

T/S10/B74

UNIVERSIDAD NACIONAL DEL CALLAO
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA PROFESIONAL DE MATEMÁTICA



**“CONVERGENCIA DEL MÉTODO DE
PUNTO PROXIMAL CON DISTANCIA
HOMOGÉNEA DE ORDEN r EN
OPTIMIZACIÓN CONVEXA”**

**TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICA**

DANTE CESAR BORDA MARCATINCO

Callao-Octubre-2013
PERÚ

ID. E = 15827
ID. P =

HOJA DE REFERENCIA DEL JURADO Y APROBACIÓN

Convergencia del Método de Punto Proximal con Distancia

Homogénea de Orden r en Optimización Convexa

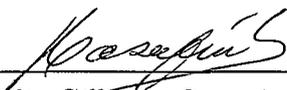
DANTE CESAR BORDA MARCATINCO

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

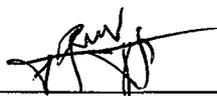
Aprobado por:



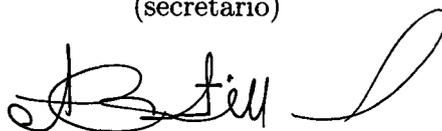
Mg. Roel Mario Vidal Guzmán
(Presidente)



Lic. Eladio Gilberto Casapia Almonte
(vocal)



Lic. Rolando Manuel Vega de la Peña
(secretario)



Lic. Absalón Castillo Valdivieso
(Suplente)



Lic. Rosa Quispe Llamoca
(Asesor)

DEDICATORIA

Para mis padres Ines y Aquiles,
para mi esposa Adela y nuestro
pequeño hijo Hector

AGRADECIMIENTO

Al finalizar un trabajo tan arduo y lleno de dificultades como el desarrollo de una tesis, es para mí un verdadero placer utilizar este espacio para ser justo y consecuente con la participación de personas e instituciones que han facilitado las cosas para que este trabajo llegue a un feliz término, expresándoles mis agradecimientos.

Primero, me gustaría agradecerte a ti Dios por bendecirme, para llegar hasta donde he llegado, porque hiciste realidad este sueño anhelado. A la UNIVERSIDAD NACIONAL DEL CALLAO por darme la oportunidad de estudiar y ser un profesional. También me gustaría agradecer a mis profesores durante toda mi carrera profesional porque todos han aportado con un granito de arena a mi formación, en especial al Dr Erik Alex Papa Quiroz, por su apoyo en gran parte de este trabajo, a la Lic. Rosa Quispe Llamoca por su participación como asesora en el desarrollo de esta tesis y al profesor Lic. Absalón Castillo Valdivieso por su minuciosa corrección en la redacción del texto des esta tesis.

Son muchas las personas que han formado parte de mi vida profesional a quienes me encantaría agradecerles su amistad, consejos, apoyo, ánimo y compañía en los momentos más difíciles de mi vida. Algunas están aquí conmigo y otras en mis recuerdos y en mi corazón, sin importar en donde estén quiero darles las gracias por formar parte de mí ser, por todo lo que me han brindado y por todas sus bendiciones. Para ellos: Muchas gracias y que Dios les bendiga.

ÍNDICE

ÍNDICE	1
LISTA DE FIGURAS	4
RESUMEN	5
ABSTRACT	6
I. PLANTEAMIENTO DE LA INVESTIGACIÓN	7
1.1. Identificación del problema	7
1.2. Formulación del problema	10
1.3. Objetivos de la investigación	11
1.3.1. Objetivo general	11
1.3.2. Objetivo específico	11
1.4. Importancia y justificación de la investigación	11
II. MARCO TEÓRICO	12
2.1. Resultados preliminares	12
2.1.1. Símbolos y notaciones	12
2.2. Análisis real	13
2.2.1. Sucesiones y series de números reales	13
2.2.2. Límite de funciones	15
2.2.3. Funciones derivables en un intervalo	16
2.3. Elementos básicos de análisis convexo	16
2.3.1. Conos asintóticos	17

