

UNIVERSIDAD CENTROCCIDENTAL  
"LISANDRO ALVARADO"

Decanato de Ciencias y Tecnología.  
Coordinación de Postgrado.  
Maestría en Ciencias Mención Optimización.



ALGORITMO PROXIMAL BASADO EN UNA  
CASI-DISTANCIA HOMOGÉNEA DE ORDEN  $p$  PARA  
PROGRAMACIÓN CASI-CONVEXA DIFERENCIABLE.

TRABAJO DE GRADO PRESENTADO POR:

LCDO. JESÚS MEDINA.

C.I.V.:17.611.280

COMO REQUISITO FINAL  
PARA OBTENER EL TÍTULO DE  
MAGÍSTER SCIENTIARUM MENCIÓN OPTIMIZACIÓN.

TUTOR: DR. RÓMULO CASTILLO.  
ÁREA DE CONOCIMIENTO: OPTIMIZACIÓN.

Barquisimeto - Venezuela

Octubre, 2013



*A mis padres.*



# AGRADECIMIENTOS

Primeramente a Dios por darme la perseverancia y la sabiduría.

A mis padres, a Dellys, a mis hermanas y a la familia Matos Bernal por su apoyo incondicional, por su paciencia y ayuda cuando más lo necesite.

A mi tutor Rómulo Castillo por su gran orientación, dedicación y asesoría. Gracias por compartir sus conocimientos que me fueron de gran ayuda.

A cada uno de los profesores y compañeros que participaron en mi desarrollo profesional, con especial agradecimiento a los profesores Hugo Lara y Rómulo Castillo.

A nuestra gran casa UCLA por abrirme sus puertas y a sus beneficios adquiridos durante todo este tiempo.



# RESUMEN

En este trabajo, proponemos un algoritmo de método de punto proximal interior basado en una casi-distancia tipo entrópica homogénea de orden  $p$ , con  $p \geq 2$ , para la minimización de la función casi-convexa, acotada inferiormente y continuamente diferenciable, sujeta a restricciones de no-negatividad. Además se establece la buena definición de la sucesión generada por el algoritmo y obtenemos dos resultados de convergencia importantes, una de éstas es que la sucesión converge a un punto estacionario y la otra son las condiciones suficientes para que el punto de convergencia sea solución del problema. Esto fue considerado por [12] para el caso  $p = 2$ .

**Palabras claves:** Método punto proximal, programación casi-convexa y casi-distancia.



# Notaciones

- $\langle \cdot, \cdot \rangle$ : Denota el producto interno Euclideo en  $\mathbb{R}^n$ .
- $\|\cdot\|$ : Denota la norma Euclidea en  $\mathbb{R}^n$ .
- $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x \geq 0\}$ .
- $\mathbb{R}_{++}^n = \{x \in \mathbb{R}^n \mid x > 0\}$ .

Dada la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  diferenciable,

- $\nabla f(x)$ : Representa el gradiente de  $f$  con respecto a  $x$ .
- $(\nabla f(x))_i$ : Denota la  $i$ -ésima derivada parcial de  $f$  con respecto a  $x$ .
- $\nabla_1 d_\phi(x, y)$ : Denota la derivada parcial de  $d_\phi$  con respecto a la primera componente.



# ÍNDICE

<b>Agradecimientos</b>	<b>i</b>
<b>Resumen</b>	<b>iii</b>
<b>Introducción</b>	<b>1</b>
<b>1. Preliminares.</b>	<b>5</b>
1.1. Conceptos básicos . . . . .	5
<b>2. Función casi-distancia <math>d_\phi</math> y sus propiedades.</b>	<b>13</b>
<b>3. Método tipo-proximal Interior.</b>	<b>25</b>
3.1. Algoritmo . . . . .	25
3.2. Análisis de Convergencia. . . . .	28
<b>4. Conclusiones y trabajo futuro.</b>	<b>33</b>
<b>Referencias</b>	<b>35</b>



# Índice de figuras

1.1. Gráfica de una función casi-convexa. . . . .	7
1.2. Gráfica de una función casi-convexa y sus conjuntos de nivel. . . . .	8
2.1. Gráfica de la función $\varphi_1$ . . . . .	14
2.2. Gráfica de la función $\varphi_2$ . . . . .	15



# INTRODUCCIÓN

El algoritmo de punto proximal para minimización de una función convexa  $f(x)$  en  $\mathbb{R}^n$  genera una sucesión  $\{x^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}^n$ , dada por el siguiente esquema iterativo:

$$x^{k+1} = \operatorname{argmín}_{x \in \mathbb{R}^n} \left\{ f(x) + \lambda_k \|x - x^k\|^2 \right\}, \quad (1)$$

donde  $\lambda_k$  es una sucesión de números positivos y  $\|\cdot\|$  denota la norma euclídeana en  $\mathbb{R}^n$  y es denominado núcleo cuadrático en este caso. El método fue originalmente introducido por Martinet [7] y está basado en la aproximación proximal de Moreau [8] (1965) de  $f$ , definida por

$$f_\mu(x) = \inf_u \left\{ f(u) + \frac{1}{2\mu} \|x - u\|^2 \right\}, \quad \mu > 0. \quad (2)$$

El método de punto proximal fue posteriormente estudiado y desarrollado por Rockafellar [10, 11] (1976) estableciendo una relación con los métodos del tipo lagrangeano aumentado. En 1992, Teboulle [13] introdujo el método de punto proximal entrópico basado en la imitación del mapeo proximal de Moreau, que sustituye a la distancia cuadrática en (1) y (2) con la siguiente casi-distancia tipo-entrópica, también llamada  $\varphi$ -divergencia:

$$d_\varphi(x, y) = \sum_{i=1}^n y_i \varphi(x_i/y_i), \quad (3)$$

donde  $\varphi : \mathbb{R} \rightarrow (-\infty, \infty]$  es una función estrictamente convexa, propia y cerrada que satisface ciertas condiciones. Una elección importante de  $\varphi$  es el caso en que  $\varphi(t) = t \ln(t) - t + 1$ , la cual, con la correspondiente  $d_\varphi$  es conocida como la función Kullback-Leibler entrópica, que involucra el término  $t \ln(t)$ , de allí el término entrópico [9, 13].

El algoritmo tipo-proximal basado en  $\varphi$ -divergencia fue originalmente usado para minimizar una función convexa sujeto a restricciones no negativas  $x \geq 0$ , la cual

genera una sucesión  $\{x^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}_{++}^n$  definida como:

$$\begin{cases} x^0 > 0 \\ x^{k+1} = \operatorname{argmín}_{x \geq 0} \{f(x) + \lambda_k d_\varphi(x, x^k)\}. \end{cases} \quad (4)$$

En [12] se enfocaron en una clase de algoritmo tipo-proximal de la forma (4) para minimizar una función casi-convexa con una función casi-distancia homogénea de segundo orden, definida por:

$$d_\phi(x, y) := \begin{cases} \sum_{i=1}^n y_i^2 \phi\left(\frac{x_i}{y_i}\right), & x, y \in \mathbb{R}_{++}^n \\ +\infty, & \text{en otro caso,} \end{cases} \quad (5)$$

donde  $\phi$  representa el núcleo y está definida con uno de dos tipos especiales de  $\varphi$  y una función cuadrática. Esta función  $\phi$  satisface ciertas propiedades que son desarrolladas para la minimización y la convergencia del siguiente problema:

$$\begin{aligned} & \operatorname{mín} f(x) \\ & \text{s.a. } x \geq 0, \end{aligned} \quad (6)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función propia, casi-convexa, acotada inferiormente y diferenciable. Puesto que no requerimos la convexidad de la función  $f$  el esquema iterativo para el algoritmo es la siguiente:

$$\begin{cases} x^0 > 0 \\ x^{k+1} \in \operatorname{argmín}_{x \geq 0} \{f(x) + \lambda_k d_\phi(x, x^k)\}, \end{cases} \quad (7)$$

donde  $\lambda_k$  es una sucesión de números positivos. En este artículo se estableció la convergencia total de la sucesión  $\{x^k\}_{k \in \mathbb{N}}$  generada por (7).

Otros trabajos consideran núcleos llamados distancias de Bregman, con diferentes propiedades a las casi-distancias que aquí consideramos, sin embargo, solamente los nombramos puesto que no serán objeto de nuestro estudio principal.

Finalmente, el problema a considerarse en el presente trabajo es el de estudiar el algoritmo proximal de tipo-entrópico considerados en [12] pero ahora con una casi-distancia homogénea de orden  $p$ , definida por:

$$d_\phi(x, y) := \begin{cases} \sum_{i=1}^n y_i^p \phi\left(\frac{x_i}{y_i}\right), & \text{con } p \geq 2, \quad x, y \in \mathbb{R}_{++}^n \\ +\infty, & \text{en otro caso,} \end{cases} \quad (8)$$

la cual fue usada en [4] para programación convexa y en este caso la usaremos para minimizar en el primer ortante con una función casi-convexa.

El propósito de este trabajo es establecer la convergencia completa de la sucesión  $\{x^k\}_{k \in \mathbb{N}}$  generada por (7) con la casi-distancia definida en (8), bajo ciertas suposiciones para el problema casi-convexo (6).

El resto de este trabajo esta organizado de la siguiente manera: en el capítulo 1, presentamos algunas definiciones y resultados básicos que serán usados en los siguientes capítulos. En el capítulo 2, presentamos la definición del núcleo  $\phi$  y las propiedades de la casi-distancia  $d_\phi$ . Basado en la función casi-distancia homogénea tipo-entrópica de orden  $p$ ,  $d_\phi$ , en el capítulo 3, se define el algoritmo tipo-proximal interior y se prueba la convergencia completa de la sucesión generada por este algoritmo. Finalmente, en el capítulo 4, presentamos las conclusiones y trabajo futuro.



# CAPÍTULO 1

## PRELIMINARES.

### 1.1. Conceptos básicos

En esta sección, presentaremos algunos resultados preliminares que serán usados en los siguientes capítulos.

Debido a que el algoritmo con que trabajaremos está basado en una casi-distancia, daremos la siguiente definición de función casi-distancia en un conjunto no vacío  $\mathcal{X}$ .

**Definición 1.1.** [5] Sea  $\mathcal{X} \neq \emptyset$ , con  $\mathcal{X} \subset \mathbb{R}^n$ , una función  $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ , se denomina casi-distancia en  $\mathcal{X}$  si, y sólo si, para todo  $x, y \in \mathcal{X}$  se cumplen:

$$(i) \quad d(x, y) \geq 0,$$

$$(ii) \quad d(x, y) = 0 \Leftrightarrow x = y.$$

Se dice que  $d$  es una casi-distancia en  $\mathcal{X}$ , debido a que le falta satisfacer la simetría y la desigualdad triangular, para que sea una distancia o métrica en  $\mathcal{X}$ .

