SELECCIONES MATEMÁTICAS
Universidad Nacional de Trujillo

http://revistas.unitru.edu.pe/index.php/SSMM

Vol. 03(02): 17-35 (2016)
ISSN: 2411-1783 (versión electrónica)

CONVEXIFICACIÓN DE FUNCIONES ESTRICTAMENTE MONÓTONAS

CONVEXIFICATION OF STRICTLY MONOTONE FUNCTIONS

JENNY ROJAS JERÓNIMO* AND JHON ANGULO VITERI**

Received, Jul. 12, 2016. Accepted, Oct. 27, 2016.

Resumen

En este artículo presentamos la teoría que nos garantiza la convexificación de una función estrictamente monótona. Se demuestra un teorema y dos corolarios para la convexificación de funciones estrictamente monótonas dos veces continuamente diferenciables, luego se generaliza estos resultados para la convexificación de funciones estrictamente monótonas no diferenciables. Ambos casos son ejemplificados. Estos resultados se usan en optimización de funciones monotonas no diferenciables.

Palabras claves. Convexificación, monotonicidad. funciones no diferenciables.

Abstract

This paper presents the theory that guarantees the convexification of a strictly monotone function. We proves a theorem and two corollaries for convexification of strictly monotones functions twice continuously differentiable, then the generalization of these results is presented for convexification of nondifferentiable strictly monotones functions. Both cases are exemplars. This results are using in optimization nonsmooth.

Keywords. Convexification, monotonicity, nonsmooth functions.

1. Introducción. En lo referente a problemas de optimización global una clase importante de estos problemas es la que estudia la optimización de funciones monotonas que aparecen en varias aplicaciones como puede verse en [7] y [8]. Generalmente las funciones de los problemas de optimización monótona son no convexas por lo tanto diseñar algoritmos que los resuelvan eficientemente resulta todo un reto. Una forma de resolver estos problemas consiste en la formulación de un problema equivalente pero convexo y aprovechar asi las propiedades de optimalidad que proporciona la convexidad de las funciones. El método de convexificación, mediante ciertas transformaciones, convierte un problema de optimización monótona en un problema de minimización cóncavo [6] el cual puede resolverse eficientemente usando métodos desarrollados para minimización cóncava.
Los conceptos de funciones convexas de rango transformable y las funciones convexas de dominio transformable, fueron introducidos por [6] y [10]. Entre los primeros trabajos sobre convexificación de problemas no convexos a problemas convexos se tiene el trabajo de Sun et.all [9]. En particular la programación geométrica y la programación fraccional son dos clases de problemas de optimización no-convexa, que se pueden convexificar.
En el presente trabajo presentamos resultados que permiten convexificar funciones de optimización estrictamente monotonas dos veces continuamente diferenciables y de funciones estrictamente

*Departamento de Matemáticas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n., Ciudad Universitaria, Trujillo-Perú (jrojas@unitru.edu.pe).
**Departamento de Matemáticas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n., Ciudad Universitaria, Trujillo-Perú (jangulo@unitru.edu.pe).

monótonas no diferenciables, demostrando un teorema y dos corolarios para cada caso con sus respectivos ejemplos.

1.1. Monotonicidad y convexidad.

Definición 1 *Una función $f : \mathbb{R}^n \to \mathbb{R}$ es creciente (decreciente) en $D \subset \mathbb{R}^n$ con respecto a x_i si*

$$f(x_1, x_2, ..., x_{i-1}, x_i^1, x_{i+1}, ..., x_n) \leq (\geq) f(x_1, x_2, ..., x_{i-1}, x_i^2, x_{i+1}, ..., x_n)$$

para $x_i^1 \leq x_i^2$, donde $x_i^1, x_i^2 \in D_i = \{x_i : \quad (x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) \in D\}$.
Una función $f : \mathbb{R}^n \to \mathbb{R}$ es estrictamente creciente (estrictamente decreciente) en $D \subset \mathbb{R}^n$ con respecto a x_i si

$$f(x_1, x_2, ..., x_{i-1}, x_i^1, x_{i+1}, ..., x_n) < (>) f(x_1, x_2, ..., x_{i-1}, x_i^2, x_{i+1}, ..., x_n)$$

para $x_i^1 < x_i^2$, donde $x_i^1, x_i^2 \in D_i = \{x_i : \quad (x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n) \in D\}$

Definición 2 *Una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es creciente (decreciente) si para cualquier $x, y \in D$ con $x_i < y_i$ para $i = 1, 2, ..., n$, se cumple $f(x) \leq (\geq) f(y)$. Una función $f(x)$ es estrictamente creciente (estrictamente decreciente)si para cualquier $x, y \in D$ con $x_i < y_i$ para $i = 1, 2, ..., n$ y $x \neq y$ se cumple $f(x) < (>) f(y)$.*

Definición 3 *$f : D \subset \mathbb{R}^n \to \mathbb{R}$ es una función monótona en $D \subset \mathbb{R}^n$ si y sólo si f es creciente o decreciente.*

Definición 4 *Sea $C \subseteq \mathbb{R}^n$, es un conjunto convexo si y sólo si para cualquier $x_1, x_2 \in C$ y $\lambda \in [0, 1]$, se cumple*

$$(1 - \lambda)x_1 + \lambda x_2 \in C.$$

Es decir, el segmento de recta que une x_1 y x_2 está en C.

ejemplo 1 *Probar que $C = \{x \in R^n : Ax = b, \ A_{n\times n}, \ b_{n\times 1}\}$ es un conjunto convexo.*
En efecto: Sea $y_1, y_2 \in C$ y $\lambda \in [0, 1]$, entonces

$$\begin{aligned} Ay_1 &= b \\ Ay_2 &= b, \end{aligned}$$

y

$$(1 - \lambda)Ay_1 = (1 - \lambda)b \tag{1.1}$$

$$\lambda Ay_2 = \lambda b. \tag{1.2}$$

Sumando (1,1) *y* (1,2)

$$\begin{aligned} (1 - \lambda)Ay_1 + \lambda Ay_2 &= (1 - \lambda)b + \lambda b \\ A[(1 - \lambda)y_1 + \lambda y_2] &= b. \end{aligned}$$

Como $(1 - \lambda)y_1 + \lambda y_2 \in C$, entonces C es un conjunto convexo.

Definición 5 *Sea $f : C \to \mathbb{R}$, donde C es un conjunto convexo no vacío en $\mathbb{R}^n$. f es una función convexa en C si y sólo si*

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

La función f es llamada estrictamente convexa si

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

Teorema 1 *Sea S un conjunto convexo no vacío en $\mathbb{R}^n$, y sea $f : S \to \mathbb{R}$ una función dos veces diferenciable en S. Entonces f es convexa si y sólo si la matriz hessiana es semidefinida positiva para cada punto en S.*

1.2. Análisis no diferenciable.

Definición 6 *Sea Y un subconjunto de $\mathbb{R}^n$. Una función $f : Y \to \mathbb{R}$ satisface la condición de Lipschitz si existe una constante k no negativa tal que*

$$|f(x) - f(y)| \leq k \parallel x - y \parallel \qquad \forall x, y \in Y.$$

Definición 7 *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz continua en $x \in \mathbb{R}^n$ y d un vector en $\mathbb{R}^n$. La derivada direccional generalizada de f en x en la dirección d, se define*

$$f^\circ(x; d) = \limsup_{y \xrightarrow[t \downarrow 0]{} x} \frac{f(y + td) - f(y)}{t}$$

donde y es un vector en $\mathbb{R}^n$ y t es un escalar positivo.

A $f^\circ(x; d)$ también se conoce como derivada generalizada de Clarke.

Definición 8 *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz continua en $x \in \mathbb{R}^n$. La subdiferencial o gradiente de Clarke de f en x es el conjunto $\partial f(x)$ de vectores $\xi \in \mathbb{R}^n$ tal que*

$$\partial f(x) = \{\xi \in \mathbb{R}^n : f^\circ(x; d) \geq \xi^T d, \forall d \in \mathbb{R}^n\}.$$

Teorema 2 *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función Lipschitz continua en $x \in \mathbb{R}^n$. Entonces*

$$f^\circ(x; d) = \max\{\xi^T d : \xi \in \partial f(x)\} \qquad \forall\, d \in \mathbb{R}^n.$$

Definición 9 *La subdiferencial de una función convexa $f : \mathbb{R}^n \to \mathbb{R}$ en $x \in \mathbb{R}^n$ es el conjunto $\partial_c f(x)$ de vectores $\xi \in \mathbb{R}^n$ tal que*

$$\partial_c f(x) = \{\xi \in \mathbb{R}^n : \ f(y) \geq f(x) + \xi^T (y - x), \forall\, y \in \mathbb{R}^n\}.$$

Cada vector $\xi \in \partial_c f(x)$ es llamado subgradiente de f en x.