2.3.2.	Funciones convexas y asintóticas	18
2.3.3.	Conjugación y sub-diferencial	20
2.4.	Optimización	22
2.5.	Método de punto proximal	27
2.5.1.	Motivación	31
2.5.2.	φ - divergencia	34
2.5.3.	Método de punto proximal con φ -divergencia	40
2.5.4.	Método punto proximal con distancia homogénea de orden dos	42
2.6.	Métodos proximales basados en núcleos homogéneos de orden r	49
2.6.1.	Métodos proximales con distancia r -homogéneas	50
2.6.2.	Análisis de convergencia	53
III. VARIABLES E HIPÓTESIS		66
3.1.	Variables de la investigación	66
3.2.	Operacionalización de variables	66
3.3.	Hipótesis	67
IV. METODOLOGÍA		68
4.1.	Tipo de investigación	68
4.2.	Diseño de la investigación	68
4.3.	Población y muestra	69
4.4.	Técnicas e instrumentación de recolección de datos	69
4.5.	Procedimiento de recolección de datos	69
4.6.	Procesamiento estadístico y análisis de datos	69
V. RESULTADOS		70
VI. DISCUSIÓN DE RESULTADOS		71
VII. CONCLUSIONES		72
VIII. RECOMENDACIONES		73
REFERENCIAS BIBLIOGRÁFICAS		74

ANEXO	75
A. Matriz de Consistencia	77
B. Mapa conceptual del trabajo	78

FIGURAS DE CONTENIDO

II.1. Figura 1	35
II.2. Figura 2	37
II.3. Figura 3	38

RESUMEN

CONVERGENCIA DEL MÉTODO DE PUNTO PROXIMAL CON DISTANCIA HOMOGÉNEA DE ORDEN r EN OPTIMIZACIÓN CONVEXA

Dante Cesar Borda Marcatinco

Octubre 2013

Título Obtenido: Licenciado en matemática

Considere el problema

$$(P) \begin{cases} \min & f(x) \\ \text{s. a.} & x \in \mathbb{R}_{++}^n. \end{cases}$$

donde $f : \mathbb{R}^n \rightarrow (-\infty; +\infty]$ es una función propia convexa y cerrada. En este trabajo presentaremos un nuevo método de punto proximal, basados en núcleos homogéneos de orden " r " de la forma $x^{k+1} \in \operatorname{argmin} \left\{ f(x) + \lambda_k^{-1} \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right); x \in \mathbb{R}_{++}^n \right\}$, donde

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right).$$

para los casos $r = 1$ y $r = 2$ tenemos el algoritmo φ -divergencia y logaritmo cuadrático respectivamente introducidos por Teboulle en 1994 y Auslender, Teboulle y Ben Tiba en 1999. Mostraremos las propiedades de convergencia del método para una solución óptima del problema (P).

Palabras Claves: Método de Punto Proximal, Phi-divergencias, Distancias Logaritmo Cuadrático, Núcleos homogéneos.

ABSTRACT

CONVERGENCE OF PROXIMAL POINT METHOD WITH ORDER HOMOGENEOUS r DISTANCE IN CONVEX OPTIMIZATION

Dante Cesar Borda Marcatinco

October 2013

Obtained Degree: Licentiate in Mathematics

Consider the problem

$$(P) \begin{cases} \min & f(x) \\ \text{s. a.} & x \in \mathbb{R}_{++}^n. \end{cases}$$

where $f : \mathbb{R}^n \rightarrow (-\infty; +\infty]$ is a closed convex functions. In this paper we present a new proximal point method based on homogeneous kernels of order “ r ” of the form $x^{k+1} \in \operatorname{argmin} \left\{ f(x) + \lambda_k^{-1} \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right); x \in \mathbb{R}_{++}^n \right\}$, where

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right).$$

for cases $r = 1$ and $r = 2$ have the algorithm *phi*-divergence and quadratic logarithm respectively introduced by Teboulle in 1994 and Auslender, Teboulle and Ben Tiba in 1999. We show the convergence properties of the method for an optimal solution of problem (P).

Keywords: Proximal Point Method, Phi-divergence, logarithm Quadratic Distances, homogeneous kernels.

CAPÍTULO I

PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1. Identificación del problema

En el campo de la Matemática Aplicada, existen diversas áreas que dan solución a problemas vinculados a las ciencias e ingenierías, una de ellas es la Optimización Matemática que estudia la forma de encontrar la mejor solución a un problema dado dentro de todas las posibles alternativas.