Ahora definimos cuando una sucesión es Fejér convergente asociada a una casi-distancia.

**Definición 1.2.** (ver [5]) Una sucesión  $\{y^k\}_{k \in \mathbb{N}}$  es Fejér convergente en un conjunto no vacío  $U \subseteq \mathbb{R}^n$  con respecto a una función casi-distancia  $d(\cdot, \cdot)$ , si para cualquier  $u \in U$  tenemos que  $d(u, y^{k+1}) \leq d(u, y^k)$ . Cuando  $d$  es la distancia euclídeana,  $\{y^k\}$  es llamada Fejér convergente en  $U$ .

Tenemos el siguiente resultado, el cual nos será útil para la demostración de uno de los teoremas importantes de análisis de convergencia.

**Proposición 1.1.** (ver [5]) Si  $\{y^k\}_{k \in \mathbb{N}}$  es Fejér convergente en un conjunto no vacío  $U \subseteq \mathbb{R}^n$  entonces  $\{y^k\}_{k \in \mathbb{N}}$  es acotado. Si un punto clausura  $y$  de  $\{y^k\}$ , con  $y \in U$  entonces  $y = \lim_{k \rightarrow +\infty} y^k$ .

*Demostración.* Como  $\{y^k\}_{k \in \mathbb{N}}$  es Fejér convergente en  $U$ , entonces  $\|y^k - u\| \leq \|y^0 - u\|$  para cualquier  $u \in U$ . Así,  $\{y^k\}$  está contenida en una bola de centro  $u$  y radio  $\|y^0 - u\|$ , por tanto es acotada.

Para la segunda parte, sea  $\{y^{j_k}\}$  una subsucesión de  $\{y^k\}$ , tal que  $\lim_{k \rightarrow \infty} y^{j_k} = y$ . Como  $y \in U$ , por ser  $\{y^k\}_{k \in \mathbb{N}}$  Fejér convergente en  $U$ , entonces la sucesión  $\{\|y^k - y\|\}$  es decreciente y no-negativo, y tiene una subsucesión  $\{\|y^{j_k} - y\|\}$  el cual converge a 0. Entonces la sucesión converge a 0, es decir,  $0 = \lim_{k \rightarrow \infty} \|y^k - y\|$ ; lo que implica  $y = \lim_{k \rightarrow +\infty} y^k$ .  $\square$

La siguiente definición básica, será usada en la mayoría de las definiciones y resultados posteriores.

**Definición 1.3.** [9] Dada la función de valor real extendida  $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ , denotamos su dominio por:

$$\text{dom} f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$$

y su epígrafo por:

$$\text{epi} f := \{(x, \beta) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \beta\}.$$

Entonces,  $f$  se dice propia si  $\text{dom} f \neq \emptyset$  y  $f(x) > -\infty, \forall x \in \text{dom} f$ , y  $f$  es una función cerrada o semicontinua inferior si  $\text{epi} f$  es un subconjunto cerrado de  $\mathbb{R}^n \times \mathbb{R}$ .

Como en este trabajo se va a trabajar con problemas de minimización diferenciable, necesitamos la siguiente definición.

**Definición 1.4.** [9] Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función diferenciable. Entonces,

- (a) Para un problema de optimización irrestricta de minimización de  $f(x)$  con  $x \in \mathbb{R}^n$ ,  $x^*$  es llamado un punto estacionario si  $\nabla f(x^*) = 0$ .

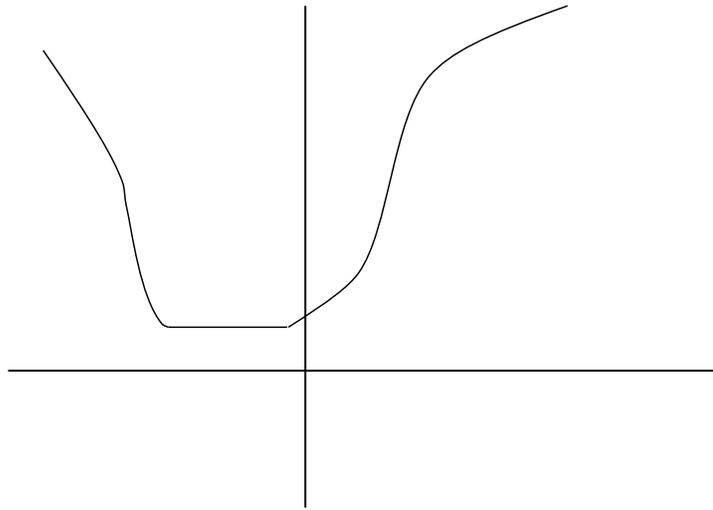
- (b) Para un problema de optimización restringida de minimización de  $f(x)$  con  $x \in \mathcal{C}$ , donde  $\mathcal{C}$  es un subconjunto no vacío y convexo de  $\mathbb{R}^n$ , es llamado punto estacionario si

$$\nabla f(x^*)^t(x - x^*) \geq 0, \forall x \in \mathcal{C}.$$

En lo que sigue presentamos la definición de un función casi-convexa, el cual es el principal objeto de estudio de nuestro trabajo.

**Definición 1.5.** [6] Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función propia. Entonces  $f$  es llamada casi-convexa si y sólo si, para cualquier  $x, y \in \text{dom}f$  y  $\beta \in (0, 1)$ , se cumple:

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



**Figura 1.1:** Gráfica de una función casi-convexa.

Con el proximo resultado, tenemos una caracterización de una función casi-convexa mediante sus conjuntos de nivel. Este resultado será utilizado para demostrar que le conjunto  $U$  usado en el análisis de convergencia, sea convexo.

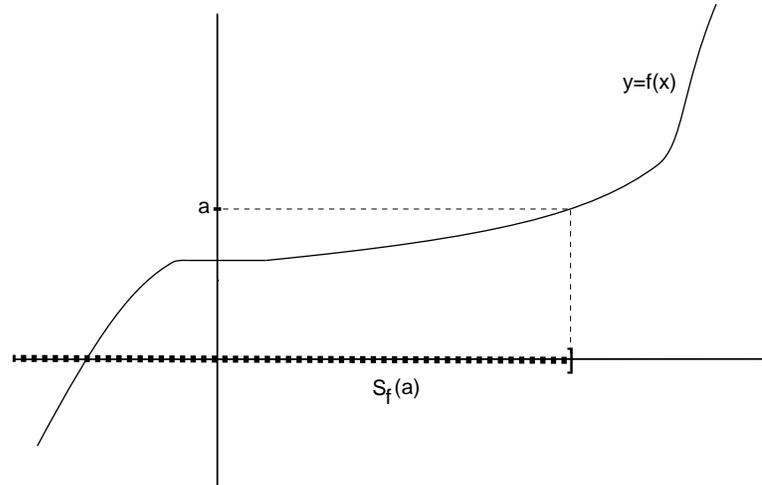
**Proposición 1.2.** [1] *La función propia  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es casi-convexa si y sólo si el conjunto de nivel  $S_f(\alpha) := \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$  es convexo para cualquier  $\alpha \in \mathbb{R}$ .*

*Demostración.* ( $\Rightarrow$ ) Supongamos que  $f$  es casi-convexa, y sean  $x, y \in S_f(\alpha)$  entonces, para  $\alpha \in \mathbb{R}$  se tiene que,  $f(x) \leq \alpha$  y  $f(y) \leq \alpha$ . Así, como  $x, y \in \text{dom} f$  entonces:

$$\begin{aligned} f(\beta x + (1 - \beta)y) &\leq \max\{f(x), f(y)\}, \forall \beta \in (0, 1), \\ &\leq \alpha. \end{aligned}$$

Por lo tanto,  $S_f(\alpha)$  es convexo.

( $\Leftarrow$ ) Ahora supongamos que  $S_f(\alpha)$  es convexo para cualquier  $\alpha \in \mathbb{R}$ . Sean  $x, y \in \text{dom} f$  entonces  $x, y \in S_f(\alpha)$ , tomando  $\alpha = \max\{f(x), f(y)\}$ , luego  $\beta x + (1 - \beta)y \in S_f(\alpha)$ ,  $\forall \beta \in (0, 1)$  (por ser  $S_f(\alpha)$  es convexo), entonces  $f(\beta x + (1 - \beta)y) \leq \alpha = \max\{f(x), f(y)\}$ ,  $\forall \beta \in (0, 1)$ . Así,  $f$  es casi-convexa en  $\mathbb{R}^n$ .  $\square$



**Figura 1.2:** Gráfica de una función casi-convexa y sus conjuntos de nivel.

Ahora presentamos otras caracterizaciones, pero para funciones casi-convexas diferenciables, el cual serán usadas para demostrar el Lema 1.1 (este lema es importante para el análisis de convergencia del algoritmo).

**Teorema 1.1.** *Sea  $\varphi$  una función diferenciable en un intervalo  $\mathcal{I} \subseteq \mathbb{R}$ . Entonces  $\varphi$  es casi-convexa en  $\mathcal{I}$  si y sólo si la siguiente implicación se cumple:*

$$t_1, t_2 \in \mathcal{I}, \quad \varphi(t_2) \leq \varphi(t_1) \Rightarrow \varphi'(t_1)(t_2 - t_1) \leq 0. \quad (1.1)$$

*Demostración.* Ver demostración en [3] página 42. □

**Proposición 1.3.** *[3] Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es diferenciable. Entonces  $f$  es casi-convexa si y sólo si para cualquier  $x, y \in \mathbb{R}^n$  tal que,*

$$f(x) \leq f(y) \Rightarrow \langle \nabla f(y), x - y \rangle \leq 0. \quad (1.2)$$

*Demostración.* ( $\Rightarrow$ ) Supongamos que  $f$  es casi-convexa y sean  $x, y \in \mathbb{R}^n$  con  $f(x) \leq f(y)$ . Consideremos la restricción  $\varphi(t) = f(y + t(x - y))$ ,  $t \in [0, 1]$ . Tenemos que  $\varphi'(t) = \nabla f(y + t(x - y))^t(x - y)$  y además  $\varphi(1) = f(x) \leq f(y) = \varphi(0)$ , así por Teorema 1.1 implica que

$$\begin{aligned} \varphi'(0)(1 - 0) \leq 0 &\Rightarrow \nabla f(y)^t(x - y) \leq 0, \\ &\Rightarrow \langle \nabla f(y), x - y \rangle \leq 0. \end{aligned}$$

Por lo tanto (1.2) se cumple.