Teorema 3 *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Entonces la derivada direccional clásica, $f'(x, d)$, existe en toda dirección $d \in \mathbb{R}^n$ y $\forall\, x \in \mathbb{R}^n$*

$$\begin{aligned} f'(x, d) &= max\{\xi^T d : \ \xi \in \partial_c f(x)\}, \forall\, d \in \mathbb{R}^n \quad y \\ \partial_c f(x) &= \{\xi \in \mathbb{R}^n : \ f'(x, d) \geq \xi^T d, \forall\, d \in \mathbb{R}^n\}. \end{aligned}$$

Teorema 4 *La subdiferencial para funciones Lipschitz continua es una generalización de la subdiferencial para funciones convexas. Entonces, si $f : \mathbb{R}^n \to \mathbb{R}$ es una función convexa, se cumple que $f^{'}(x, d) = f^\circ(x; d)$ para todo $d \in \mathbb{R}^n$, y $\partial_c f(x) = \partial f(x)$.*

Teorema 5 *La subdiferencial para funciones Lipschitz continua es una generalización de la derivada clásica. Entonces, si $f : \mathbb{R}^n \to \mathbb{R}$ es una función Lipschitz continua y diferenciable en $x \in \mathbb{R}^n$, se cumple que $\nabla f(x) \in \partial f(x)$.*
Además, si $f : \mathbb{R}^n \to \mathbb{R}$ es continuamente diferenciable en $x \in \mathbb{R}^n$, entonce $\partial f(x) = \{\nabla f(x)\}$.

Definición 10 *Una función f es regular en x^* si la derivada direccional $f'(x^*; d)$ existe para todo $d \in \mathbb{R}^n$ y $f'(x^*; d) = f^\circ(x^*; d)$ para todo $d \in \mathbb{R}^n$, donde $f^\circ(x^*; d)$ es la derivada direccional de Clarke.*

Teorema 6 *Sea la función $g : Y \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, definida como $g = f(h)$, donde $h : Y \longrightarrow \mathbb{R}^n$ y $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Si f es una función regular en $x = h(y)$, $y \in Y$, y h una función continuamente diferenciable, entonces g es una función regular.*

Definición 11 *Suponer que $\{x^k\} \subset \mathbb{R}^n$ converge a $x^* \in \mathbb{R}^n$ con $x^k \neq x^*$ para cada k.*
La sucesión $\{x^k\}$ es convergente a x^* en la dirección u si existe $\{t_k\}$ satisfaciendo: $t_k > 0$ y $t_k \longrightarrow 0$ para $k \longrightarrow \infty$, tal que $\{\frac{x^k - x^*}{t_k}\} \longrightarrow u$.

Se observa que $\{x^k\}$ converge a x^* en la dirección u si y sólo si $\{\frac{x^k-x^*}{\|x^k-x^*\|}\} \longrightarrow \frac{u}{\|u\|}$.

Definición 12 *Sea $u \in \mathbb{R}^n$. Entonces se define $\partial_u f(x^*)$ como el conjunto de vectores v tal que, para cada uno de ellos, existen sucesiones $\{x^k\}$ y $\{v^k\}$ con $\{x^k\}$ que converge a x^* en la dirección u, $\{v^k\}$ que converge a v y v^k pertenece a $\partial f(x^k)$ para cada k.*

Luego, $\partial_u f(x^*) \subseteq \partial f(x^*)$.

Definición 13 *Sea $v^* \in \partial_u f(x^*)$. La derivada direccional inferior de segundo orden de f en x^* y v^* en la dirección u, $f''_-(x^*, v^*, u)$ se define como*

$$\liminf (v^k - v^*)^T (x^k - x^*)/t_k^2,$$

tomado sobre las sucesiones $\{x^k\}$, $\{v^k\}$, $\{t_k\}$ para los cuales
(i) $t_k > 0$ para cada k y $\{x^k\}$ converge a x^*.
(ii) $\{\frac{(x^k-x^*)}{t_k}\}$ converge a u.
(iii) $\{v^k\}$ converge a v^* con $v^k \in \partial f(x^k)$ para cada k.

Definición 14 *Una función f es semisuave en x^* si la sucesión $\{(v^k)^T u\}$ converge siempre que $\{x^k\}$ converge a x^* en la dirección u y $v^k \in \partial f(x^k)$ para cada k.*

Teorema 7 *Sea la función $g : Y \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, definida como $g = f(h)$, donde $h : Y \longrightarrow \mathbb{R}^n$ y $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Si f es una función semisuave en $x = h(y)$, $y \in Y$, y h una función continuamente diferenciable, entonces g es una función semisuave.*

Lema 1 *Suponer que f es semisuave y regular en un conjunto abierto convexo W. Suponer también que $f''_-(x^*, v^*, u) \geq 0$ para cada tripleta (x^*, v^*, u) con $x^* \in W$, $u \in \mathbb{R}^n$, y $v^* \in \partial_u f(x^*)$. Entonces, f es convexa en W.*

2. Material y Métodos.

2.1. Objeto de estudio.

a) Funciones estrictamente monótonas.
b) Convexificación de funciones estrictamente monótonas dos veces continuamente diferenciables.
c) Convexificación de funciones estrictamente monótonas no diferenciables

2.2. Instrumentos de recolección de datos.

Sea f una función definida en:

$$X = \prod_{j=1}^{n} [l_j, L_j] = [l_1, L_1] \times [l_2, L_2] \times ... \times [l_n, L_n] \tag{2.1}$$

con $0 \leq l_j \leq L_j$ para cada j. La función f es llamada función estrictamente monótona si es una función estrictamente monótona para cada x_j.
Sea $h : (y_1, y_2, ..., y_n) \longmapsto (h_1(y_1), h_2(y_2), ..., h_n(y_n))$ una función biyectiva y sea la siguiente transformación de variable de la función f,

$$g(y) = f(h(y)), \tag{2.2}$$

donde el dominio de g es

$$Y = \prod_{j=1}^{n} Y_j = \prod_{j=1}^{n} h^{-1}([l_j, L_j]). \tag{2.3}$$

2.3. Métodos y técnicas.

2.3.1. Convexificación de funciones estrictamente monótonas dos veces continuamente diferenciables. Sea f una función dos veces continuamente diferenciable y $S^n = \{d \in R^n : \|d\| = 1\}$, se define

$$\sigma = min\{d^T \nabla^2 f(x) d \ : \quad x \in X, d \in S^n\}$$

$$\eta = min\left\{\frac{\partial f}{\partial x_j} \ : \quad x \in X, j = 1, 2, ..., n\right\}.$$

Teorema 8 *Sea f una función dos veces continuamente diferenciable y estrictamente creciente en X con $\eta > 0$ y h_j, $j = 1, 2, ..., n$ funciones dos veces continuamente diferenciable y estrictamente monótona que satisfacen la siguiente condición*

$$\frac{h_j''(y_j)}{[h_j{}'(y_j)]^2} \geq -\frac{\sigma}{\eta}, \qquad \forall\, y_j \in Y_j, \quad j = 1, 2, ..., n, \tag{2.4}$$

entonces $g(y) = f(h(y))$ es una función convexa en Y.