La Optimización Matemática estudia el problema de maximizar o minimizar una función, llamada función objetivo, sujeta a algunas restricciones sobre su dominio. La Optimización, como línea de investigación, surgió a mediados del siglo pasado y en este intervalo de tiempo ha demostrado diversas aplicaciones en diferentes áreas de las Ciencias e Ingenierías, donde una elección óptima de los parámetros y variables con lleva al mejoramiento de las técnicas para resolver el problema planteado. Dado el siguiente problema

$$(P) \begin{cases} \min & f(x) \\ \text{s. a.} & x \in X, \end{cases}$$

donde $f : X \rightarrow \mathbb{R}$ es una función arbitraria y X es un subconjunto de \mathbb{R}^n .

En la actualidad existen diversos software para resolver problemas de optimización, estos se basan en algoritmos iterativos, para un punto inicial x^0 se genera una sucesión de puntos $\{x^k\}$ tal que $x^{k+1} = \varphi(x^k)$, donde φ es una cierta función.

Los métodos de optimización en general tienen el siguiente esquema:

Datos de entrada: $x^0 \in X$ un punto arbitrario y ϵ el error de aproximación.

Para $k = 0, 1, 2, 3, \dots$

Si x^k satisface alguna condición de optimalidad entonces parar.

Caso contrario, hacer $x^{k+1} = \varphi(x^k)$. Fin (Para).

Para resolver (P) existen diversos métodos iterativos por ejemplo:

Método del gradiente proyectado $\varphi(x^k) = P_X (x^k - \lambda_k \nabla f(x^k))$.

Método de Newton proyectado $\varphi(x^k) = P_X (x^k - \lambda_k (\nabla^2 f(x^k))^{-1} \nabla f(x^k))$.

Métodos Cuasi-Newton proyectado $\varphi(x^k) = P_X (x^k - \lambda_k H_k \nabla f(x^k))$ con $H_k \approx (\nabla^2 f(x^k))^{-1}$,

donde $\lambda_k > 0$ y P_X denota el operador proyección ortogonal sobre X .

Cuando la función objetivo es convexa y diferenciable y el conjunto de restricciones X es convexo existen métodos que resuelven problemas en un tiempo razonable y los algoritmos finalizan generalmente en un mínimo global.

Cuando la función es no diferenciable existen algunos métodos que obtienen solución del problema (P) por ejemplo:

El método subgradiente proyectado, $x^{k+1} = P_X (x^k - \lambda_k g^k)$ donde $g^k \in \partial f(x^k)$ y $\partial f(x^k)$ es el subdiferencial de f en x^k .

El método del punto proximal, el algoritmo clásico de punto proximal, para minimizar una función convexa f en \mathbb{R}^n genera una sucesión $\{x^k\}$ por el esquema iterativo: comenzar con un punto inicial $x^0 \in \mathbb{R}^n$ y resolver

$$x^k = \operatorname{argmin} \left\{ f(x) + \frac{1}{2} \lambda_k \|x - x^{k-1}\|^2 : x \in X \right\},$$

donde $\lambda_k > 0$ y $\|\cdot\|$ denota la norma euclidiana. La notación *argmin* significa que x^k es el único punto de mínimo global de $\left\{ f(x) + \frac{1}{2} \lambda_k \|x - x^{k-1}\|^2 \right\}$ sobre X .

Se han realizado varias propuestas para sustituir el término de segundo grado $\|x - x^{k-1}\|^2$ por distancias similares, por ejemplo:

Censor y Zenios véase [5] propuso la minimización con D-funciones, generando la sucesión $\{x^k\}$ por $x^k = \operatorname{argmin} \{ f(x) + \lambda_k^{-1} D_h(x, x^{k-1}) : x \in \mathbb{R}_+^n \}$, donde $\{\lambda_k\}$ es una sucesión de números positivos, la función D es llamada distancia de Bregman

y definida por:

$$D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle$$

donde h es una función continua, estrictamente convexa y diferenciable en el interior de su dominio y ∇h denota el gradiente de h .

En 1992, Teboulle ver [14, 15], modifica el método de punto proximal reemplazando la distancia cuadrática por la casi-distancia, también conocida con el nombre de φ -divergencia, generando la sucesión $\{x^k\}$ por

$$x^k = \operatorname{argmin} \left\{ f(x) + \lambda_k^{-1} \tilde{d}_\varphi(x, x^{k-1}) : x \in \mathbb{R}_+^n \right\}, \quad (1.1.1)$$

donde $\{\lambda_k\}$ es una sucesión de números positivos y

$$\tilde{d}_\varphi(x, y) := \sum_{j=1}^n y_j \varphi(y_j^{-1} x_j). \quad (1.1.2)$$

Es la cuasi-distancia φ -divergencia, llamada distancia homogénea de primer orden, basada en una función estrictamente convexa φ , una importante elección de φ es $\varphi(t) = t \ln t - t + 1$, para cual la correspondiente d_φ es bien conocida como Kullback-Leibler función entrópica de estadísticas.