( $\Leftarrow$ ) Supongamos por absurdo que  $f$  no es casi-convexa, entonces existe una restricción  $\varphi(t) = f(y + t(x - y))$ ,  $x, y \in \mathbb{R}^n$ ,  $t \in [0, 1]$ , la cual no es casi-convexa (por Teorema 2.2.10 de [3]) y en consecuencia del Teorema 1.1, existe  $t_1, t_2 \in [0, 1]$ , tal que,  $\varphi(t_2) \leq \varphi(t_1)$  y  $\varphi'(t_1)(t_2 - t_1) > 0$ . Sea  $\bar{y} = y + t_1(x - y)$  y  $\bar{x} = y + t_2(x - y)$ , entonces,  $f(\bar{x}) = f(y + t_2(x - y)) = \varphi(t_2) \leq \varphi(t_1) = f(y + t_1(x - y)) = f(\bar{y})$  y además,

$$\begin{aligned} \frac{1}{t_2 - t_1}(\bar{x} - \bar{y})^t &= \frac{1}{t_2 - t_1} (y + t_2(x - y) - (y + t_1(x - y)))^t, \\ &= \frac{1}{t_2 - t_1} ((t_2 - t_1)(x - y))^t, \\ &= (x - y)^t, \end{aligned}$$

luego,

$$\begin{aligned}\varphi'(t_1) &= (x - y)^t \nabla f(\bar{y}), \\ &= \frac{1}{t_2 - t_1} (\bar{x} - \bar{y})^t \nabla f(\bar{y}).\end{aligned}$$

En consecuencia,

$$\varphi'(t_1)(t_2 - t_1) = (\bar{x} - \bar{y})^t \nabla f(\bar{y}) > 0,$$

lo cual es una contradicción (ya que contradice a (1.2)).  $\square$

**Lema 1.1.** [12] *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es casi-convexa y diferenciable. Si  $\nabla f(x) \neq 0$  y  $f(y) \leq f(x)$ , entonces  $\langle \nabla f(x), y - x \rangle < 0$ .*

*Demostración.* Como  $f$  es continua y casi-convexa, entonces existe  $\delta > 0$ , tal que  $z \in \mathbf{B}(y, \delta)$  implica  $f(z) \leq f(x)$ , donde  $\mathbf{B}(y, \delta) := \{w \in \mathbb{R}^n : \|y - w\| < \delta\}$ .

Tomando

$$\hat{z} = y + \frac{\delta}{2} \frac{\nabla f(x)}{\|\nabla f(x)\|}.$$

Luego,

$$\|y - \hat{z}\| = \left\| y - y - \frac{\delta}{2} \frac{\nabla f(x)}{\|\nabla f(x)\|} \right\| = \frac{\delta}{2} \frac{\|\nabla f(x)\|}{\|\nabla f(x)\|} = \frac{\delta}{2} < \delta.$$

Así,  $\hat{z} \in \mathbf{B}(y, \delta)$  lo cual implica que  $f(\hat{z}) \leq f(x)$ , por la Proposición 1.3 se tiene que:

$$\begin{aligned}\langle \nabla f(x), \hat{z} - x \rangle \leq 0 &\Rightarrow \langle \nabla f(x), y + \frac{\delta}{2} \frac{\nabla f(x)}{\|\nabla f(x)\|} - x \rangle \leq 0, \\ &\Rightarrow \langle \nabla f(x), y - x \rangle + \frac{\delta}{2} \frac{\|\nabla f(x)\|^2}{\|\nabla f(x)\|} \leq 0, \\ &\Rightarrow \langle \nabla f(x), y - x \rangle + \frac{\delta}{2} \|\nabla f(x)\| \leq 0, \\ &\Rightarrow \langle \nabla f(x), y - x \rangle \leq -\frac{\delta}{2} \|\nabla f(x)\| < 0.\end{aligned}$$

$\square$

Para finalizar este capítulo, presentamos los siguientes resultados que son usados para probar que las funciones definidas en el capítulo siguiente, sean estrictamente convexas o fuertemente convexas.

**Proposición 1.4.** [2] Sea  $C \subset \mathbb{R}^n$  un conjunto convexo, sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función dos veces continuamente diferenciable en  $C$  y  $\alpha$  un escalar positivo, se tiene que

- (i) Si  $\nabla^2 f(x)$  es definida positiva para cualquier  $x \in C$ , entonces  $f$  es estrictamente convexa en  $C$ .
- (ii)  $f$  es fuertemente convexa si y sólo si  $\nabla^2 f(x) - \alpha I$  es semidefinido positivo,  $\forall x \in \mathbb{R}^n$ , donde  $I$  es la matriz identidad.

*Demostración.* (i) Por la Proposición A.23 (b) del apéndice A de [2], para cualquier  $x, y \in C$  tenemos

$$f(y) = f(x) + (y - x)^t \nabla f(x) + \frac{1}{2} (y - x)^t \nabla^2 f(x + \alpha(y - x))(y - x),$$

para algún  $\alpha \in [0, 1]$ . Luego usando la hipótesis que  $\nabla^2 f(x)$  es definida positiva para cualquier  $x \in C$ , tenemos que,

$$f(y) > f(x) + (y - x)^t \nabla f(x), \quad \forall x, y \in C.$$

Por lo tanto  $f$  es estrictamente convexa en  $C$ , (Por Proposición B.3 (a) de [2]).

- (ii) Ver prueba de la Proposición B.4 de [2].

□



# CAPÍTULO 2

## FUNCIÓN CASI-DISTANCIA $d_\phi$ Y SUS PROPIEDADES.

En este capítulo presentaremos la definición del núcleo  $\phi$  y estudiaremos las propiedades de la función de dos variables  $d_\phi$  inducida por  $\phi$  usando la fórmula:

$$d_\phi(x, y) := \begin{cases} \sum_{i=1}^n (y_i)^p \phi\left(\frac{x_i}{y_i}\right), & \text{con } p \geq 2, \quad x, y \in \mathbb{R}_{++}^n \\ +\infty, & \text{en otro caso.} \end{cases}$$

Ahora presentamos la definición de las clases de funciones  $\Phi$  y algunas propiedades que son usadas en el contexto de optimización. Esto es debido a que nuestra función casi-distancia depende de una función  $\phi \in \Phi$ .

**Definición 2.1.** [12]

Sea  $\varphi : \mathbb{R} \rightarrow (-\infty, +\infty]$  una función propia convexa cerrada y  $\text{dom}\varphi \subseteq [0, +\infty)$ . Asumimos que

- (i)  $\varphi$  es dos veces continuamente diferenciable en  $\text{int}(\text{dom}\varphi) = (0, +\infty)$ ,
- (ii)  $\varphi$  es estrictamente convexa en su dominio,
- (iii)  $\lim_{t \rightarrow 0^+} \varphi'(t) = -\infty$ ,
- (iv)  $\varphi(1) = \varphi'(1) = 0$  y  $\varphi''(1) > 0$ .

Denotaremos por  $\Phi$  las clases de funciones que satisfacen (i) – (iv). Dado  $\varphi \in \Phi$ , definimos las siguientes subclases de  $\Phi$ :

$$\Phi_1 = \{\varphi \in \Phi : \varphi''(1)(1 - 1/t) \leq \varphi'(t) \leq \varphi''(1) \ln(t), \quad \forall t > 0\} \quad (2.1)$$

$$\Phi_2 = \{\varphi \in \Phi : \varphi''(1)(1 - 1/t) \leq \varphi'(t) \leq \varphi''(1)(t - 1), \quad \forall t > 0\} \quad (2.2)$$

Como  $\ln(t) \leq (t - 1)$ ,  $\forall t > 0$  y  $\varphi''(1) > 0$ , claramente,  $\Phi_1 \subseteq \Phi_2 \subseteq \Phi$

**Ejemplo 2.1.** [9] Bajo suposición que  $0 \ln(0) = 0$  definimos las siguientes funciones propias, convexas y cerradas:

$$\varphi_1(t) = t \ln(t) - t + 1, \quad \text{dom}\varphi_1 = [0, +\infty), \quad (2.3)$$

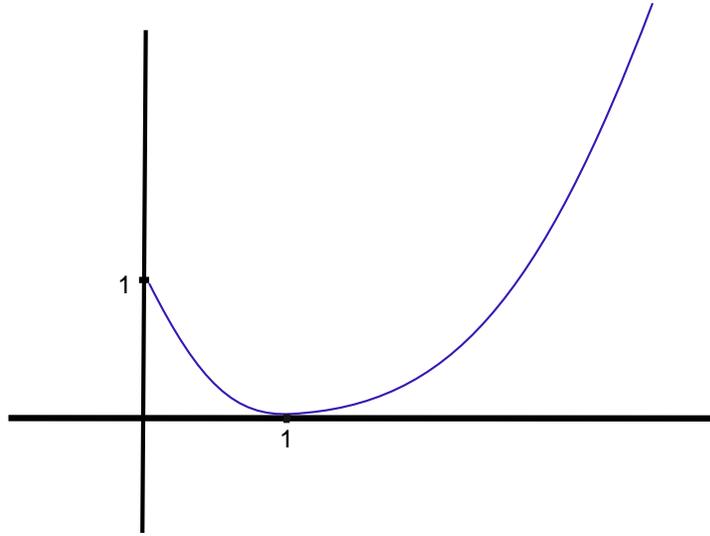
$$\varphi_2(t) = -\ln(t) + t - 1, \quad \text{dom}\varphi_2 = (0, +\infty), \quad (2.4)$$

Ahora demostraremos que  $\varphi_1, \varphi_2 \in \Phi_1$

$\varphi_1'(t) = \ln(t) + 1 - 1 = \ln(t)$  y  $\varphi_1''(t) = 1/t$ . Luego,  $\varphi_1$  es dos veces continuamente diferenciable en  $\text{int}(\text{dom}\varphi_1) = (0, +\infty)$ , además es estrictamente convexa en su dominio,  $\lim_{t \rightarrow 0^+} \varphi_1'(t) = -\infty$  y  $\varphi_1(1) = \varphi_1'(1) = 0$  y  $\varphi_1''(1) = 1 > 0$ . Así,  $\varphi_1 \in \Phi$  y

$$\varphi_1''(1)(1 - 1/t) = (1 - 1/t) \leq \varphi_1'(t) = \ln(t) = \varphi_1''(1) \ln(t), \quad \forall t > 0$$

Luego,  $\varphi_1 \in \Phi_1 \subseteq \Phi_2$ .