Demostración
Debido a que f y h son funciones dos veces continuamente diferenciable, entonces g es una función dos veces continuamente diferenciable, por lo que será suficiente probar que la matriz hessiana de g es semidefinida positiva en Y.
Para cualquier $y \in Y$, sea $x = h(y)$, entonces $x \in X$, luego:

$$\frac{\partial g}{\partial y_j} = h_j'(y_j)\frac{\partial f}{\partial x_j}, \qquad j = 1, 2, ..., n.$$

$$\frac{\partial^2 g}{\partial y_i^2} = h_i''(y_i)\frac{\partial f}{\partial x_i} + (h_i'(y_i))^2\frac{\partial^2 f}{\partial x_i^2}, \qquad i = 1, 2, ..., n. \tag{2.5}$$

$$\frac{\partial^2 g}{\partial y_i \partial y_j} = h_i'(y_i)h_j'(y_j)\frac{\partial^2 f}{\partial x_i \partial x_j}, \qquad i \neq j, \quad j, i = 1, 2, ..., n. \tag{2.6}$$

Luego la matriz hessiana, $H(y)$, de g es

$$H(y) = \nabla^2 g(y) = \begin{bmatrix} \frac{\partial^2 g}{\partial y_1 \partial y_1}(y) & \frac{\partial^2 g}{\partial y_1 \partial y_2}(y) & \frac{\partial^2 g}{\partial y_1 \partial y_3}(y) & \cdots & \frac{\partial^2 g}{\partial y_1 \partial y_n}(y) \\ \frac{\partial^2 g}{\partial y_2 \partial y_1}(y) & \frac{\partial^2 g}{\partial y_2 \partial y_2}(y) & \frac{\partial^2 g}{\partial y_2 \partial y_3}(y) & \cdots & \frac{\partial^2 g}{\partial y_2 \partial y_n}(y) \\ \frac{\partial^2 g}{\partial y_3 \partial y_1}(y) & \frac{\partial^2 g}{\partial y_3 \partial y_2}(y) & \frac{\partial^2 g}{\partial y_3 \partial y_3}(y) & \cdots & \frac{\partial^2 g}{\partial y_3 \partial y_n}(y) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 g}{\partial y_n \partial y_1}(y) & \frac{\partial^2 g}{\partial y_n \partial y_2}(y) & \frac{\partial^2 g}{\partial y_n \partial y_3}(y) & \cdots & \frac{\partial^2 g}{\partial y_n \partial y_n}(y) \end{bmatrix}$$

Sea
$h_i' = h_i'(y_i)$,
$h_i'' = h_i''(y_i)$,
$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_i}(x)$,
$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 f}{\partial x_i^2}(x)$,
$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \qquad \forall i, j = 1, 2, ..., n.$

Reemplazando (2.5)y (2.6) en $H(y)$

$$\begin{bmatrix} h_1''\frac{\partial f}{\partial x_1}+(h_1')^2\frac{\partial^2 f}{\partial x_1^2} & h_1'h_2'\frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & h_1'h_n'\frac{\partial^2 f}{\partial x_1\partial x_n} \\ h_2'h_1'\frac{\partial^2 f}{\partial x_2\partial x_1} & h_2''\frac{\partial f}{\partial x_2}+(h_2')^2\frac{\partial^2 f}{\partial x_2^2} & \cdots & h_2'h_n'\frac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ h_n'h_1'\frac{\partial^2 f}{\partial x_n\partial x_1} & h_n'h_2'\frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & h_n''\frac{\partial f}{\partial x_n}+(h_n')^2\frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$$= \begin{bmatrix} h_1'\left[\frac{h_1''}{(h_1')^2}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial x_1^2}\right]h_1' & h_1'\frac{\partial^2 f}{\partial x_1\partial x_2}h_2' & \cdots & h_1'\frac{\partial^2 f}{\partial x_1\partial x_n}h_n' \\ h_2'\frac{\partial^2 f}{\partial x_2\partial x_1}h_1' & h_2'\left[\frac{h_2''}{(h_2')^2}\frac{\partial f}{\partial x_2}+\frac{\partial^2 f}{\partial x_2^2}\right]h_2' & \cdots & h_2'\frac{\partial^2 f}{\partial x_2\partial x_n}h_n' \\ \vdots & \vdots & \ddots & \vdots \\ h_n'\frac{\partial^2 f}{\partial x_n\partial x_1}h_1' & h_n'\frac{\partial^2 f}{\partial x_n\partial x_2}h_2' & \cdots & h_n'\left[\frac{h_n''}{(h_n')^2}\frac{\partial f}{\partial x_n}+\frac{\partial^2 f}{\partial x_n^2}\right]h_n' \end{bmatrix}$$

$$= \begin{bmatrix} h_1' & 0 & \cdots & 0 \\ 0 & h_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n' \end{bmatrix} \begin{bmatrix} \frac{h_1''}{(h_1')^2}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{h_2''}{(h_2')^2}\frac{\partial f}{\partial x_2}+\frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{h_n''}{(h_n')^2}\frac{\partial f}{\partial x_n}+\frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \begin{bmatrix} h_1' & 0 & \cdots & 0 \\ 0 & h_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n' \end{bmatrix}$$

Sea

$$A(y) = \begin{bmatrix} h_1' & 0 & \cdots & 0 \\ 0 & h_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n' \end{bmatrix}$$

$$\begin{aligned} C(x) &= \begin{bmatrix} \frac{h_1''}{(h_1')^2}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{h_2''}{(h_2')^2}\frac{\partial f}{\partial x_2}+\frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{h_n''}{(h_n')^2}\frac{\partial f}{\partial x_n}+\frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \\ &= \left(\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} + \begin{bmatrix} \frac{h_1''}{(h_1')^2}\frac{\partial f}{\partial x_1} & 0 & \cdots & 0 \\ 0 & \frac{h_2''}{(h_2')^2}\frac{\partial f}{\partial x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{h_n''}{(h_n')^2}\frac{\partial f}{\partial x_n} \end{bmatrix} \right) \\ &= (\nabla^2 f(x) + B(x)), \end{aligned}$$

donde

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix},$$

$$B(x) = \begin{bmatrix} \frac{h_1''}{(h_1')^2}\frac{\partial f}{\partial x_1} & 0 & \cdots & 0 \\ 0 & \frac{h_2''}{(h_2')^2}\frac{\partial f}{\partial x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{h_n''}{(h_n')^2}\frac{\partial f}{\partial x_n} \end{bmatrix}$$

En consecuencia $H(y) = A(y)(\nabla^2 f(x) + B(x))A(y)$.
Ahora para todo $d \in S^n$

$$d^T H(y) d = d^T A(y) C(x) A(x) d.$$

Es claro que $A(y)C(x)A(y)$ es una matriz semidefinida positiva si y sólo si $C(x)$ es una matriz semidefinida positiva.
Por la definición de σ y η, se tiene

$$d^T\nabla^2 f(x)d \geq \sigma \qquad \forall\, d \in S^n,$$

$$\frac{\partial f}{\partial x_j} \geq \eta > 0 \qquad \forall\, x_j \in [l_j, L_j],$$

y para cualquier $d \in S^n$,

$$\begin{aligned} d^T C(x)d &= d^T(\nabla^2 f(x) + B(x))d \\ &= d^T\nabla^2 f(x)d + d^T B(x)d \\ &\geq \sigma + \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j}\frac{h_j''(y_j)}{[h_j'(y_j)]^2} d_j^2 \\ &\geq \sigma + \left(\eta\left(-\frac{\sigma}{\eta}\right)\right)\sum_{j=1}^{n} d_j^2 \\ &= \sigma - \sigma \parallel d \parallel \\ &= \sigma - \sigma \\ &= 0, \end{aligned}$$

es decir, $d^T C(x)d \geq 0$. Luego $C(x)$ es una matriz semidefinida positiva.
Entonces $H(y) = \nabla^2 g(y)$ es una matriz semidefinida positiva.
Por lo tanto g es una función convexa. ∇
Similarmente, una función estrictamente decreciente se puede convertir a una función convexa a través de trasformaciones de variables satisfaciendo

$$\frac{h_j^{''}(y_j)}{[{h_j}'(y_j)]^2} \leq -\frac{\sigma}{\eta}, \qquad \forall\, y_j \in Y_j, \qquad j = 1, 2, ..., n.$$

Hay funciones especificas que satisfacen la condición (2.4), en particular, considerar las siguientes dos funciones

$$h_j(y_j) = \frac{1}{p}\ln\left(1 - \frac{1}{y_j}\right), \qquad p > 0, \qquad j = 1, 2, ..., n. \tag{2.7}$$

$$h_j(y_j) = y_j^{-\frac{1}{p}}, \qquad p > 0, \qquad j = 1, 2, ..., n. \tag{2.8}$$

Corolario 1 *Sea f una función que satisface las condiciones del teorema anterior, con*

$$h_j(y_j) = \frac{1}{p}\ln\left(1 - \frac{1}{y_j}\right), \qquad p > 0, \qquad j = 1, 2, ..., n,$$

entonces existe una constante $p_1 \geq 0$ tal que $g(y) = f(h(y))$ es una función convexa en $Y = \left[\frac{1}{(1-e^{pl_j})}, \frac{1}{(1-e^{pL_j})}\right]$ cuando $p \geq p_1$.