El algoritmo generado por 1.1.2, también es llamado cuasi-proximal, este algoritmo con $\varphi(t) = -\ln t + t - 1$ fue propuesto primero por Eggermont en [7]. Similares resultados de convergencia del método casi-proximal se obtienen usando distancias de Bregman. El análisis del método casi-proximal basado en distancias de Bregman no puede ser llevado a cabo igual que en el algoritmo definido en 1.1.1, excepto en el caso que $\varphi(t) = t \ln t - t + 1$ donde las dos distancias coinciden.

En 1999, Auslender, Teboulle y Ben-tiba ver [1], propusieron sustituir en la iteración proximal (1.1.2) a la función \tilde{d}_φ por

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^2 \varphi(y_j^{-1} x_j),$$

llamada distancia homogénea de segundo orden. Observe que d_φ tiene la misma propiedad del método introducido por Teboulle salvo el término cuadrático que permite que d_φ sea dos veces diferenciable, la fundamental diferencia aquí es que el término d_φ es usado para forzar que las iteraciones $\{x^k\}$ se mantengan en el

ortante no negativo \mathbb{R}_{++}^n evitando realizar proyecciones en cada iteración, es decir el algoritmo genera automáticamente términos positivos. Auslender, Teboulle y Ben-Tiba, comienzan generando un algoritmo que llaman "Basic Iterations Scheme" (BIS), que es una versión aproximada del método de punto proximal, el cual usa la distancia homogénea de segundo orden. En [1] se estudian dos clases de algoritmos basados en BIS, estos son "Interior Proximal Method" (IPM) y "Regularized Interior Proximal Method" (RIPM).

Una idea que surge naturalmente es explorar las propiedades de convergencia y ventajas del método del punto proximal (1.1.2) considerando la función:

$$d_{\varphi}(x, y) := \sum_{j=1}^p y_j^r \varphi(y_j^{-1} x_j),$$

con $r \geq 1$, la que llamamos distancia homogénea de orden r . Esta distancia incluye en particular las distancias de primer y segundo orden introducidas por Teboulle (1992) y Auslender, Teboulle y Ben-tiba (1999), además φ tiene las mismas condiciones que en los algoritmos anteriores tanto para distancia de primer orden como distancia de segundo orden.

1.2. Formulación del problema

Lo que se pretende analizar y responder es la siguiente interrogante:
 ¿Será posible obtener la convergencia del método de punto proximal que usan las distancias homogéneas d_{φ} de orden r , en problemas de optimización convexa?

1.3. Objetivos de la investigación

1.3.1. Objetivo general

Analizar la convergencia del método de punto proximal cuya sucesión de puntos $\{x^k\}$ es dado por el algoritmo

$$x^k = \operatorname{argmin} \{f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) : x \in \mathbb{R}_+^n\}$$

para la distancia homogénea de orden r

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right),$$

con $r > 2$ en problemas de optimización de convexa

1.3.2. Objetivo específico

- a. Estudiar las propiedades de convergencia del método proximal usando las distancias homogéneas d_φ de orden r , para $r = 1$ y $r = 2$.
- b. Demostrar las propiedades de las distancias homogéneas d_φ de orden r , para distancias $r > 2$.

1.4. Importancia y justificación de la investigación

La aplicación de la optimización convexa se encuentra en diversas áreas de las ciencias e ingeniería, por ejemplo en la Economía para maximizar ganancias y minimizar costos, en la Física, para la estabilidad de sistemas físicos, en Administración, para problemas de cartera de inversión, en ingeniería para telecomunicaciones, circuitos eléctricos, transportes, localización y otros. Es por eso, que resulta importante la investigación de métodos de optimización, en particular del método de punto proximal. El desarrollo de este trabajo de investigación esta justificado desde que no existen estudios realizados para las distancias homogéneas d_φ de orden r , con $r > 2$.

CAPÍTULO II

MARCO TEÓRICO

2.1. Resultados preliminares

2.1.1. Símbolos y notaciones

A lo largo de la tesis adoptaremos las siguientes terminologías,

$\mathbb{N} = \{1, 2, 3, 4, \dots\}$; el conjunto de los números naturales.