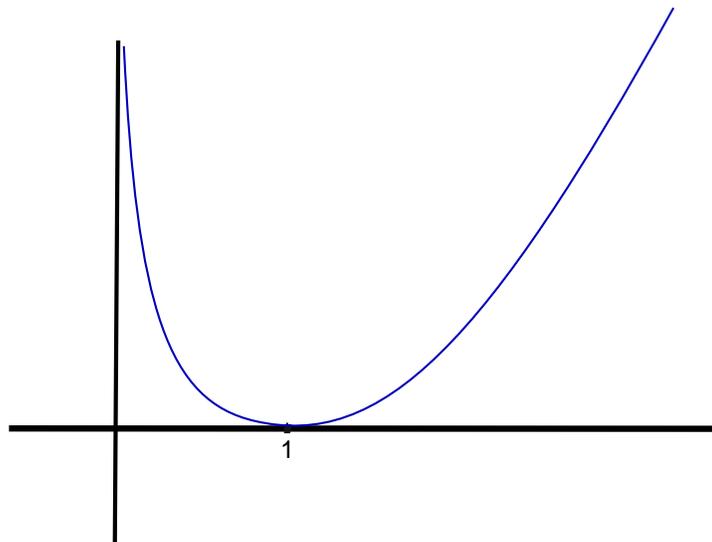
**Figura 2.1:** Gráfica de la función  $\varphi_1$ .

Por otro lado

$\varphi_2'(t) = -1/t + 1$  y  $\varphi_2''(t) = 1/t^2$ . Luego,  $\varphi_2$  es dos veces continuamente diferenciable en  $\text{int}(\text{dom}\varphi_2) = (0, +\infty)$ , además es estrictamente convexa en su dominio,  $\lim_{t \rightarrow 0^+} \varphi_2'(t) = -\infty$  y  $\varphi_2(1) = \varphi_2'(1) = 0$  y  $\varphi_2''(1) = 1 > 0$ . Así,  $\varphi_2 \in \Phi$  y

$$\varphi_2''(1)(1 - 1/t) = (1 - 1/t) = \varphi_2'(t) \leq \ln(t) = \varphi_1''(1) \ln(t), \quad \forall t > 0$$

Por lo tanto,  $\varphi_2 \in \Phi_1 \subseteq \Phi_2$ .



**Figura 2.2:** Gráfica de la función  $\varphi_2$ .

A continuación presentamos la definición del núcleo  $\phi$ , definida como la suma de una función  $\varphi \in \Phi$  y una función cuadrática. Luego en las propiedades de  $\phi$ , tenemos que  $\phi \in \Phi$  y además bajo algunas condiciones es fuertemente convexa en su dominio.

**Definición 2.2.** [12] Dado  $\varphi \in \Phi$  y los parámetros  $\mu > 0$  y  $\nu \geq 0$ . Se define como núcleo, a la función convexa propia cerrada dada por:

$$\phi(t) := \mu\varphi(t) + \frac{\nu}{2}(t-1)^2. \quad (2.5)$$

### Observación 1. (Propiedades de $\phi$ )

(a)  $\phi \in \Phi$ .

En efecto, veamos que se cumple las condiciones (i) – (iv) de la definición 2.1

(i)  $\phi$  es dos veces continuamente diferenciable en  $\text{int}(\text{dom}\phi) = (0, +\infty)$ , ya que  $\varphi \in \Phi$  y  $g(t) = (t-1)^2$  son dos veces continuamente diferenciable en  $(0, +\infty)$ .

(ii)  $\phi$  es estrictamente convexa en su dominio, ya que  $\varphi$  y  $g$  son estrictamente convexas en el  $\text{dom}\phi$ .

(iii) Como  $\phi'(t) = \mu\varphi'(t) + \nu(t - 1)$ , entonces,

$$\begin{aligned} \lim_{t \rightarrow 0^+} \phi'(t) &= \lim_{t \rightarrow 0^+} (\mu\varphi'(t) + \nu(t - 1)) \\ &= -\infty, \quad (\text{ya que } \mu > 0 \text{ y } \lim_{t \rightarrow 0^+} \varphi'(t) = -\infty). \end{aligned}$$

(iv)  $\phi(1) = \phi'(1) = 0$  y  $\phi''(1) > 0$ ,

En efecto

Como  $\varphi(1) = 0 = \varphi'(1)$ , entonces:

$$\phi(1) = \mu\varphi(1) + \frac{\nu}{2}(1 - 1)^2 = 0 = \mu\varphi'(1) + \nu(1 - 1) = \phi'(1),$$

y además,  $\phi''(1) = \mu\varphi''(1) + \nu > \nu > 0$ , (ya que  $\mu\varphi''(1) > 0$ ).

(b)  $\phi$  es fuertemente convexa en su dominio, si  $\nu > 0$ .

En efecto,

Sabemos que  $\varphi''(t) > 0$ ,  $\forall t \in \text{dom}\varphi = \text{dom}\phi$ , entonces,  $\phi''(t) = \mu\varphi''(t) + \nu > \nu > 0$ ,  $\forall t \in \text{dom}\phi$ . Así,  $\phi$  es fuertemente convexa en su dominio (por la proposición 1.4).

Los siguientes lemas están demostrados para  $p = 2$  en [9] y [12], y nosotros lo generalizamos para  $p \geq 2$ .

En el siguiente resultado mostraremos las propiedades de la función casi-distancia  $d_\phi$ , las cuales serán empleadas en el capítulo siguiente para el análisis de convergencia.

**Lema 2.1.** *Dado  $\varphi \in \Phi$  y los parámetros  $\mu > 0$  y  $\nu \geq 0$ , y sea  $\phi$  el núcleo definido como en (2.5) y*

$$d_\phi(x, y) := \begin{cases} \sum_{i=1}^n (y_i)^p \phi\left(\frac{x_i}{y_i}\right), & \text{con } p \geq 2, \quad x, y \in \mathbb{R}_{++}^n \\ +\infty, & \text{en otro caso.} \end{cases} \quad (2.6)$$

Entonces,

(a)  $d_\phi$  es una función homogénea de orden  $p$  en  $\mathbb{R}_{++}^n$ .

(b) Para cualquier  $y \in \mathbb{R}_{++}^n$  fijo, la función  $d_\phi(\cdot, y)$  es estrictamente convexa en  $\mathbb{R}_{++}^n$ . Si, adicionalmente,  $\nu > 0$ , entonces  $d_\phi(\cdot, y)$  es fuertemente convexa en  $\mathbb{R}_{++}^n$ .

- (c) Para cualquier  $(x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ ,  $d_\phi(x, y) \geq 0$ , y  $d_\phi(x, y) = 0$  si y sólo si  $x = y$ .
- (d) Si  $\varphi \in \Phi_1$  o  $\Phi_2$  y  $\{y^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}_{++}^n$  una sucesión que converge a  $\bar{y} \in \mathbb{R}_+^n$ , entonces para cualquier  $x \in \mathbb{R}_{++}^n$ , la sucesión  $\{d_\phi(x, y^k)\}_{k \in \mathbb{N}}$  es acotada.

*Demostración.* (a) Sean  $x, y \in \mathbb{R}_{++}^n$  y Para  $\alpha > 0$ , se tiene:

$$\begin{aligned}
 d_\phi(\alpha x, \alpha y) &= \sum_{i=1}^n (\alpha y_i)^p \phi\left(\frac{\alpha x_i}{\alpha y_i}\right), \\
 &= \sum_{i=1}^n \alpha^p y_i^p \phi\left(\frac{x_i}{y_i}\right), \\
 &= \alpha^p \sum_{i=1}^n y_i^p \phi\left(\frac{x_i}{y_i}\right), \\
 &= \alpha^p d_\phi(x, y).
 \end{aligned}$$

- (b) Sean  $x, z \in \mathbb{R}_{++}^n$  arbitrarios,  $y \in \mathbb{R}_{++}^n$  fijo y para  $\forall \alpha \in (0, 1)$  y como  $\phi$  es estrictamente convexa entonces, se tiene que:

$$\begin{aligned}
 d_\phi(\alpha x + (1 - \alpha)z, y) &= \sum_{i=1}^n (y_i)^p \phi\left(\frac{\alpha x_i + (1 - \alpha)z_i}{y_i}\right), \\
 &< \sum_{i=1}^n \left[ \alpha y_i^p \phi\left(\frac{x_i}{y_i}\right) + (1 - \alpha) y_i^p \phi\left(\frac{z_i}{y_i}\right) \right], \\
 &= \alpha \sum_{i=1}^n y_i^p \phi\left(\frac{x_i}{y_i}\right) + (1 - \alpha) \sum_{i=1}^n y_i^p \phi\left(\frac{z_i}{y_i}\right), \\
 &= \alpha d_\phi(x, y) + (1 - \alpha) d_\phi(z, y).
 \end{aligned}$$

Ahora veamos que si  $\nu > 0$  entonces  $d_\phi(\cdot, y)$  es fuertemente convexa, lo cual es lo mismo probar que  $h_\xi(t) := \xi^\nu \phi(t/\xi)$ , para  $\xi > 0$  es fuertemente convexa en  $\mathbb{R}_{++}$ .

En efecto,

$$\begin{aligned}
 h'_\xi(t) &= \xi^\nu \phi'\left(\frac{t}{\xi}\right) \frac{1}{\xi} = \xi^{\nu-1} \phi'\left(\frac{t}{\xi}\right), \\
 & \quad y \\
 h''_\xi(t) &= \xi^{\nu-2} \phi''\left(\frac{t}{\xi}\right) = \xi^{\nu-2} \left( \mu \phi''\left(\frac{t}{\xi}\right) \right) + \xi^{\nu-2} \nu > \xi^{\nu-2} \nu > 0.
 \end{aligned}$$

Como  $h_\xi''(t) > \xi^{p-2}\nu > 0$ ,  $\forall t \in \mathbb{R}_{++}$ , entonces  $h_\xi$  es fuertemente convexa en  $\mathbb{R}_{++}$  (por la proposición 1.4).

(c) Como  $\phi$  es estrictamente convexa en su dominio y  $\phi'(1) = 0$  (ya que  $\phi \in \Phi$ ) entonces  $\phi$  alcanza un mínimo global en  $t = 1$ , entonces:

$$\phi(t) \geq \phi(1) = 0, \quad \forall t \in \text{dom } \phi.$$

Así,  $d_\phi(x, y) \geq 0$ ,  $\forall (x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ .

Por otro lado,

$$\begin{aligned} d_\phi(x, y) = 0 &\Leftrightarrow \sum_{i=1}^n y_i^p \phi\left(\frac{x_i}{y_i}\right) = 0, \quad \forall i \\ &\Leftrightarrow y_i^p \phi\left(\frac{x_i}{y_i}\right) = 0, \quad \forall i \\ &\Leftrightarrow \phi\left(\frac{x_i}{y_i}\right) = 0, \quad \forall i \\ &\Leftrightarrow \frac{x_i}{y_i} = 1, \quad (\text{ya que } \phi(t) = 0 \Leftrightarrow t = 1) \\ &\Leftrightarrow x_i = y_i, \quad \forall i \\ &\Leftrightarrow x = y. \end{aligned}$$

(d) Por definición de  $\phi$  y  $d_\phi$  tenemos que:

$$\begin{aligned} d_\phi(x, y^k) &= \sum_{i=1}^n (y_i^k)^p \phi\left(\frac{x_i}{y_i^k}\right), \\ &= \sum_{i=1}^n \left[ \mu(y_i^k)^p \varphi\left(\frac{x_i}{y_i^k}\right) + \frac{\nu}{2} (y_i^k)^p \left(\frac{x_i}{y_i^k} - 1\right)^2 \right], \\ &= \sum_{i=1}^n \left[ \mu(y_i^k)^p \varphi\left(\frac{x_i}{y_i^k}\right) + \frac{\nu}{2} (y_i^k)^{p-2} (x_i - y_i^k)^2 \right]. \end{aligned}$$

Así,

$$d_\phi(x, y^k) = \sum_{i=1}^n \left[ \mu(y_i^k)^p \varphi\left(\frac{x_i}{y_i^k}\right) + \frac{\nu}{2} (y_i^k)^{p-2} (x_i - y_i^k)^2 \right]. \quad (2.7)$$

Si  $\varphi(t)$  es acotado para cualquier  $t > 0$  entonces la conclusión es obvia. Discutimos los siguientes casos:

Caso (i) :  $\bar{y}_i > 0$  para cada  $i \in \{1, \dots, n\}$ . Como  $\{y_i^k\}_{k \in \mathbb{N}} \rightarrow \bar{y}_i$  para cada  $i$ ,

$$\begin{aligned} \lim_{k \rightarrow +\infty} d_\phi(x, y^k) &= \lim_{k \rightarrow +\infty} \sum_{i=1}^n \left[ \mu(y_i^k)^p \varphi\left(\frac{x_i}{y_i^k}\right) + \frac{\nu}{2} (y_i^k)^{p-2} (x_i - y_i^k)^2 \right], \\ &= \sum_{i=1}^n \left[ \mu(\bar{y}_i)^p \varphi\left(\frac{x_i}{\bar{y}_i}\right) + \frac{\nu}{2} (\bar{y}_i)^{p-2} (x_i - \bar{y}_i)^2 \right]. \end{aligned}$$

Luego,  $\{d_\phi(x, y^k)\}_{k \in \mathbb{N}}$  es convergente, y por tanto acotada.

Caso (ii) : Existe un índice  $i_0 \in \{1, \dots, n\}$ , tal que,  $\bar{y}_{i_0} = 0$ . Dadas las suposiciones y el Caso (i) es suficiente probar que la sucesión  $\{(y_{i_0}^k)^p \varphi(x_i/y_{i_0}^k)\}$  es acotada. Para cualquier  $k \in \mathbb{N}$ , usando la convexidad de  $\varphi$  y el hecho de que  $\varphi(1) = 0$ , tenemos que:

$$\begin{aligned} \varphi\left(\frac{x_i}{y_{i_0}^k}\right) - \varphi(1) &\leq \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) \left(\frac{x_i}{y_{i_0}^k} - 1\right), \\ \Rightarrow \varphi\left(\frac{x_i}{y_{i_0}^k}\right) &\leq \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) \left(\frac{x_i}{y_{i_0}^k} - 1\right), \\ \Rightarrow (y_{i_0}^k)^p \varphi\left(\frac{x_i}{y_{i_0}^k}\right) &\leq (y_{i_0}^k)^p \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) \left(\frac{x_i}{y_{i_0}^k} - 1\right), \\ \Rightarrow (y_{i_0}^k)^p \varphi\left(\frac{x_i}{y_{i_0}^k}\right) &\leq (y_{i_0}^k)^{p-1} \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) (x_i - y_{i_0}^k), \\ \Rightarrow (y_{i_0}^k)^p \varphi\left(\frac{x_i}{y_{i_0}^k}\right) &\leq \left| (y_{i_0}^k)^{p-1} \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) (x_i - y_{i_0}^k) \right|, \end{aligned}$$

Luego,

$$(y_{i_0}^k)^{p-1} \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) (x_i - y_{i_0}^k) \geq (y_{i_0}^k)^p \varphi\left(\frac{x_i}{y_{i_0}^k}\right), \quad (2.8)$$

ó

$$(y_{i_0}^k)^{p-1} \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) (x_i - y_{i_0}^k) \leq -(y_{i_0}^k)^p \varphi\left(\frac{x_i}{y_{i_0}^k}\right). \quad (2.9)$$

Si  $\varphi \in \Phi_2$  entonces

$$\varphi''(1)(1 - 1/t) \leq \varphi'(t) \leq \varphi''(1)(t - 1), \quad \forall t > 0.$$

Sustituyendo  $t = x_i/y_{i_0}^k$  en la inecuación anterior

$$\begin{aligned} \varphi''(1) \left(1 - \frac{y_{i_0}^k}{x_i}\right) &\leq \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) \leq \varphi''(1) \left(\frac{x_i}{y_{i_0}^k} - 1\right), \\ \Rightarrow \varphi''(1) \frac{y_{i_0}^k}{x_i} (x_i - y_{i_0}^k) &\leq y_{i_0}^k \varphi'\left(\frac{x_i}{y_{i_0}^k}\right) \leq \varphi''(1) (x_i - y_{i_0}^k). \end{aligned} \quad (2.10)$$

De la desigualdad derecha de (2.10), suponiendo que  $x_i - y_{i_0}^k \geq 0$  y usando (2.8),

$$\begin{aligned} \varphi''(1)(x_i - y_{i_0}^k) &\geq y_{i_0}^k \varphi' \left( \frac{x_i}{y_{i_0}^k} \right), \\ (y_{i_0}^k)^{p-2} \varphi''(1)(x_i - y_{i_0}^k)^2 &\geq (y_{i_0}^k)^{p-1} \varphi' \left( \frac{x_i}{y_{i_0}^k} \right) (x_i - y_{i_0}^k), \\ &\geq (y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right). \end{aligned}$$

Así,

$$(y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right) \leq (y_{i_0}^k)^{p-2} \varphi''(1)(x_i - y_{i_0}^k)^2. \quad (2.11)$$

Ahora, de la desigualdad derecha de (2.10), suponiendo que  $x_i - y_{i_0}^k < 0$  y usando (2.9),

$$\begin{aligned} \varphi''(1)(x_i - y_{i_0}^k) &\geq y_{i_0}^k \varphi' \left( \frac{x_i}{y_{i_0}^k} \right), \\ 0 < (y_{i_0}^k)^{p-2} \varphi''(1)(x_i - y_{i_0}^k)^2 &\leq (y_{i_0}^k)^{p-1} \varphi' \left( \frac{x_i}{y_{i_0}^k} \right) (x_i - y_{i_0}^k), \\ &\leq -(y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right), \\ &< 0, \end{aligned}$$

lo cual es una contradicción.

De la desigualdad izquierda de (2.10), suponiendo que  $x_i - y_{i_0}^k > 0$  y usando (2.9),

$$\begin{aligned} \varphi''(1) \frac{y_{i_0}^k}{x_i} (x_i - y_{i_0}^k) &\leq y_{i_0}^k \varphi' \left( \frac{x_i}{y_{i_0}^k} \right), \\ 0 < (y_{i_0}^k)^{p-2} \varphi''(1) \frac{y_{i_0}^k}{x_i} (x_i - y_{i_0}^k)^2 &\leq (y_{i_0}^k)^{p-1} \varphi' \left( \frac{x_i}{y_{i_0}^k} \right) (x_i - y_{i_0}^k), \\ &\leq -(y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right), \\ &< 0, \end{aligned}$$

lo cual es una contradicción.

Ahora de la desigualdad izquierda de (2.10), suponiendo que  $x_i - y_{i_0}^k \leq 0$  y usando (2.8),

$$\begin{aligned}\varphi''(1) \frac{y_{i_0}^k}{x_i} (x_i - y_{i_0}^k) &\leq y_{i_0}^k \varphi' \left( \frac{x_i}{y_{i_0}^k} \right), \\ (y_{i_0}^k)^{p-2} \varphi''(1) \frac{y_{i_0}^k}{x_i} (x_i - y_{i_0}^k)^2 &\geq (y_{i_0}^k)^{p-1} \varphi' \left( \frac{x_i}{y_{i_0}^k} \right) (x_i - y_{i_0}^k), \\ \varphi''(1) \frac{(y_{i_0}^k)^{p-1}}{x_i} (x_i - y_{i_0}^k)^2 &\geq (y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right),\end{aligned}$$

Así,

$$(y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right) \leq \varphi''(1) \frac{(y_{i_0}^k)^{p-1}}{x_i} (x_i - y_{i_0}^k)^2. \quad (2.12)$$

Combinando las inecuaciones (2.11) y (2.12) se tiene:

$$(y_{i_0}^k)^p \varphi \left( \frac{x_i}{y_{i_0}^k} \right) \leq \max \left\{ \varphi''(1) (y_{i_0}^k)^{p-2} (x_i - y_{i_0}^k)^2, \varphi''(1) \frac{(y_{i_0}^k)^{p-1}}{x_i} (x_i - y_{i_0}^k)^2 \right\}.$$

Por lo tanto,  $\{(y_{i_0}^k)^p \varphi(x_i/y_{i_0}^k)\}$  es acotada para cualquier  $\varphi \in \Phi_2$  y en consecuencia la sucesión  $\{d_\phi(x, y^k)\}_{k \in \mathbb{N}}$  es acotada. Como  $\Phi_1 \subseteq \Phi_2$ , la sucesión  $\{d_\phi(x, y^k)\}_{k \in \mathbb{N}}$  también es acotada para  $\varphi \in \Phi_1$ . □

**Observación 2.** Por el Lema 2.1 (a) – (c) se tiene que  $d_\phi$  definido por (2.6) es una función casi-distancia convexa homogénea de orden  $p$ , con  $p \geq 2$ .

Finalmente, presentaremos una propiedad favorable de  $d_\phi$ , el cual juega un papel importante en el análisis de convergencia en el proximo capítulo.

**Lema 2.2.** Para cualquier  $a, b \in \mathbb{R}_{++}^n$  y  $c \in \mathbb{R}_+^n$ , con  $\varphi \in \Phi_2$  y  $\phi$  definido como en (2.5), entonces,

$$\langle c - b, \nabla_1 d_\phi(b, a) \rangle \leq \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [\theta(\|c - a\|^2 - \|c - b\|^2) - \tau \|b - a\|^2],$$

con  $\theta := (\nu + \mu\varphi''(1))/2$  y  $\tau := (\nu - \mu\varphi''(1))/2$ .