Demostración
Será suficiente probar que se cumple (2.4), notar que

$$\begin{aligned} h_j'(y_j) &= \frac{1}{p}\left(\frac{1}{1 - \frac{1}{y_j}}\right)\left(1 - \frac{1}{y_j}\right)' \\ &= \frac{1}{p}\left(\frac{1}{\frac{y_j - 1}{y_j}}\right)\left(\frac{1}{y_j^2}\right) \\ &= \frac{1}{p}\left(\frac{1}{y_j(y_j - 1)}\right) \end{aligned}$$

$$\begin{aligned}
h_j''(y_j) &= \frac{1}{p}\left(\frac{-(y_j(y_j-1))'}{y_j^2(y_j-1)^2}\right)\\
&= \frac{1}{p}\left(\frac{-((y_j-1)+y_j)}{y_j^2(y_j-1)^2}\right)\\
&= \frac{1}{p}\left(\frac{1-2y_j}{y_j^2(y_j-1)^2}\right)
\end{aligned}$$

Ya que $y_j < 0$ para $y_j \in \left[\frac{1}{(1-e^{pl_j})}, \frac{1}{(1-e^{pL_j})}\right]$, se tiene

$$\begin{aligned}
\frac{h_j''(y_j)}{[h_j'(y_j)]^2} &= \frac{\frac{1}{p}\left(\frac{1-2y_j}{y_j^2(y_j-1)^2}\right)}{\left(\frac{1}{p}\left(\frac{1}{y_j(y_j-1)}\right)\right)^2}\\
&= \frac{p^2(1-2y_j)(y_j(y_j-1))^2}{py_j^2(y_j-1)^2}\\
&= p(1-2y_j)\\
&> p.
\end{aligned}$$

Para que la condición (2.4) se cumpla, se debe satisfacer que $p \geq p_1 = max\left\{0, -\frac{\sigma}{\eta}\right\}$.
Por lo tanto g es una función convexa en Y cuando $p \geq p_1$. ∇

De forma similar se enuncia el siguiente corolario usando la función (2. 6).

Corolario 2 *Sea f una función que satisface las condiciones del teorema anterior, con*

$$h_j(y_j) = y_j^{-\frac{1}{p}}, \qquad p > 0, \qquad j = 1,2,...,n.$$

entonces existe un $p_2 \geq 0$ tal que $g(y) = f(h(y))$ es una función convexa en $Y = \prod_{j=1}^{n}[L_j^{-p}, l_j^{-p}]$ cuando $p \geq p_2$.

Demostración
Será suficiente probar que se cumple (2.4), notar que

$$\begin{aligned}
h_j'(y_j) &= -\frac{1}{p}y_j^{\left(-\frac{1}{p}-1\right)}\\
h_j''(y_j) &= \left(-\frac{1}{p}\right)\left(-\frac{1}{p}-1\right)y_j^{\left(-\frac{1}{p}-2\right)}\\
&= \left(\frac{1}{p}\right)\left(\frac{1}{p}+1\right)y_j^{\left(-\frac{1}{p}-2\right)}.
\end{aligned}$$

Ahora para todo $y_j \in \left[L_j^{-p}, l_j^{-p}\right]$, se tiene

$$\begin{aligned}
\frac{h_j''(y_j)}{(h_j'(y_j))^2} &= \frac{\left(\frac{1}{p}\right)\left(\frac{1}{p}+1\right)y_j^{\left(-\frac{1}{p}-2\right)}}{\left(-\frac{1}{p}y_j^{\left(-\frac{1}{p}-1\right)}\right)^2}\\
&= (1+p)y_j^{\frac{1}{p}}\\
&\geq (1+p)\frac{1}{L_j}
\end{aligned}$$

pues,

$$L_j^{-p} \leq y_j \leq l_j^{-p}$$

$$L_j^{-1} \leq y_j^{\frac{1}{p}} \leq l_j^{-1}$$

entonces, $y_j^{\frac{1}{p}} \geq L_j^{-1} = \frac{1}{L_j}$.
Sea $L = max_{1\leq j\leq n} L_j$,
luego

$$\begin{aligned}(1+p)\frac{1}{L_j} &\geq -\frac{\sigma}{\eta}\\ (1+p) &\geq -\frac{\sigma}{\eta}L\\ p &\geq -\frac{\sigma}{\eta}L - 1.\end{aligned}$$

Por lo tanto la condición (2.4) se cumple cuando $p \geq p_2 = max\left\{0, -\frac{\sigma}{\eta}L - 1\right\}$. ∇

Para ilustrar la convexificación de una función estrictamente monótona se considera el siguiente ejemplo.

ejemplo 2 *Sea la función*

$$f(x) = \frac{1}{3}(x-2)^3 + x, \qquad x \in X = [1, 3].$$

Usando los corolarios (1) y (2), hallar los valores de p para garantizar la convexificación de f.
En efecto
La función f es no convexa, pues

$$\begin{aligned}f'(x) &= (x-2)^2 + 1\\ f''(x) &= 2(x-2),\end{aligned}$$

al evaluar para $x = 1$, se tiene que $f''(x) = -2$. Entonces $f(x)$ es una función no convexa.
Notar que

$$\begin{aligned}f'(x) &= (x-2)^2 + 1 \geq 1, & \forall x \in [1,3]\\ f''(x) &= 2(x-2) \geq -2, & \forall x \in [1,3]\end{aligned}$$

Usando la función definida en (2.7) y el corolario (1),

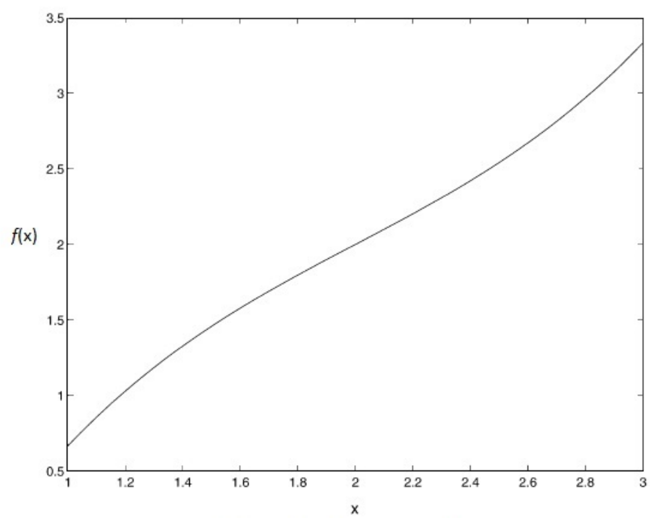

Figura 1. función no convexa $f(x)$.

hallar

$$\begin{aligned}p_1 &= max\left\{0, -\frac{\sigma}{\eta}\right\}\\ p_1 &= max\left\{0, -\frac{-2}{1}\right\} = 2,\end{aligned}$$

entonces para cualquier $p \geq p_1 = 2$ garantiza la convexidad de g.
Usando la función definida en (2.8) y el corolario (2),

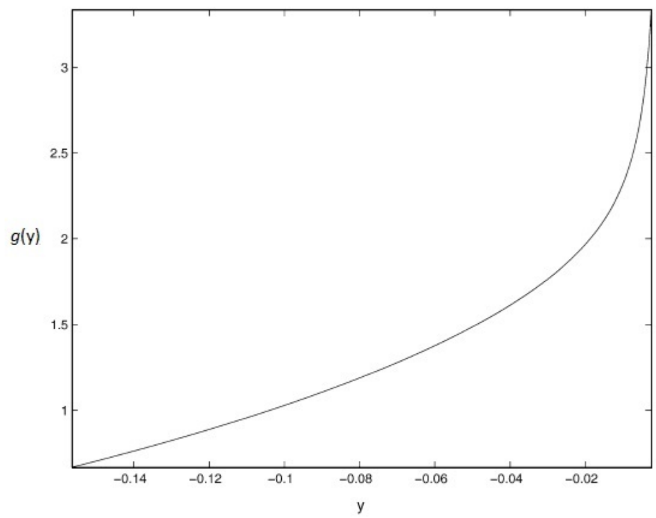

Figura 2. Función convexificada g(y) con h(y) definida en (2.7) y p=2.

hallar

$$p_2 = max\left\{0, -\frac{\sigma}{\eta}L - 1\right\}$$
$$p_2 = max\left\{0, -\frac{(-2)}{1}3 - 1\right\} = 5,$$

entonces para cuaquier $p \geq p_2 = 5$ se garantiza la convexidad de g.