$\mathbb{R}^n = \{x = (x_1, x_2, \dots, x_n); x_i \in \mathbb{R}, \forall i = 1, \dots, n\}$: el espacio vectorial euclidiano.

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

$\mathbb{R}_{++}^n = \{x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n; x_i > 0\}$: el ortante positivo.

Dado $x \in \mathbb{R}^n$, $x \geq 0$ significa $x_i \geq 0, \forall i = 1, \dots, n$.

Dado $x \in \mathbb{R}^n$, $x > 0$ significa $x_i > 0, \forall i = 1, \dots, n$.

$\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

$\partial(\Omega)$: la frontera de Ω .

$L^2(\Omega) = \{f : \Omega \rightarrow \mathbb{C}, \text{ medible en el sentido de Lebesgue y } \int_{\Omega} |f|^2 < \infty\}$.

$H_0^1(\Omega) = \{u \in L^2(\Omega) / u' \in L^2(\Omega) \wedge u(x) = 0, x \in \partial(\Omega)\}$.

Definamos las siguientes operaciones:

$$\begin{aligned}
 r \cdot (+\infty) &= +\infty, & \text{si } r > 0. \\
 r \cdot (+\infty) &= -\infty, & \text{si } r < 0. \\
 r + (+\infty) &= +\infty, & \text{si } r \in \mathbb{R}. \\
 r - (+\infty) &= -\infty, & \text{si } r \in \mathbb{R}. \\
 0 \cdot (+\infty) &= 0 \cdot (-\infty) = 0. \\
 +\infty - \infty &= \text{indefinido}.
 \end{aligned}$$

2.2. Elementos básicos de análisis real

2.2.1. Sucesiones y series de números reales

Definición II.1 Una sucesión de números reales es una función $x : \mathbb{N} \rightarrow \mathbb{R}$.

Escribiremos $\{x_1, \dots, x_n, \dots\}$, o $\{x_n\}_{n \in \mathbb{N}}$, o simplemente $\{x_n\}$ para indicar a la sucesión x .

Diremos que una sucesión $\{x_n\}$ es *limitada*, cuando existe $M \in \mathbb{R}$ tal que $|x_n| \leq M$.

Definición II.2 Diremos que un número real a es *límite* de una sucesión $\{x_n\}$ de números reales cuando para $\epsilon > 0$ existe un entero $n_0 \in \mathbb{N}$ tal que $|x_n - a| < \epsilon$, para todo $n > n_0$. Denotamos $a = \lim_{n \rightarrow \infty} x_n$. Simbólicamente:

$$\lim_{n \rightarrow \infty} x_n = a \equiv \forall \epsilon > 0 \exists n_0 \in \mathbb{N}; n > n_0 \Rightarrow |x_n - a| < \epsilon.$$

Teorema II.1 (*Unicidad del límite*). Si $\lim_{n \rightarrow \infty} x_n = a$ y $\lim_{n \rightarrow \infty} x_n = b$, entonces $a=b$.

Demostración. Véase [13], página 85. ■

Teorema II.2 Si $\lim_{n \rightarrow \infty} x_n = 0$ y $\{y_n\}$ es una sucesión limitada, entonces $\lim_{n \rightarrow \infty} x_n y_n = 0$.

Demostración. Véase [13], página 156. ■

Teorema II.3 Si $\lim_{n \rightarrow \infty} x_n = a$ y $\lim_{n \rightarrow \infty} y_n = b$, entonces

1. $\lim_{n \rightarrow \infty} (x_n + y_n) = a + b$.

$$2. \lim_{n \rightarrow \infty} x_n y_n = ab.$$

$$3. \lim_{n \rightarrow \infty} \left(\frac{x_n}{y_n} \right) = \left(\frac{a}{b} \right), \text{ si } b \neq 0.$$