*Demostración.* Como  $\varphi \in \Phi_2$ , entonces de (2.2) se tiene que:

$$\varphi'(t) \leq \varphi''(1)(t - 1), \forall t > 0.$$

Sustituyendo  $t = b_j/a_j$  en la inecuación anterior y multiplicando por  $c_j(a_j)^{p-1}$

$$c_j(a_j)^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) \leq c_j(a_j)^{p-2}\varphi''(1)(b_j - a_j), j = 1, \dots, n. \quad (2.13)$$

Además también por (2.2) se tiene que:

$$-\varphi'(t) \leq -\varphi''(1)(1 - 1/t), \forall t > 0.$$

Sustituyendo  $t = b_j/a_j$  en la inecuación anterior y multiplicando por  $(a_j)^{p-1}$  obtenemos que:

$$\begin{aligned} -(a_j)^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) &\leq (a_j)^{p-1}\varphi''(1) \left( \frac{a_j}{b_j} - 1 \right), j = 1, \dots, n \\ -(a_j)^{p-1}b_j\varphi' \left( \frac{b_j}{a_j} \right) &\leq (a_j)^{p-1}\varphi''(1) (a_j - b_j), j = 1, \dots, n. \end{aligned} \quad (2.14)$$

Por otro lado definimos:

$$\Psi'(a, b) := \left( (a_1)^{p-1}\varphi' \left( \frac{b_1}{a_1} \right), \dots, (a_n)^{p-1}\varphi' \left( \frac{b_n}{a_n} \right) \right)^t, \forall a, b \in \mathbb{R}_{++}^n. \quad (2.15)$$

Luego, usando (2.7), con  $y^k = y$ , tenemos,

$$d_\phi(x, y) = \sum_{i=1}^n \left[ \mu(y_i)^p \varphi \left( \frac{x_i}{y_i} \right) + \frac{\nu}{2} (y_i)^{p-2} (x_i - y_i)^2 \right].$$

En consecuencia,

$$\begin{aligned} \nabla_1 d_\phi(x, y) &= \nabla_1 \left[ \sum_{i=1}^n \left[ \mu(y_i)^p \varphi \left( \frac{x_i}{y_i} \right) + \frac{\nu}{2} (y_i)^{p-2} (x_i - y_i)^2 \right] \right], \\ &= \nabla_1 \left[ \mu(y_1)^p \varphi \left( \frac{x_1}{y_1} \right) + \frac{\nu}{2} (y_1)^{p-2} (x_1 - y_1)^2 + \dots + \mu(y_n)^p \varphi \left( \frac{x_n}{y_n} \right) + \frac{\nu}{2} (y_n)^{p-2} (x_n - y_n)^2 \right], \\ &= \left( \mu(y_1)^{p-1} \varphi' \left( \frac{x_1}{y_1} \right) + \nu(y_1)^{p-2} (x_1 - y_1), \dots, \mu(y_n)^{p-1} \varphi' \left( \frac{x_n}{y_n} \right) + \nu(y_n)^{p-2} (x_n - y_n) \right)^t, \\ &= \mu \left( (y_1)^{p-1} \varphi' \left( \frac{x_1}{y_1} \right), \dots, (y_n)^{p-1} \varphi' \left( \frac{x_n}{y_n} \right) \right)^t + \nu \left( (y_1)^{p-2} (x_1 - y_1), \dots, (y_n)^{p-2} (x_n - y_n) \right)^t. \end{aligned}$$

Usando la igualdad anterior, con  $x = b$ ,  $y = a$  y usando también (2.15), se tiene lo siguiente:

$$\nabla_1 d_\phi(b, a) = \mu \Psi'(a, b) + \nu \left( (a_1)^{p-2} (b_1 - a_1), \dots, (a_n)^{p-2} (b_n - a_n) \right)^t. \quad (2.16)$$

Ahora sumando las dos inecuaciones (2.13) y (2.14), con  $j = 1, \dots, n$  se obtiene,

$$\begin{aligned}
c_j(a_j)^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) - b_j a_j^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) &\leq c_j(a_j)^{p-2}\varphi''(1)(b_j - a_j) \\
&\quad + (a_j)^{p-1}\varphi''(1)(a_j - b_j), \\
\mu(c_j - b_j)(a_j)^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) + \nu(a_j)^{p-2}(c_j - b_j)(b_j - a_j) &\leq \mu\varphi''(1)[(a_j)^{p-2}(c_j(b_j - a_j) \\
&\quad + a_j(a_j - b_j))] + \nu(a_j)^{p-2}(c_j - b_j)(b_j - a_j) \\
&= (a_j)^{p-2}[\mu\varphi''(1)(c_j - a_j)(b_j - a_j)) \\
&\quad + \nu(c_j - b_j)(b_j - a_j)].
\end{aligned}$$

Así,

$$\begin{aligned}
(c_j - b_j) \left[ \mu(a_j)^{p-1}\varphi' \left( \frac{b_j}{a_j} \right) + \nu(a_j)^{p-2}(b_j - a_j) \right] &\leq a_j^{p-2}[\mu\varphi''(1)(c_j - a_j)(b_j - a_j)) \\
&\quad + \nu(c_j - b_j)(b_j - a_j)], \quad j = 1, \dots, n.
\end{aligned}$$

Luego, sumando sobre  $j = 1, \dots, n$  la última desigualdad y usando (2.16), obtenemos:

$$\begin{aligned}
\langle c - b, \nabla_1 d_\phi(b, a) \rangle &\leq \sum_{j=1}^n (a_j)^{p-2}[\mu\varphi''(1)(c_j - a_j)(b_j - a_j)) + \nu(c_j - b_j)(b_j - a_j)], \\
&\leq \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} \sum_{j=1}^n [\mu\varphi''(1)(c_j - a_j)(b_j - a_j)) + \nu(c_j - b_j)(b_j - a_j)], \\
&= \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [\mu\varphi''(1)\langle c - a, b - a \rangle + \nu\langle c - b, b - a \rangle].
\end{aligned}$$

Entonces,

$$\langle c - b, \nabla_1 d_\phi(b, a) \rangle \leq \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [\mu\varphi''(1)\langle c - a, b - a \rangle + \nu\langle c - b, b - a \rangle].$$

Usando las identidades,

$$\langle c - a, b - a \rangle = \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 + \|b - a\|^2)$$

y

$$\langle c - b, b - a \rangle = \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 - \|b - a\|^2),$$

en la inecuación anterior se tiene que,

$$\begin{aligned}
\langle c - b, \nabla_1 d_\phi(b, a) \rangle &\leq \frac{1}{2} \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [\mu \varphi''(1) (\|c - a\|^2 - \|c - b\|^2 + \|b - a\|^2) \\
&\quad + \nu (\|c - a\|^2 - \|c - b\|^2 - \|b - a\|^2)], \\
&= \frac{1}{2} \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [(\mu \varphi''(1) + \nu) (\|c - a\|^2 - \|c - b\|^2) \\
&\quad - (\nu - \mu \varphi''(1)) \|b - a\|^2], \\
&= \max_{1 \leq j \leq n} \{(a_j)^{p-2}\} [\theta (\|c - a\|^2 - \|c - b\|^2) - \tau \|b - a\|^2],
\end{aligned}$$

donde  $\theta = (\nu + \mu \varphi''(1))/2$  y  $\tau = (\nu - \mu \varphi''(1))/2$ .

□

# CAPÍTULO 3

## MÉTODO TIPO-PROXIMAL INTERIOR.

### 3.1. Algoritmo

En este capítulo, consideraremos un algoritmo tipo-proximal basado en la función homogénea de orden  $p$ , con  $p \geq 2$ ,  $d_\phi$  para el siguiente problema de optimización casi-convexa:

$$\begin{cases} \text{mín } f(x) \\ \text{s.a. } x \geq 0, \end{cases} \quad (3.1)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función propia y casi-convexa.

El algoritmo **RIPM** basado en una casi-distancia homogénea de orden  $p = 2$ , fue propuesto en [12] para problemas de minimización casi-convexa diferenciable con restricciones de no negatividad y también fue propuesto por [9] pero para problemas de minimización casi-convexa que no necesariamente son diferenciables, con restricciones de no negatividad.

#### Método Proximal Interior Regularizado (RIPM)

Sea  $\phi$  definida como en (2.5) con  $\nu \geq \mu\phi''(1) > 0$ ,  $\varphi \in \Phi_2$  y  $d_\phi$  definida como en (2.6). Generando la sucesión  $\{x^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}_{++}^n$  dada en el siguiente esquema iterativo,

$$\begin{cases} x^0 > 0 \\ x^{k+1} \in \text{argmín}_{x \geq 0} \{f(x) + \lambda_k d_\phi(x, x^k)\}, \end{cases} \quad (3.2)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función propia y casi-convexa; y  $\{\lambda_k\}$  una sucesión de números reales positivos que satisfacen  $0 < \lambda_k < \bar{\lambda}$ , para algún  $\bar{\lambda} > 0$ .

Para establecer la convergencia de **RIPM**, a lo largo de este capítulo, tomaremos las siguientes hipótesis para el problema de optimización casi-convexa (3.1):

(H<sub>1</sub>)  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es acotada inferiormente en  $\mathbb{R}_+^n$ , esto es existe  $\beta \in \mathbb{R}$ , tal que,  
 $\beta \leq f(x), \forall x \in \mathbb{R}_+^n$ .

(H<sub>2</sub>)  $f$  es continuamente diferenciable.

A continuación, se estudian las propiedades de la sucesión  $\{x^k\}_{k \in \mathbb{N}}$  generadas por **RIPM**. Para este fin, se define el siguiente conjunto,

$$U := \{x \in \mathbb{R}_+^n : f(x) \leq f(x^k), k = 0, 1, \dots\}. \quad (3.3)$$

Veamos que el conjunto  $U$  es convexo y cerrado (ya que para el análisis de convergencia del algoritmo **RIPM** necesitamos que este conjunto sea convexo y cerrado).

En efecto,

Si  $U = \emptyset$  entonces  $U$  es convexo en  $\mathbb{R}_+^n$ .

Ahora supongamos que  $U \neq \emptyset$ , como la función propia  $f$  es casi-convexa si y sólo si el conjunto de nivel  $S_f(\alpha) := \{x \in \text{dom}f \mid f(x) \leq \alpha\}$  es convexo para cualquier  $\alpha \in \mathbb{R}$ , (por Proposición 1.2). Tomando  $\alpha = f(x^k)$  se tiene que:  $U := \{x \in \mathbb{R}_+^n : f(x) \leq f(x^k)\}$  es convexo en  $\mathbb{R}_+^n$ .

Para probar que  $U$  es cerrado, basta con tomar una sucesión cualquiera  $\{y^m\}_{m \in \mathbb{N}}$  en  $U$ , que converge a  $y$  y demostrar que  $y \in U$ .