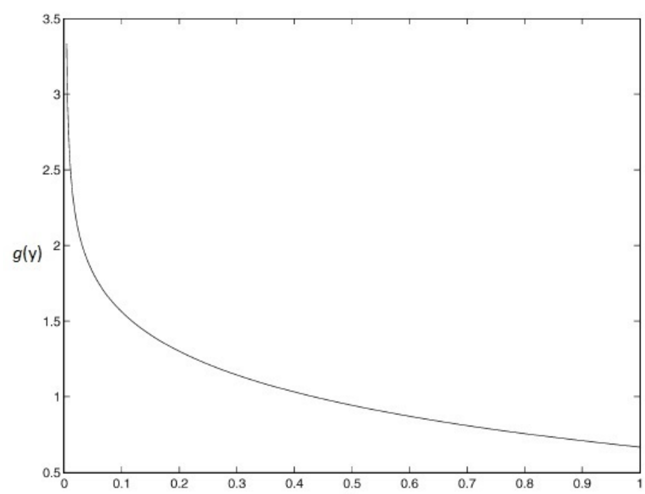

Figura 3. Función convexificada g(y) con h(y) definida en (2.8) y p=5.

Por lo tanto en esta sección se probó que se puede convexificar una función dos veces continuamente diferenciable estrictamente creciente, en la siguiente sección se ampliará los resultados de convexificación para funciones no suaves. Especificamente se probará que funciones estrictamente monótonas semisuaves y regulares pueden ser convertidas en funciones convexas.

2.3.2. Convexificación de funciones monótonas no diferenciables.. Se considera a X y Y definidos como en (2.1) y (2.3) respectivamente y g como (2.2).
Sea W un conjunto abierto convexo tal que $Y \subseteq W$.

Teorema 9 *Asumir que*
i) f *es una función semisuave y regular en un conjunto abierto convexo que contiene a* X.
ii) f *es una función estrictamente creciente en* X *y*

$$\inf \min_{i=1,...,n} = \{\xi_i : \xi = (\xi_1, \xi_2, ..., \xi_n)^T \in \partial f(x), x \in X\} \geq \epsilon > 0 \tag{2.9}$$

$$\rho = \inf\{f''_-(x, v, w) : x \in X, \| w \|= 1, v \in \partial_w f(x)\} > -\infty \tag{2.10}$$

iii) $h_i \in C^2(W), \quad i = 1, 2, ..., n$ *son funciones estrictamente monótonas convexas que satisfacen*

$$\frac{h''_i(y_i)}{[h'_i(y_i)]^2} \geq \frac{\sigma}{\epsilon}, \quad y \in W, \quad i = 1, 2, ..., n, \tag{2.11}$$

donde $\sigma = -\rho$.
Entonces g *es una función convexa en cualquier subconjunto convexo de* Y.

Demostración
•) Asumir que $\rho \geq 0$.
Tenemos por $i)$ que f es una función semisuave y regular en un conjunto abierto convexo contenido en X.
Por la condición $ii)$

$$\rho = \inf\{f''_-(x, v, w) : x \in X, \| w \|= 1, v \in \partial_w f(x)\}\} \leq f''_-(x, v, w),$$

entonces $f''_-(x, v, w) \geq \rho \geq 0$, consecuentemente $f''_-(x, v, w) \geq 0$. Ahora por lema (1) concluimos que f es una función convexa.

•) Sin pérdida de generalidad, asumir que $\sigma = -\rho > 0$, es decir, $\rho < 0$.
Para $y \in Y$, sea $x = h(y)$, entonces $x \in X$.
Como f es una función semisuave en $x = h(y)$ y h es una función continuamente diferenciable, por teorema (7) se tiene que $g(y) = f(h(y))$ es semisuave en Y.
También como f es una función regular en $x = h(y)$ y h una función continuamente diferenciable, entonces por teorema (6) se tiene que $g(y) = f(h(y))$ es regular en Y.
Luego g es una función semisuave y regular, por lo que será suficiente probar que $g''_-(y^*, v^*, d) \geq 0$ para cualquier $y^* \in Y$, $d \in \mathbb{R}^n$ y $v^* \in \partial_d g(y^*)$.
En efecto:
dado cualquier $y^* \in Y$, $d \in \mathbb{R}^n$, $d \neq 0$ y $v^* \in \partial_d g(y^*)$, suponer que $\{y^k\}$, $\{v^k\}$ y $\{t_k\}$ son sucesiones tales que

1. $t_k > 0$, $\forall k$.
2. $\{y^k\}$ converge a y^*.
3. $\{\frac{y^k - y^*}{t_k}\}$ converge a d.
4. $\{v^k\}$ converge a v^*, con $v^k \in \partial g(y^k)$, para cada k.

Sea $x^* = h(y^*) \in X$ y $x^k = h(y^k)$ para cada k, entonces $\{x^k\} \longrightarrow x^*$, cuando $k \longrightarrow \infty$, pues h es uno a uno. Si $h_i \in C^2(W)$, existe $\theta^k_1, \theta^k_2, ..., \theta^k_n \in (0, 1)$ tal que

$$h_i(y^k_i) - h_i(y^*_i) = h'_i(y^*_i + \theta^k_i(y^k_i - y^*_i))(y^k_i - y^*_i). \tag{2.12}$$

Consecuentemente

$$x^k - x^* = h(y^k) - h(y^*) = \nabla h(\eta^k)(y^k - y^*) \tag{2.13}$$

donde $\eta^k = (\eta^k_1, \eta^k_2, ..., \eta^k_n), \quad \eta^k_i = y^*_i + \theta^k_i(y^k_i - y^*_i), \quad \forall i = 1, 2, ..., n$.
También, se tiene que

$$h'_i(y_j) = 0, \quad \forall i \neq j,$$

por lo que el jacobiano de h en η^k es

$$\nabla h(\eta^k) = diag(h'_1(\eta^k_1), h'_2(\eta^k_2), ..., h'_n(\eta^k_n)).$$

Dividiendo ambos lados de (2.13) entre t_k y tomando límite cuando $k \longrightarrow \infty$, se obtiene

$$\begin{aligned}\lim_{k\longrightarrow\infty} \frac{x^k - x^*}{t_k} &= \lim_{k\longrightarrow\infty} \nabla h(\eta^k)\frac{(y^k - y^*)}{t_k} \\ &= \nabla h(y^*)d,\end{aligned} \tag{2.14}$$

sea $u = \nabla h(y^*)d$. Entonces por la definición (11), $\{x^k\}$ converge a x^* en la dirección u.
Por la regla de la cadena de la derivada generalizada del gradiente de Clarke, se cumple que

$$\partial g(y) \subseteq \{\nabla h(y)\xi : \xi \in \partial f(h(y))\}, \qquad y \in Y.$$

Ya que $v^k \in \partial g(y^k)$, entonces existe $\xi^k \in \partial f(x^k)$ tal que

$$v^k = \nabla h(y^k)\xi^k \tag{2.15}$$

Por la condición $iii)$ del teorema, h_i es una función estrictamente monótona para cada i, entonces, $h_i'(y_i) \neq 0$ para todo i.
Luego

$$\xi^k = \nabla h(y^k)^{-1}v^k,$$

y

$$\begin{aligned}\lim_{k\longrightarrow\infty} \xi^k &= \lim_{k\longrightarrow\infty} \nabla h(y^k)^{-1}v^k \\ &= \nabla h(y^*)^{-1} \lim_{k\longrightarrow\infty} v^k \\ &= \nabla h(y^*)^{-1}v^*,\end{aligned}$$

por tanto $\{\xi^k\}$ converge, sea

$$\xi^* = \lim_{k\longrightarrow\infty} \xi^k,$$

por la cerradura de $\partial f(x)$, $\xi^* \in \partial f(x^*)$, además, $\xi^* \in \partial_u f(x^*)$ y

$$v^* = \nabla h(y^*)\xi^*. \tag{2.16}$$

Usando (2.15) y (2.16), se tiene

$$\begin{aligned}&(v^k - v^*)^T(y^k - y^*) \\ &= (\nabla h(y^k)\xi^k - \nabla h(y^*)\xi^*)^T(y^k - y^*) \\ &= (\nabla h(y^k)\xi^k - \nabla h(y^k)\xi^* + \nabla h(y^k)\xi^* - \nabla h(y^*)\xi^*)^T(y^k - y^*) \\ &= [(\nabla h(y^k)\xi^k - \nabla h(y^k)\xi^*)^T + (\nabla h(y^k)\xi^* - \nabla h(y^*)\xi^*)^T](y^k - y^*) \\ &= (\nabla h(y^k)\xi^k - \nabla h(y^k)\xi^*)^T(y^k - y^*) + (\nabla h(y^k)\xi^* - \nabla h(y^*)\xi^*)^T(y^k - y^*) \\ &= (\xi^k - \xi^*)^T\nabla h(y^k)(y^k - y^*) + (\xi^*)^T(\nabla h(y^k) - \nabla h(y^*))(y^k - y^*),\end{aligned}$$

por lo tanto,

$$(v^k - v^*)^T(y^k - y^*) = (\xi^k - \xi^*)^T\nabla h(y^k)(y^k - y^*) + (\xi^*)^T(\nabla h(y^k) - \nabla h(y^*))(y^k - y^*). \tag{2.17}$$

