entonces podemos concluir que $g_0(x) \geq g_0(x^*)$, es decir, x^* es solución global para el problema (3.1). □

Ejemplo 3.9. *Minimizar la función $g_0(x, y) = x - \operatorname{sen}(y)$ sujeto a las restricciones:*

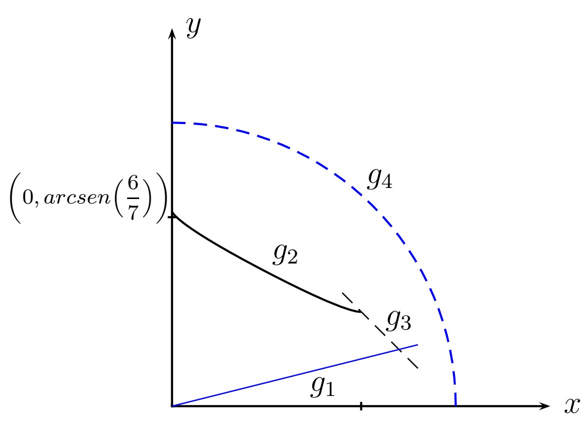
$$
\begin{aligned}
g_1 := \operatorname{sen}(x) - 4\operatorname{sen}(y) &\leq 0,\\
g_2 := 2\operatorname{sen}(x) + 7\operatorname{sen}(y) + x - 6 &\leq 0,\\
g_3 := 2x + 2y - 3 &\leq 0,\\
g_4 := 4x^2 + 4y^2 - 9 &\leq 0,\\
g_5 := -\operatorname{sen}(x) &\leq 0,\\
g_6 := -\operatorname{sen}(y) &\leq 0,\\
(x, y) &\in \mathbb{R}^2,
\end{aligned}
$$

Se puede ver que la función a minimizar y las funciones de las restricciones no son todas convexas. Sin embargo, son invexas con $\eta(z, u)$ dada por:

$$
\eta(z, u) = \Big(\frac{\operatorname{sen}(u_1) - \operatorname{sen}(x)}{\cos(x)}, \frac{\operatorname{sen}(u_2) - \operatorname{sen}(y)}{\cos(y)}\Big)
$$

donde $z = (x, y)$ y $u = (u_1, u_2)$ (Vea también el Teorema 2.20 y el Corolario 2.21). Veamos la siguiente figura:

Figura 3.2: Gráfica del ejemplo 3.9



Podemos ver, en la gráfica anterior, que las funciones no tienen puntos estacionarios, luego por el Corolario 2.21 son invexas.

La función Lagrangeana $L(z, \lambda) = g_0(z) + \sum_{i=1}^{6} \lambda_i g_i(z)$ es:

$$
\begin{aligned}
L(z,\lambda) &= x-\operatorname{sen}(y)\\
&+ \lambda_1(\operatorname{sen}(x)-4\operatorname{sen}(y))\\
&+ \lambda_2(2\operatorname{sen}(x)+7\operatorname{sen}(y)+x-6)\\
&+ \lambda_3(2x+2y-3)\\
&+ \lambda_4(4x^2+4y^2-9)\\
&+ \lambda_5(-\operatorname{sen}(x))\\
&+ \lambda_6(-\operatorname{sen}(y)).
\end{aligned}
$$

Podemos ver que las condiciones de KKT son satisfechas para $\lambda_0 = \left(0, \frac{1}{7}, 0, 0, \frac{10}{7}, 0\right)$ *y* $z_0 = \left(0, \operatorname{arc\,sen}\left(\frac{6}{7}\right)\right)$.
De hecho, la condición (3.9) es satisfecha; ahora veamos la condición (3.10):

$$
\begin{aligned}
\lambda_1(\operatorname{sen} x - 4\operatorname{sen} y) &= 0\\
\lambda_2(2\operatorname{sen} x + 7\operatorname{sen} y + x - 6) &= \frac{1}{7}\left(2\operatorname{sen}\ 0 + 7\operatorname{sen}\left(\operatorname{arc\,sen}\frac{6}{7}\right) + 0 - 6\right) = 0\\
\lambda_3(3x+2y-3) &= 0\\
\lambda_4(4x^2+4y^2-9) &= 0\\
\lambda_5(-\operatorname{sen} x) &= \frac{10}{7}(-\operatorname{sen}\ 0) = 0\\
\lambda_6(y) &= 0.
\end{aligned}
$$

Ahora, la función Langrangiana

$$
\begin{aligned}
L(z_0,\lambda_0) &= x-\operatorname{sen}(y)+\frac{2}{7}\operatorname{sen}(x)+\operatorname{sen}(y)+\frac{x}{7}-\frac{6}{7}-\frac{10}{7}\operatorname{sen}(x)\\
&= \frac{8}{7}x-\frac{8}{7}\operatorname{sen}(x)-\frac{6}{7},
\end{aligned}
$$

luego $L(z_0,\lambda_0) = -\frac{6}{7}$. *Así, aplicando el Teorema 3.8, el punto* z_0 *es un punto de mínimo global de* $g_0(x,y)$, *sujeto a las restricciones* $g_1,\ldots,g_6$.

REFERENCIAS

[1] T. Apostol. Cálculo vol 2, Ed. Reverté Colombiana, Colombia. 1989.

[2] R. Barbolla; E. Cerdá; P. Sanz. Optimización. Cuestiones, Ejercicios y Aplicaciones a la Economía. Prentice Hall, Madrid. 2000.

[3] A. Ben-Israel; B. Mond. "What's Invexity?". J. Austral. Math. Soc. Ser. B, **28**, 1986, 1-9.

[4] A.J. Brandão; M.A. Rojas-Medar; G.N. Silva. "Uma introdução às funções invexas diferenciáveis com aplicações em otimização". Boletim da Sociedade Paranaense de Matemática, Vol.19, **1-2**, 1999, 51-65.

[5] J. Cervelati; M.A. Rojas-Medar. "Funções Invexas Diferenciáveis e o Teorema de Karush-Kuhn-Tucker". Departamento de Matemática Aplicada, IMECC, UNICAMP, 13081-970 Campinas, SP, Brasil, 1-9.

[6] G. Giorgi. "A note on the relashionships between convexity and invexity". J. Austral. Math. Soc. Ser. B, **32**, 1990, 97-99.

[7] M.A. Hanson. "On sufficiency of the Kuhn-Tucker conditions". J. Math. Anal. Appl., **80**, 1981, 545-550.

[8] D. Luenberger, Programación lineal y no lineal. John Wiley & Sons, New York. 1997.

[9] O.L. Mangasarian. "Nonlinear Programming". Classics in Applied Mathematics, SIAM, **10**, 1994.

[10] D.H. Martin. "The essence of invexity". J. Math. Anal. Appl., **47**, 1985, 65-76.

[11] S. Monsalve. Matemáticas Básicas para Economistas III. Departamento de Matemáticas, Universidad Nacional de Colombia, Colombia. 2004.

[12] A.C. Moretti; M-A. Rojas-Medar. "Condiciones suficientes de optimalidad en programación no lineal". Cubo Matemática Educacional, Vol.3, **2**, 2001, 129-146.

[13] G. Strang, Álgebra lineal y sus aplicaciones. Addison Wesley, Iberoamericana, México. 1986.

[14] E.J. Villamizar-Roa. Algunos métodos para la optimización de funciones no lineales. Trabajo de grado. Escuela de Matemáticas. Universidad Industrial de Santander, Bucaramanga, Colombia. 2000.