*Demostración.* Como  $f$  es continua y  $y^m \in U$  para cada  $m \in \mathbb{N}$ , se obtiene,

$$\begin{aligned} f(y^m) \leq f(x^k) &\Rightarrow \lim_{m \rightarrow +\infty} f(y^m) \leq \lim_{m \rightarrow +\infty} f(x^k), \\ &\Rightarrow f(y) \leq f(x^k), \\ &\Rightarrow y \in U. \end{aligned}$$

Por lo tanto,  $U$  es un conjunto cerrado y convexo en  $\mathbb{R}_+^n$ . □

En lo que sigue nos concentraremos en la convergencia del **RIPM**. El resultado siguiente en la parte (a) establecemos la buena definición del algoritmo **RIPM** y las partes (b) y (c) serán usados en el análisis de convergencia. Las demostraciones de los items (a), (b) y (c) son extraídos de [12], pero la parte (c) la modificamos para  $p \geq 2$ .

**Proposición 3.1.** *Suponiendo que  $H_1$  y  $H_2$  se cumplen y sea la sucesión  $\{x^k\}$  generada por el algoritmo **RIPM** entonces:*

- (a) La sucesión  $\{x^k\}_{k \in \mathbb{N}}$  está bien definida.
- (b)  $\{f(x^k)\}_{k \in \mathbb{N}}$  es una sucesión decreciente y convergente.
- (c) Si  $U \neq \emptyset$  entonces  $\{x^k\}_{k \in \mathbb{N}}$  es Fejér convergente en el conjunto  $U$ .

*Demostración.* (a) La prueba se procede por inducción. Para  $k = 0$ , se cumple, (ya que  $x^0 > 0$  por la inicialización del algoritmo **RIPM**). Asumimos que  $x^k$  está bien definida, como  $x^k > 0$  entonces por  $H_1$  existe  $\beta \in \mathbb{R}$ , tal que,

$$f(x) \geq \beta, \forall x \in \mathbb{R}_+^n \Leftrightarrow F_k(x) := f(x) + \lambda_k d_\phi(x, x^k) \geq \beta + \lambda_k d_\phi(x, x^k), \forall x \in \mathbb{R}_+^n.$$

Por la definición de  $d_\phi$  en (2.6) y  $(x^k)_i > 0$ , tenemos,

$$\lambda_k d_\phi(x, x^k) \rightarrow +\infty, \text{ cuando } x \rightarrow \partial \mathbb{R}_{++}^n. \quad (3.4)$$

De (3.4) y como  $f$  es acotada inferiormente y se tiene que:

$$F_k(x) \rightarrow +\infty, \text{ cuando } x \rightarrow \partial \mathbb{R}_{++}^n \quad (3.5)$$

Como  $F_k$  es continua, acotada inferiormente y cumple (3.5),  $F_k$  alcanza un mínimo en un punto  $w > 0$ . Por lo tanto, existe  $x^{k+1} \in \text{argmín}\{F_k(x)\}$  con  $x^{k+1} = w > 0$ , el cual no es único debido a la no convexidad de  $f$ .

- (b) De (3.2),  $x^{k+1}$  es una solución óptima global del problema:

$$\min_{x \geq 0} \{f(x) + \lambda_k d_\phi(x, x^k)\},$$

y en consecuencia, para cualquier  $x \in \mathbb{R}_+^n$ , se tiene lo siguiente:

$$f(x^{k+1}) + \lambda_k d_\phi(x^{k+1}, x^k) \leq f(x) + \lambda_k d_\phi(x, x^k). \quad (3.6)$$

Sustituyendo  $x = x^k$  en (3.6) y usando el lema 2.6 (c), entonces obtenemos que:

$$\begin{aligned} f(x^{k+1}) + \lambda_k d_\phi(x^{k+1}, x^k) &\leq f(x^k) + \lambda_k d_\phi(x^k, x^k) = f(x^k), \\ \Rightarrow 0 &\leq \lambda_k d_\phi(x^{k+1}, x^k) \leq f(x^k) - f(x^{k+1}), \\ \Rightarrow 0 &\leq f(x^k) - f(x^{k+1}), \forall k \in \mathbb{N}. \end{aligned}$$

Por lo tanto,  $\{f(x^k)\}_{k \in \mathbb{N}}$  es una sucesión decreciente, entonces ésta sucesión tiene cota superior. Por la suposición  $H_1$  garantizamos que  $f$  está acotada inferiormente. Luego,  $\{f(x^k)\}_{k \in \mathbb{N}}$  es monótona y acotada y por lo tanto convergente.

(c) Dado  $x \in U$  entonces  $f(x) \leq f(x^{k+1})$ ,  $\forall k \in \mathbb{N}$  donde  $\{x^k\}$  es generado por **RIPM**. Por la suposición  $H_2$  y la Proposición 1.3, obtenemos que:

$$\langle \nabla f(x^{k+1}), x - x^{k+1} \rangle \leq 0 \quad (3.7)$$

Por otro lado, si  $z \in \text{argmín}\{F_k(x)\}$ , con  $z > 0$ . Por lo tanto por las condiciones de optimalidad de este problema, resulta que  $\nabla F_k(z) = 0$  y así  $\nabla f(z) = -\lambda_k \nabla_1 d_\phi(z, x^k)$ ,  $\forall k \in \mathbb{N}$ . Particularmente,

$$\nabla f(x^{k+1}) = -\lambda_k \nabla_1 d_\phi(x^{k+1}, x^k), \forall k \in \mathbb{N} \quad (3.8)$$

Por (3.7), (3.8) y el Lema 2.2, con  $b = x^{k+1}$ ,  $a = x^k$  y  $c=x$ , tenemos

$$\begin{aligned} 0 \geq \langle \nabla f(x^{k+1}), x - x^{k+1} \rangle &= \langle -\lambda_k \nabla_1 d_\phi(x^{k+1}, x^k), x - x^{k+1} \rangle \\ &= -\lambda_k \langle \nabla_1 d_\phi(x^{k+1}, x^k), x - x^{k+1} \rangle \\ &\geq -\lambda_k \max_{1 \leq j \leq n} \{(x_j^k)^{p-2}\} [\theta(\|x - x^k\|^2 - \|x - x^{k+1}\|^2) \\ &\quad - \tau \|x^{k+1} - x^k\|^2], \\ &= \lambda_k \max_{1 \leq j \leq n} \{(x_j^k)^{p-2}\} [\theta(\|x - x^{k+1}\|^2 - \|x - x^k\|^2) \\ &\quad + \tau \|x^{k+1} - x^k\|^2], \\ &\geq \lambda_k \max_{1 \leq j \leq n} \{(x_j^k)^{p-2}\} \theta(\|x - x^{k+1}\|^2 - \|x - x^k\|^2), \end{aligned}$$

como  $\lambda_k > 0$ ,  $\theta > 0$ ,  $\tau \geq 0$  y  $\max_{1 \leq j \leq n} \{(x_j^k)^{p-2}\} > 0$ , para cada  $k \in \mathbb{N}$ , se tiene,

$$\|x - x^{k+1}\|^2 \leq \|x - x^k\|^2, \quad \forall k \in \mathbb{N}.$$

Así, se concluye que  $\{x^k\}_{k \in \mathbb{N}}$  es Fejér convergente en  $U$ . □

### 3.2. Análisis de Convergencia.

En esta sección demostraremos los resultados importantes de convergencia. En la primera, se establece la convergencia a un punto estacionario del problema (3.1), cuando  $U \neq \emptyset$ . Este teorema fue demostrado por [12] usando la casi-distancia con  $p = 2$ , y nosotros la generalizamos para  $p \geq 2$ .

**Teorema 3.1.** *Si las hipótesis  $H_1$  y  $H_2$  se cumplen y  $U \neq \emptyset$ , entonces la sucesión  $\{x^k\}$  generada por el algoritmo **RIPM** converge a un punto estacionario del problema (3.1).*

*Demostración.* Primero, demostraremos que  $\{x^k\}$  generada por el algoritmo **RIPM** es convergente. Sea  $\bar{x}$  un punto de clausura de  $\{x^k\}$  y  $\{x^{k_j}\}$  una subsucesión de  $\{x^k\}$  que converge a  $\bar{x}$ . De la continuidad de  $f$  y la Proposición 3.1 (b) tenemos que

$$f(\bar{x}) = \lim_{j \rightarrow +\infty} f(x^{k_j}) = \inf\{f(x^k)\}_{k \in \mathbb{N}} \leq f(x^k), \quad \forall k \in \mathbb{N},$$

por la definición de  $U$ , implica que  $\bar{x} \in U$ .

Luego usando la Proposición 3.1 (c) y la Proposición 1.1, tenemos que  $\{x^k\}$  es convergente.

Ahora mostraremos que  $\bar{x}$  es un punto estacionario del problema 3.1, esto es por Definición 1.4:

$$\nabla f(\bar{x})^t(x - \bar{x}) \geq 0,$$

lo que en nuestro caso sería lo mismo probar que:

$$\bar{x} \geq 0, \quad \nabla f(\bar{x}) \geq 0 \quad y \quad \bar{x}_i(\nabla f(\bar{x}))_i = 0, \quad \forall i = 1, \dots, n. \quad (3.9)$$

La primera condición en (3.9) es inmediata, ya que  $\lim_{k \rightarrow +\infty} x^k = \bar{x}$  y  $x^k > 0$ , para cada  $k \in \mathbb{N}$ . Para la prueba de las otras dos condiciones de (3.9) consideramos los siguientes conjuntos:

$$I(\bar{x}) := \{i \in \{1, \dots, n\} : \bar{x}_i = 0\} \quad y \quad J(\bar{x}) := \{i \in \{1, \dots, n\} : \bar{x}_i > 0\}.$$

Caso 1 Si  $i \in I(\bar{x})$ , supongamos por absurdo que  $(\nabla f(\bar{x}))_i < 0$ . Por ser  $f$  continuamente diferenciable tenemos que,  $(\nabla f(x^{k+1}))_i \rightarrow (\nabla f(\bar{x}))_i < 0$ . Por lo tanto  $(\nabla f(x^{k+1}))_i < 0$  para  $k$  suficientemente grande. Por la ecuación (3.8) y la definición de  $d_\phi$  en (2.6) tenemos

$$(\nabla f(x^{k+1}))_i = -\lambda_k(\nabla_1 d_\phi(x^{k+1}, x^k))_i = -\lambda_k(x_i^k)^{p-1} \phi' \left( \frac{x_i^{k+1}}{x_i^k} \right). \quad (3.10)$$

Luego,

$$0 > (\nabla f(x^{k+1}))_i = -\lambda_k(x_i^k)^{p-1} \phi' \left( \frac{x_i^{k+1}}{x_i^k} \right),$$

así,  $\phi' \left( \frac{x_i^{k+1}}{x_i^k} \right) > 0$  (ya que  $\lambda_k > 0$  y  $(x^k)^{p-1} > 0$ ) para  $k$  suficientemente grande. Lo cual implica que para  $k$  suficientemente grande se tiene,

$$\frac{x_i^{k+1}}{x_i^k} > 1 \Rightarrow x_i^{k+1} > x_i^k > 0,$$

con el resultado anterior hemos probado que  $\{x_i^k\}$  es estrictamente creciente y positiva. Luego, como esta sucesión es convergente, se tiene que  $\lim_{k \rightarrow +\infty} \{x_i^k\} = \sup\{x_i^k\} > 0$ , lo cual contradice el hecho que  $\{x_i^k\}$  converge a  $\bar{x}_i = 0$ . Por lo tanto  $(\nabla f(\bar{x}))_i \geq 0, \forall i \in I(\bar{x})$ .