Como $h_i \in C^2(W)$, se tiene

$$x_i^* - x_i^k = h_i(y_i^*) - h_i(y_i^k) = h_i'(y_i^k)(y_i^* - y_i^k) + \frac{1}{2}h_i''(\zeta_i^k)(y_i^* - y_i^k)^2, \tag{2.18}$$

$$h_i'(y_i^k) - h_i'(y_i^*) = h_i''(\delta_i^k)(y_i^k - y_i^*), \tag{2.19}$$

donde
$\zeta_i^k = y_i^k + \alpha_i^k(y_i^* - y_i^k), \qquad 0 < \alpha_i^k < 1,$

$\delta_i^k = y_i^* + \beta_i^k(y_i^k - y_i^*), \qquad 0 < \beta_i^k < 1, \quad i = 1, 2, ..., n.$
También

$$\nabla h(y^k) = diag(h_1'(y_1^k), h_2'(y_2^k), ..., h_n(y_n^k)), \tag{2.20}$$

$$\nabla h(y^*) = diag(h_1'(y_1^*), h_2'(y_2^*), ..., h_n(y_n^*)). \tag{2.21}$$

Por lo tanto, la ecuación (2.17) puede ser reescrita usando (2.12), (2.18), (2.19), (2.20) y (2.21) como

$$\begin{aligned}
&(v^k - v^*)(y^k - y^*)\\
&= \begin{bmatrix}(\xi_1^k - \xi_1^*) & \cdots & (\xi_n^k - \xi_n^*)\end{bmatrix} diag(h_1'(y_1^k), h_2'(y_2^k), ..., h_n(y_n^k)) \begin{bmatrix}(y_1^k - y_1^*) & \cdots & (y_n^k - y_n^*)\end{bmatrix}\\
&+ \begin{bmatrix}\xi_1^* & \cdots & \xi_n^*\end{bmatrix} diag(h_1'(y_1^k) - h_1'(y_1^*), ..., h_n'(y_n^k) - h_n(y_n^*)) \begin{bmatrix}(y_1^k - y_1^*) & \cdots & (y_n^k - y_n^*)\end{bmatrix}\\
&= [(\xi_1^k - \xi_1^*)h_1'(y_1^k)(y_1^k - y_1^*) + (\xi_2^k - \xi_2^*)h_2'(y_2^k)(y_2^k - y_2^*) + \cdots + (\xi_n^k - \xi_n^*)h_n'(y_n^k)(y_n^k - y_n^*)]\\
&+ [\xi_1^*(h_1'(y_1^k) - h_1'(y_1^*))(y_1^k - y_1^*) + \cdots + \xi_n^*(h_n'(y_n^k) - h_n'(y_n^*))(y_n^k - y_n^*)]\\
&= [(\xi_1^k - \xi_1^*)(h_1(y_1^k) - h_1(y_1^*)) - (\xi_1^k - \xi_1^*)h_1(y_1^k) + (\xi_1^k - \xi_1^*)h_1(y_1^*)\\
&+ (\xi_1^k - \xi_1^*)h_1'(y_1^k)(y_1^k - y_1^*)] + \cdots\\
&+ [(\xi_n^k - \xi_n^*)(h_n(y_n^k) - h_n(y_n^*)) - (\xi_n^k - \xi_n^*)h_n(y_n^k) + (\xi_n^k - \xi_n^*)h_n(y_n^*)\\
&+ (\xi_n^k - \xi_n^*)h_n'(y_n^k)(y_n^k - y_n^*)]\\
&+ \left[\xi_1^* \frac{(h_1'(y_1^k) - h_1'(y_1^*))}{(y_1^k - y_1^*)}(y_1^k - y_1^*)^2 + \cdots + \xi_n^* \frac{(h_n'(y_n^k) - h_n'(y_n^*))}{(y_n^k - y_n^*)}(y_n^k - y_n^*)^2\right]\\
&= \sum_{i=1}^n (\xi_i^k - \xi_i^*)(h_i(y_i^k) - h_i(y_i^*))\\
&+ \sum_{i=1}^n (\xi_i^k - \xi_i^*)h_i(y_i^*) - (\xi_i^k - \xi_i^*)h_i(y_i^k) - (\xi_i^k - \xi_i^*)h_i'(y_i^k)(y_i^* - y_i^k)\\
&+ \sum_{i=1}^n \xi_i^* \frac{(h_i'(y_i^k) - h_i'(y_i^*))}{(y_i^k - y_i^*)}(y_i^k - y_i^*)^2\\
&= \sum_{i=1}^n (\xi_i^k - \xi_i^*)(x_i^k - x_i^*) + \sum_{i=1}^n (\xi_i^k - \xi_i^*)[h_i(y_i^*) - h_i(y_i^k) - h_i'(y_i^k)(y_i^* - y_i^k)]\\
&+ \sum_{i=1}^n \xi_i^*(h_i''(\delta_i^k))(y_i^k - y_i^*)^2\\
&= (\xi^k - \xi^*)(x^k - x^*) + \sum_{i=1}^n (\xi_i^k - \xi_i^*)\left(\frac{1}{2}h_i''(\zeta_i^k)\right)(y_i^* - y_i^k)^2 + \sum_{i=1}^n \xi_i^* h_i''(\delta_i^k)\frac{(x_i^k - x_i^*)^2}{h_i'(\eta_i^k)^2}\\
&= (\xi^k - \xi^*)(x^k - x^*) + \frac{1}{2}\sum_{i=1}^n (\xi_i^k - \xi_i^*)h_i''(\zeta_i^k)(y_i^* - y_i^k)^2 + \sum_{i=1}^n \xi_i^* \frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}(x_i^k - x_i^*)^2,
\end{aligned}$$

así

$$\begin{aligned}
(v^k - v^*)(y^k - y^*) = (\xi^k - \xi^*)(x^k - x^*) &+ \frac{1}{2}\sum_{i=1}^n (\xi_i^k - \xi_i^*)h_i''(\zeta_i^k)(y_i^* - y_i^k)^2\\
&+ \sum_{i=1}^n \xi_i^* \frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}(x_i^k - x_i^*)^2,
\end{aligned} \tag{2.22}$$

donde
$\eta_i^k \longrightarrow y_i^*, \quad \zeta_i^k \longrightarrow y_i^*, \quad \delta_i^k \longrightarrow y_i^*, \qquad \text{cuando } k \longrightarrow \infty, \quad i = 1, 2, ..., n,$

ahora como

$$a)\ \lim_{k\longrightarrow\infty}\frac{(y_i^k-y_i^*)^2}{t_k^2}=d_i^2,$$
$$b)\ \lim_{k\longrightarrow\infty}(\xi_i^k-\xi_i^*)=0,$$
$$c)\ \lim_{k\longrightarrow\infty}h_i''(\zeta_i^k)=h_i''(y_i^*),$$

combinando $a)$, $b)$ y $c)$ y por teorema (6)

$$\lim_{k\longrightarrow\infty}(\xi_i^k-\xi_i^*)h_i''(\zeta_i^k)\frac{(y_i^k-y_i^*)^2}{t_k^2}=0. \tag{2.23}$$