Demostración. Véase [13], página 90. ■

Corolario II.1 Sean $\{x_n\}$ y $\{y_n\}$ dos sucesiones convergentes. Si $x_n \leq y_n$ para todo $n \in \mathbb{N}$ entonces,

$$\lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n.$$

Demostración. Véase [13], página 92. ■

Definición II.3 Sea una sucesión x_n limitada superiormente y tomemos $M_n = \sup \{x_n, x_{n+1}, \dots\}$. Si $\{M_n\}$ converge, definimos el límite superior

$$\limsup_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} M_n.$$

Definición II.4 Sea una sucesión x_n limitada inferiormente y tomemos $m_n = \inf \{x_n, x_{n+1}, \dots\}$. Si $\{m_n\}$ converge, definimos el límite inferior

$$\liminf_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} m_n.$$

Sea $\{x_n\}$ una sucesión limitada; digamos con $\alpha \leq x_n \leq \beta$ para todo $n \in \mathbb{N}$. Escribiendo $X_n = \{x_n, x_{n+1}, \dots\}$, tenemos

$$[\alpha, \beta] \supset X_1 \supset X_2 \supset \dots \supset X_n \supset \dots$$

Para $a_n = \inf X_n$ y $b_n = \sup X_n$, obtenemos:

$$\alpha \leq a_1 \leq a_2 \leq \dots \leq a_n \leq \dots \leq b_n \leq \dots \leq b_2 \leq b_1 \leq \beta.$$

Por lo tanto existen los siguientes límites:

$$\lim a_n = \sup_n a_n = \sup \inf X_n.$$

$$\lim b_n = \inf_n b_n = \inf \sup X_n.$$

También se puede notar

$$\sup_n \inf X_n \leq \inf_n \sup X_n.$$

Teorema II.4 Si $\{x_n\}$ es una sucesión convergente de números reales, entonces

$$\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n.$$

Demostración. Véase [13], página 96. ■

Teorema II.5 Sean $\{x_n\}$ y $\{y_n\}$ dos sucesiones limitadas de números reales, entonces las siguientes relaciones son válidas

1. $\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n.$
2. $\liminf_{n \rightarrow \infty} x_n + \liminf_{n \rightarrow \infty} y_n \leq \liminf_{n \rightarrow \infty} (x_n + y_n).$
3. $\limsup_{n \rightarrow \infty} (x_n + y_n) \leq \limsup_{n \rightarrow \infty} x_n + \limsup_{n \rightarrow \infty} y_n.$
4. Si $x_n \leq y_n$ para todo $n \in \mathbb{N}$, entonces $\liminf_{n \rightarrow \infty} (x_n) \leq \liminf_{n \rightarrow \infty} (y_n).$
Además, $\limsup_{n \rightarrow \infty} (x_n) \leq \limsup_{n \rightarrow \infty} (y_n).$

Demostración. Véase [2], página 148. ■

Teorema II.6 Sea $\sum_{n=1}^{\infty} a_n$, $a_n \geq 0$, una serie convergente, entonces $\lim_{n \rightarrow \infty} a_n = 0.$

Demostración. Véase [13], página 106. ■

2.2.2. Límite de funciones

Definición II.5 Diremos que un número real L es el límite de f cuando x tiende para a denotado por $\lim_{x \rightarrow a} f(x) = L$, si para cada $\epsilon > 0$ existe un $\delta > 0$, tal que $|f(x) - L| < \epsilon$ cuando $0 < |x - a| < \delta.$

Teorema II.7 (Teorema del Sandwich). Sea $X \subset \mathbb{R}$ y $f, g, h : X \rightarrow \mathbb{R}.$ Si para todo $x \in X$, $x \neq a$ tal que $f(x) \leq g(x) \leq h(x)$ y $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = L$, entonces

$$\lim_{x \rightarrow a} h(x) = L.$$

Demostración. Véase [13], página 155. ■

2.2.3. Funciones derivables en un intervalo

Teorema II.8 (*Teorema del Valor Medio*). Si $f : [a, b] \rightarrow \mathbb{R}$ es continua, y derivable en (a, b) , entonces existe $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Demostración. Véase [13], página 213. ■

2.3. Elementos básicos de análisis convexo

Definición II.6 Sea $C \subset \mathbb{R}^n$ y $K \subset \mathbb{R}^n$ entonces

1. Un subconjunto C es llamado **convexo** si $\forall x, y \in C, \forall t \in [0, 1]$

$$tx + (1 - t)y \in C.$$

2. Un subconjunto K es llamado **cono** $\forall d \in K \wedge \forall t \geq 0$ se tiene

$$td \in K.$$

Definición II.7 Dado un conjunto $C \subset \mathbb{R}^n$. Se dice que x es una combinación convexa de elementos de C si existen: $p \in \mathbb{N}$, $\{t_i\}_{i=1}^p \subset [0, 1]$ y $\{x_i\}_{i=1}^p \subset C$ tales que:

$$\sum_{i=1}^p t_i x_i = x \quad \sum_{i=1}^p t_i = 1$$

Definición II.8 Dado un conjunto $C \subset \mathbb{R}^n$ no vacío, diremos que la cápsula convexa de C , la cual se denotará como $C_0(C)$, es la intersección de todos los conjuntos convexos que contiene a C .