Caso 2 Si  $i \in J(\bar{x})$ , entonces

$$\phi' \left( \frac{x_i^{k+1}}{x_i^k} \right) \rightarrow \phi'(1) = 0,$$

luego usando  $H_2$ , (3.10) y además como  $\{\lambda_k\}$  es acotada y  $\{x^k\}$  es convergente, tenemos,

$$(\nabla f(x^{k+1}))_i \rightarrow (\nabla f(\bar{x}))_i = 0, \quad \forall i \in J(\bar{x}).$$

De los Casos 1 y 2 se concluye que,

$$\bar{x} \geq 0, \quad \nabla f(\bar{x}) \geq 0 \quad \text{y} \quad \bar{x}_i (\nabla f(\bar{x}))_i = 0, \quad \forall i = 1, \dots, n.$$

Así,

$$\nabla f(\bar{x})^t (x - \bar{x}) \geq 0, \quad x \geq 0.$$

Por lo tanto,  $\bar{x}$  es un punto estacionario del problema (3.1). □

Con la hipótesis del teorema anterior, el nuevo resultado establece dos condiciones suficientes para que  $\bar{x}$  sea una solución del problema 3.1. De aquí en adelante  $\bar{x} = \lim_{k \rightarrow +\infty} x^k$ , donde  $\{x^k\}_{k \in \mathbb{N}}$  es una sucesión generada por **RIPM**.

**Teorema 3.2.** [12] Si las hipótesis  $H_1$  y  $H_2$  se cumplen,  $U \neq \emptyset$  y sea la sucesión  $\{x^k\}$  generada por el algoritmo **RIPM** que converge a  $\bar{x}$  entonces,

- (a) Si  $\nabla f(\bar{x}) \neq 0$  entonces  $\bar{x}$  es solución del problema (3.1).
- (b) Si  $\lambda_k \rightarrow 0$  entonces  $\bar{x}$  es solución del problema (3.1).

*Demostración.* (a) Supongamos por absurdo que  $\bar{x}$  no es un minimizador de  $f$ , entonces existe  $\hat{x} \geq 0$  tal que  $f(\hat{x}) < f(\bar{x})$  y con  $\nabla f(\bar{x}) \neq 0$  por Lema 1.1, tenemos que,

$$\langle \nabla f(\bar{x}), \hat{x} - \bar{x} \rangle < 0.$$

En consecuencia y usando el Teorema 3.1,

$$0 > \langle \nabla f(\bar{x}), \hat{x} - \bar{x} \rangle = \langle \nabla f(\bar{x}), \hat{x} \rangle + \langle \nabla f(\bar{x}), \bar{x} \rangle = \langle \nabla f(\bar{x}), \hat{x} \rangle.$$

Lo cual es una contradicción ya que  $\nabla f(\bar{x}) \geq 0$  y  $\hat{x} \geq 0$  (Por Teorema 3.1). Por lo tanto  $\bar{x}$  es solución del problema (3.1).

(b) Como  $x^{k+1}$  es el minimizador de  $f(x) + \lambda_k d_\phi(x, x^k)$ , tenemos:

$$f(x^{k+1}) + \lambda_k d_\phi(x^{k+1}, x^k) \leq f(x) + \lambda_k d_\phi(x, x^k), \forall x \in \mathbb{R}_{++}^n,$$

tomando el límite  $k \rightarrow +\infty$  en la última inecuación, usando la continuidad de  $f$ , el Lema 2.1 (c)-(d) y  $\lambda_k \rightarrow 0$  se tiene,

$$f(\bar{x}) \leq f(x), \forall x \in \mathbb{R}_{++}^n. \quad (3.11)$$

Ahora consideramos una sucesión  $\{y^k\}$  en  $\mathbb{R}_{++}^n$  que converge a  $y \in \mathbb{R}_+^n$  y usando (3.11) se concluye que,

$$f(\bar{x}) \leq f(y^k), \forall k \in \mathbb{N}, \quad (3.12)$$

tomando el límite  $k \rightarrow +\infty$  en la inecuación (3.12), usando la continuidad de  $f$ , tenemos,

$$f(\bar{x}) \leq f(y), \forall y \in \mathbb{R}_+^n.$$

Por lo tanto  $\bar{x}$  es solución del problema (3.1). □

Finalmente, consideramos el caso cuando  $U = \emptyset$ .

**Proposición 3.2.** [12] *Supongamos que  $H_1$  y  $H_2$  se cumplen. Si  $U = \emptyset$  entonces  $\lim_{k \rightarrow +\infty} f(x^k) = \inf_{x \geq 0} f(x)$  y  $\{x^k\}$  no es acotada, donde  $\{x^k\}$  es generada por **RIPM**.*

*Demostración.* Por la Proposición 3.1,  $\{f(x^k)\}$  es una sucesión decreciente y convergente. Luego, sea  $\alpha := \lim_{k \rightarrow +\infty} f(x^k) = \inf\{f(x^k)\}_{k \in \mathbb{N}}$  y supongamos por absurdo que  $\alpha \neq \inf_{x \geq 0} f(x)$ . Entonces existe  $x^* \in \mathbb{R}_+^n$ , tal que,  $f(x^*) < \alpha$ . Así,

$$f(x^*) \leq \lim_{k \rightarrow +\infty} f(x^k) = \inf\{f(x^k)\}_{k \in \mathbb{N}} \leq f(x^k), \quad \forall k \in \mathbb{N},$$

lo que implica que  $x^* \in U$ , lo cual es una contradicción con  $U = \emptyset$ , en consecuencia,

$$\lim_{k \rightarrow +\infty} f(x^k) = \inf_{x \geq 0} f(x). \quad (3.13)$$

Por otro lado, supongamos por absurdo que  $\{x^k\}$  es acotada, entonces existe una subsucesión  $\{x^{k_j}\}$  de  $\{x^k\}$ , tal que,

$$\lim_{j \rightarrow +\infty} x^{k_j} = w \geq 0.$$

De la continuidad de  $f$ ,

$$\lim_{j \rightarrow +\infty} f(x^{k_j}) = f(w).$$

Por otro lado, de (3.13) tenemos,

$$\lim_{j \rightarrow +\infty} f(x^{k_j}) = \inf_{x \geq 0} f(x).$$

Luego por la unicidad del límite se tiene que,

$$f(w) = \inf_{x \geq 0} f(x) = \lim_{k \rightarrow +\infty} f(x^k) = \inf\{f(x^k)\}_{k \in \mathbb{N}} \leq f(x^k), \quad \forall k \in \mathbb{N},$$

entonces  $w \in U$ , lo cual es una contradicción (ya que  $U = \emptyset$ ). Por lo tanto  $\{x^k\}$  no es acotada.  $\square$

## CAPÍTULO 4

# CONCLUSIONES Y TRABAJO FUTURO.

Hemos generalizado el algoritmo **RIPM** definido en (3.2) asociado a la función casi-distancia homogénea de orden  $p$ , con  $p \geq 2$ , definida en (2.6), para la minimización de funciones casi-convexas, continuamente diferenciables y acotadas inferiormente en el ortante no negativo, que fue estudiado originalmente en [12] para  $p = 2$  y a su vez algunos resultados de [12] y [9] fueron generalizados para  $p \geq 2$ . También mostramos que la sucesión generada por el algoritmo **RIPM** converge a un punto estacionario  $\bar{x}$  del problema 3.1. Además probamos que, si  $\nabla f(\bar{x}) \neq 0$  o si el parámetro proximal  $\lambda_k$  satisface la condición que  $\lambda_k \rightarrow 0$ , entonces  $\bar{x}$  es una solución del problema 3.1. Como investigación futura, estamos interesados en extender este trabajo, considerando que la función objetivo sólo sea casi-convexa y semicontinua inferiormente y además se realizarán experimentos numéricos que validen el uso de los algoritmos.



# REFERENCIAS

- [1] Bazaraa, M. S. and Shetty, C. M.: *Nonlinear Programming Theory and Algorithms*, New York: John Wiley and Sons, 3rd ed. 2006.
- [2] Bertsekas, D. P.: *Nonlinear Programming*. SECOND EDITION. Massachusetts Institute of Technology. 1995.
- [3] Cambini, A. and Martein, L.: *Generalized Convexity and Optimization (Theory and Applications)*. Springer. 2009.
- [4] Castillo, R., Hernández, E. and Campos, J.: *A  $p$ -order like-distances in Convex Programming*, submitted to *Opuscula Matemática*.2012.
- [5] Iusem, A.: *Métodos de ponto proximal em otimização*, 20°, Colóquio Brasileiro de Matemática, IMPA - Rio de Janeiro, 1995.
- [6] Mangasarian, O.L.: *Nonlinear Programming (Classics in Applied Mathematics)*, Society for Industrial Mathematics, Philadelphia, 1987.
- [7] Martinet, B.: *Perturbation des methodes d'Optimisation*. Appl. R.A.I.R.O. Numer. Anal. 12, 154–171 (1978).
- [8] Moreau, J. J.: *Promimité et Dualité dans un Espace Hilbertien*. Bull. Soc. Math. France 93, 273–299 (1965).
- [9] Pan, S. and Chen, J.: *Entropy-like proximal algorithms based on a second-order homogeneous distance function for quasi-convex programming*. Journal of Global Optimization, vol. 39, pp. 555-575, 2007.
- [10] Rockafellar, R.T.: *Augmented lagrangians and applications of proximal point algorithm in convex programming*. Math. Oper. Res. 1, 97–116 (1976).
- [11] Rockafellar, R.T.: *Monotone operators and the proximal point algorithm*. SIAM J. Control Optim. 14, 877–898 (1976).

- [12] Silva Brito, A., Oliveira, P. R., Lopes J. O. and da Cruz Neto, J.X.: *Generalized Proximal Point Algorithms for Quasiconvex Programming*. Partially supported by CAPES. PESC-COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. E-mail: arnaldo@cos.urfj.br, 2008.
- [13] Teboulle, M.: *Entropic proximal mappings with applications to nonlinear programming*. Math. Oper. Res. 17, 670–690 (1992).