Dividiendo por t_k^2 a (2.22), se tiene

$$\begin{aligned}\frac{(v^k-v^*)(y^k-y^*)}{t_k^2}&=\frac{(\xi^k-\xi^*)(x^k-x^*)}{t_k^2}+\frac{1}{2}\sum_{i=1}^n(\xi_i^k-\xi_i^*)h_i''(\zeta_i^k)\frac{(y_i^*-y_i^k)^2}{t_k^2}\\&+\sum_{i=1}^n\xi_i^*\frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}\frac{(x_i^k-x_i^*)^2}{t_k^2},\end{aligned} \tag{2.24}$$

aplicando $\liminf$ a (2.24), cuando $k\longrightarrow\infty$ y aplicando el teorema (**??**), se obtiene

$$\begin{aligned}&\liminf_{k\longrightarrow\infty}\frac{(v^k-v^*)(y^k-y^*)}{t_k^2}\\&=\liminf_{k\longrightarrow\infty}[\frac{(\xi^k-\xi^*)(x^k-x^*)}{t_k^2}+\frac{1}{2}\sum_{i=1}^n(\xi_i^k-\xi_i^*)h_i''(\zeta_i^k)\frac{(y_i^*-y_i^k)^2}{t_k^2}\\&+\sum_{i=1}^n\xi_i^*\frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}\frac{(x_i^k-x_i^*)^2}{t_k^2}]\end{aligned}$$

$$\begin{aligned}&\geq\liminf_{k\longrightarrow\infty}\frac{(\xi^k-\xi^*)(x^k-x^*)}{t_k^2}+\liminf_{k\longrightarrow\infty}\frac{1}{2}\sum_{i=1}^n(\xi_i^k-\xi_i^*)h_i''(\zeta_i^k)\frac{(y_i^*-y_i^k)^2}{t_k^2}\\&+\liminf_{k\longrightarrow\infty}\sum_{i=1}^n\xi_i^*\frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}\frac{(x_i^k-x_i^*)^2}{t_k^2}\\&=\liminf_{k\longrightarrow\infty}\frac{(\xi^k-\xi^*)(x^k-x^*)}{t_k^2}+\frac{1}{2}\sum_{i=1}^n\lim_{k\longrightarrow\infty}(\xi_i^k-\xi_i^*)h_i''(\zeta_i^k)\frac{(y_i^*-y_i^k)^2}{t_k^2}\\&+\sum_{i=1}^n\lim_{k\longrightarrow\infty}\xi_i^*\frac{h_i''(\delta_i^k)}{h_i'(\eta_i^k)^2}\frac{(x_i^k-x_i^*)^2}{t_k^2},\end{aligned}$$

por (2.23), y usando (2.14)

$$\liminf_{k\longrightarrow\infty}\frac{(v^k-v^*)(y^k-y^*)}{t_k^2}\geq\liminf_{k\longrightarrow\infty}\frac{(\xi^k-\xi^*)(x^k-x^*)}{t_k^2}+\sum_{i=1}^n\xi_i^*\frac{h_i''(y_i^*)}{h_i'(y_i^*)^2}u_i^2$$
$$g_-''(y^*,v^*,d)\geq f_-''(x^*,\xi^*,u)+\sum_{i=1}^n\xi_i^*\frac{h_i''(y_i^*)}{h_i'(y_i^*)^2}u_i^2.$$

Por propiedad se cumple que $f_-''(x,v,\lambda w)=\lambda^2f_-''(x,v,w)$ para cualquier $\lambda>0$ y usando (2.9) y (2.10),se tiene

$$\begin{aligned}g_-''(y^*,v^*,d)&\geq f_-''(x^*,\xi^*,u)+\frac{\sigma}{\epsilon}\sum_{i=1}^n\xi_i^*u_i^2\\&\geq\|u\|^2f_-''(x^*,\xi^*,u/\|u\|)+\frac{\sigma}{\epsilon}\epsilon\|u\|^2\\&\geq\|u\|^2(-\sigma+\sigma)\\&=0.\end{aligned}$$

Entonces $g''_{-}(y^*, v^*, d) \geq 0$.
Aplicando el lema (1), se concluye que g es una función convexa. ∇

A continuación se definen dos clases de funciónes que satisfacen la condición (2.11).

Corolario 3 *Sea f una función definida en X. Suponer que f satisface las condiciones del teorema (9). Definir*

$$h_j(y_j) = \frac{1}{p} \ln\left(1 - \frac{1}{y_j}\right), \qquad p > 0, \qquad j = 1, 2, ..., n,$$

entonces $g(y) = f(h(y))$ es una función convexa en $Y = \left[\frac{1}{(1-e^{pl_j})}, \frac{1}{(1-e^{pL_j})}\right]$ si existe una constante $p_1 \geq 0$ tal que $p \geq p_1$.

Demostración
Será suficiente probar que se cumple la condición (2.11).
En efecto:

$$\begin{aligned} h'_j(y_j) &= \frac{1}{p}\left(\frac{1}{1-\frac{1}{y_j}}\right)\left(1 - \frac{1}{y_j}\right)' \\ &= \frac{1}{p}\left(\frac{1}{\frac{y_j-1}{y_j}}\right)\left(\frac{1}{y_j^2}\right) \\ &= \frac{1}{p}\left(\frac{1}{y_j(y_j-1)}\right) \end{aligned}$$

$$\begin{aligned} h''_j(y_j) &= \frac{1}{p}\left(\frac{-(y_j(y_j-1))'}{y_j^2(y_j-1)^2}\right) \\ &= \frac{1}{p}\left(\frac{-((y_j-1)+y_j)}{y_j^2(y_j-1)^2}\right) \\ &= \frac{1}{p}\left(\frac{1-2y_j}{y_j^2(y_j-1)^2}\right), \end{aligned}$$

se tiene que $y_j < 0$ para $y_j \in \left[\frac{1}{(1-e^{pl_j})}, \frac{1}{(1-e^{pL_j})}\right]$, entonces

$$\begin{aligned} \frac{h''_j(y_j)}{[h'_j(y_j)]^2} &= \frac{\frac{1}{p}\left(\frac{1-2y_j}{y_j^2(y_j-1)^2}\right)}{\left(\frac{1}{p}\left(\frac{1}{y_j(y_j-1)}\right)\right)^2} \\ &= \frac{p^2(1-2y_j)(y_j(y_j-1))^2}{py_j^2(y_j-1)^2} \\ &= p(1-2y_j) \\ &> p. \end{aligned}$$

Ahora para que la condición (2.11) se cumpla, se debe tener que $p \geq \frac{\sigma}{\epsilon}$, pero $p > 0$.
Sea

$$p_1 = max\left\{0, \frac{\sigma}{\epsilon}\right\}.$$

Por lo tanto g es convexa cuando $p \geq p_1$. ∇

Corolario 4 *Sea f una función definida en X. Suponer que f satisface las condiciones del teorema (9). Definir*

$$h_j(y_j) = y_j^{-\frac{1}{p}}, \qquad p > 0, \qquad j = 1, 2, ..., n.$$

entonces $g(y) = f(h(y))$ *es una función convexa en* $Y = \prod_{j=1}^{n}[L_j^{-p}, l_j^{-p}]$ *si existe un* $p_2 \geq 0$ *tal que* $p \geq p_2$.

Demostración
Será suficiente probar que se cumple la condición (2.11).
En efecto:

$$\begin{aligned} h_j'(y_j) &= -\frac{1}{p}y_j^{\left(-\frac{1}{p}-1\right)} \\ h_j''(y_j) &= \left(-\frac{1}{p}\right)\left(-\frac{1}{p}-1\right)y_j^{\left(-\frac{1}{p}-2\right)} \\ &= \left(\frac{1}{p}\right)\left(\frac{1}{p}+1\right)y_j^{\left(-\frac{1}{p}-2\right)}, \end{aligned}$$

ahora para todo $y_j \in \left[L_j^{-p}, l_j^{-p}\right]$, se tiene

$$L_j^{-p} \leq y_j \leq l_j^{-p}$$

$$L_j^{-1} \leq y_j^{\frac{1}{p}} \leq l_j^{-1},$$

luego $y_j^{\frac{1}{p}} \geq L_j^{-1} = \frac{1}{L_j}$ y sea $L = \text{máx}_{1 \leq j \leq n} L_j$.
Entonces

$$\begin{aligned} \frac{h_j''(y_j)}{(h_j'(y_j))^2} &= \frac{\left(\frac{1}{p}\right)\left(\frac{1}{p}+1\right)y_j^{\left(-\frac{1}{p}-2\right)}}{\left(-\frac{1}{p}y_j^{\left(-\frac{1}{p}-1\right)}\right)^2} \\ &= (1+p)y_j^{\frac{1}{p}} \\ &\geq (1+p)\frac{1}{L_j} \\ &\geq (1+p)\frac{1}{L}. \end{aligned}$$

Ahora para q la condición (2.11) se cumpla,se debe tener que $(1+p)\frac{1}{L} \geq \frac{\sigma}{\epsilon}$,
entonces $p \geq \frac{\sigma}{\epsilon}L - 1$ pero $p > 0$.
Sea

$$p_2 = max\left\{0, \frac{\sigma}{\epsilon}L - 1\right\}.$$

Por lo tanto g es convexa cuando $p \geq p_2$. ∇

A continuación se presenta un ejemplo.

ejemplo 3 *Considerar la función*

$$f(x) = max\{f_1(x), f_2(x)\} \qquad\qquad x \in X = [2, 5],$$

donde

$$f_1(x) = -(1/4)(x-5)^2 + 4, \qquad\qquad f_2(x) = (x-4)|x-4| + 2x - 2.$$

Aplicando el corolario (3), hallar el valor de p para garantizar la convexidad de $g(y) = f(h(h))$.
Solución
Se observa que f *es una función no convexa y no diferenciable, además se verifica que* f *es una función semisuave y regular en* x.

Reescribiendo la función f, se tiene

$$f(x) = \begin{cases} -(1/4)(x-5)^2+4, & x \in [2,3\rangle, \\ -(x-4)^2+2x-2, & x \in [3,4], \\ (x-4)^2+2x-2, & x \in \langle 4,5]. \end{cases}$$

Ahora hallar la subdiferencial de f
∗ Para $x \in [2,3\rangle$, se tiene por teorema(5), que
$\partial f(x) = -(1/2)(x-5)$, pues f es una función continuamente diferenciable en $[2,3\rangle$.
∗ Para $x = 3$ y por teorema (4) se tiene que $\partial f(x) = \partial_c f(x)$, entonces

$$\begin{aligned} \partial f(3) &= \lambda f_1'(3) + (1-\lambda) f_2'(3), && \forall\, \lambda \in [0,1] \\ &= \lambda(1) + (1-\lambda)(4) \\ &= 4 - 3\lambda && \forall\, \lambda \in [0,1], \end{aligned}$$

por tanto $\partial f(x) = [1,4]$.
∗ Para $x \in \langle 3,4]$, se tiene por teorema(5), que
$\partial f(x) = -2(x-5)$, pues f es una función continuamente diferenciable en $\langle 3,4]$.
∗ Para $x \in \langle 4,5]$, se tiene por teorema(5), que
$\partial f(x) = 2(x-3)$, pues f es una función continuamente diferenciable en $\langle 4,5]$.
Notar que $\xi \geq 1$ para cualquier $\xi \in \partial f(x)$.

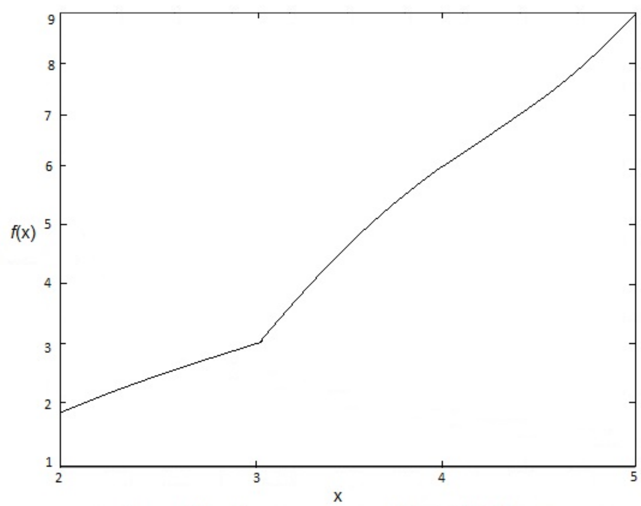

Figura 4. función no convexa y no diferenciable f(x)

Así $\epsilon = 1$. Para $|u| = 1$ y $v \in \partial_u f(x)$, se tiene

$$f_{-}''(x,v,u) = \begin{cases} -1/2, & x \in [2,3\rangle, \\ -2, & x \in \langle 3,4\rangle, \\ 2, & x \in \langle 4,5]. \end{cases}$$

Faltaría analizar para $x = 3$ y $x = 4$.
Para $x = 4$, se tiene que $v = 2$, donde $v \in \partial_u f(4)$, luego

$$\begin{aligned} &f_{-}''(4,2,-1) = -2, \\ &f_{-}''(4,2,1) = 2. \end{aligned}$$

Para $x = 3$, $|u| = 1$ y $v \in \partial_u f(3)$, se tiene

$$f_{-}''(3,v,u) = max\{\lambda_1 f_1''(3) + \lambda_2 f_2''(3) :\ (\lambda_1,\lambda_2) \in S(3,v)\},$$

donde

$$S(3,v)) = \{(\lambda_1,\lambda_2) :\ v = \lambda_1 f_1'(3) + \lambda_2 f_2'(3),\ \lambda_1,\ \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1\}.$$

Para

$$f_1'(3) = 1, \qquad f_2'(3) = 4, \qquad f_1''(3) = -1/2, \qquad f_2''(3) = -2,$$

se tiene

$$f''_-(3,1,-1) = -1/2,$$
$$f''_-(3,4,1) = -2,$$

entonces para $v \in (1,4)$ *y* $|u| = 1$,

$$f''_-(3,v,u) = -1/2.$$

Por lo tanto, para cualquier $x \in [2,5]$, *y* $v \in \partial_u f(x)$,

$$f''_-(x,v,u) \geq -2 = \rho,$$

pero $\sigma = -\rho = 2$.
Luego, por el corolario (3),

$$p_1 = max\left\{0, \frac{\sigma}{\epsilon}\right\} = max\left\{0, \frac{2}{1}\right\} = 2.$$

Por lo tanto, para cualquier $p \geq 2$ *garantiza la convexidad de g(y)=f(h(y)).*

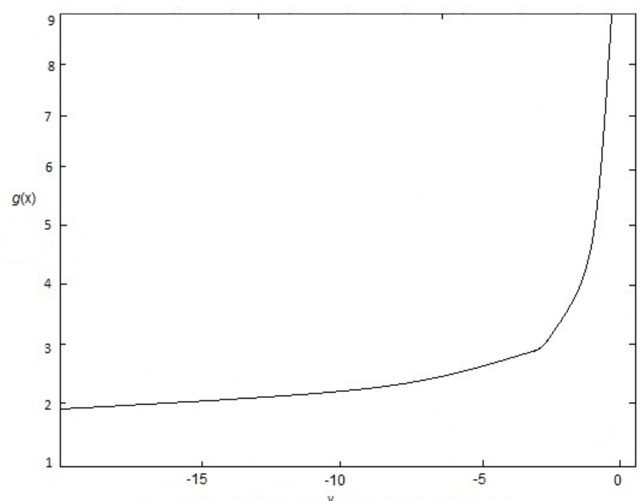

Figura 5. Función convexificada *g*(y) con *p*=2

Los resultados dados en esta sección proporcionan una base teórica para extender el alcance de los métodos de convexificación para funciones estrictamente monótonas no diferenciables.

3. Resultados. Los principales resultados del presente trabajo están dados en el teorema 8 y los corolarios 1 y 2 para la convexificación de funciones estrictamente monotonas dos veces continuamente diferenciables y el teorema 9 con los corolarios 3 y 4 para la convexificación de funciones estrictamente monotonas no diferenciables. Proporcionan una base teórica para extender el alcance de los métodos de convexificación para funciones estrictamente monótonas dos veces continuamente diferenciables y no diferenciables que se presentan en problemas de optimización global.

Agradecimientos. Mi agradecimiento a todos los amigos, docentes y estudiantes del Departamento de Matemáticas de la Facultad de Ciencias Físicas y Matemáticas de la Universidad Nacional de Trujillo por su apoyo moral y académico en el desarrollo del presente trabajo.

Referencias

[1] Bazaraa, M., *Programación Lineal y Flujos y Redes*, segunda edición, Noriega Editores, México, 2004.
[2] Bazaraa, M., *Nonlinear Programming: Theory and Algorithms*, Third edition, Wiley Interscience, New Jersey, 2006.
[3] Clarke, F., *Optimization and Nonsmooth Analysis*, University of Montreal, New York, 1983.
[4] Chaney, R., *Second-Order Directional Derivatives for Nonsmooth Functions*, Journal of Mathematical Analysis and Applications 128, 495-511, 1987.
[5] Dara, A. and Dutta, J., *Optimality conditions in convex optimization a Finite-Dimensional View*, Taylor and Francis Group, London, 2012.
[6] Hofman, K., *A Method for Globally Minimizaing Concave Functions Over Convex Sets*, Mathematical Programming 20 (1981) 22-32.
[7] Li, D., Sun, X. and McKinnon, K., *An Exact Solution Method for Reliability Optimization Problems in Complex Systems*, Annals of Operation Research 133, 129-148, 2005.
[8] Sun, X., Luo, H. and Li, D., *Convexification of Nonsmooth Monotone Functions*, Journal of Optimization Theory and Applications, Vol.132, No 2, 2007.
[9] Sun, X., Mckinnon, H. and Li, D., *A Convexification method for a class of global optimization problems with applications to reliability optimization*, Journal of Global Optimization 21: 185-199,2001.
[10] Wu, Z., Zhang, L., Bai, F. AND Yang, X., *Convexification and Concavification Methods for some Global Optimization Problems*, Journal of Systems Science and Complexity, Vol. 17, No. 3, 2004.