Observación II.1 La cápsula convexa $C_0(C)$ de un conjunto convexo C es también convexa. Mas aún, la cápsula convexa de un conjunto C , es el menor conjunto convexo que contiene a C .

Definición II.9 *Los hiperplanos son conjuntos definidos mediante:*

$$S = \{x \in \mathbb{R}^n / p^T \cdot x = \alpha\}$$

Donde p es un vector no nulo de \mathbb{R}^n y $\alpha \in \mathbb{R}$. Empleando la definición de conjunto convexo, se demuestra fácilmente que S es un conjunto convexo. Tomando $x, y \in S$, por la definición de S , se obtiene que $p^T \cdot x = \alpha$ y $p^T \cdot y = \alpha$. Para cualquier $\lambda \in [0, 1]$, se obtiene:

$$p^T[\lambda x + (1 - \lambda)y] = \lambda p^T \cdot x + (1 - \lambda)p^T \cdot y = \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

Definición II.10 *Un conjunto C no vacío es un cono si $x \in C$ implica que $\alpha x \in C$ para todo $\alpha > 0$. Si además, C es convexo, se dice que es un cono convexo.*

Definición II.11 *Sea $C \subset \mathbb{R}^n$ un conjunto convexo y $\bar{x} \in C$. El cono normal (cono de direcciones normales) en el punto \bar{x} en relación al conjunto C es dado por:*

$$N_C(\bar{x}) = \{d \in \mathbb{R}^n / \langle d, x - \bar{x} \rangle \leq 0 \forall x \in C\}$$

2.3.1. Conos asintóticos

Definición II.12 *Dado $C \subset \mathbb{R}^n$ convexo y no vacío, se llama cono asintótico (cono de recesión) al conjunto*

$$C_\infty = \bigcap_{a \in C} C_\infty(a),$$

donde $C_\infty(a) = \{d \in \mathbb{R}^n : a + \lambda d \in C, \forall \lambda > 0\}$.

Los elementos $d \in C_\infty$ son llamados direcciones de recesión de C .

Teorema II.9 *Sea $C \subset \mathbb{R}^n$ no vacío convexo y cerrado, entonces*

$$C_\infty(a) = C_\infty(b), \forall a, b \in C.$$

Además, $C_\infty = C_\infty(a) = C_\infty(b)$ es cerrado.

Demostración. Véase [6], página 10. ■

Proposición II.1 *Sea $C \subset \mathbb{R}^n$ no vacío, convexo y cerrado, entonces C es acotado, si y solo si, $C_\infty(a) = \{0\}$.*

Demostración. Véase [6], página 10. ■

2.3.2. Funciones convexas y asíntóticas

Definición II.13 Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, una función de valores extendidos, definimos al dominio efectivo de f , denotado por $\text{dom}f$, a

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

Definición II.14 $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es propia si

i) $\text{dom}f \neq \emptyset$,

ii) $\forall x \in \text{dom}f, f(x) > -\infty$.

Definición II.15 Una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es llamada **convexa** si

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \forall x, y \in \mathbb{R}^n \text{ y } \forall \alpha \in [0, 1].$$

Se dice que f es **estrictamente convexa** cuando la desigualdad es estricta $\forall x, y \in \mathbb{R}^n$ tal que $x \neq y$.

se dice que f se **fuertemente convexa** con módulo $\gamma > 0$, si: $\forall x, y \in \mathbb{R}^n \text{ y } \forall \lambda \in (0, 1), f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) - \gamma \lambda(1 - \lambda) \|x - y\|^2$

Se nota que una función fuertemente convexa es estrictamente convexa y una función estrictamente convexa es convexa.

Definición II.16 $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ es **semicontinua inferior (sci)** si $\forall \{x^k\} \subset \mathbb{R}^n$ tal que $\lim_{k \rightarrow \infty} x^k = \bar{x}$ se tiene que

$$f(\bar{x}) \leq \liminf_{k \rightarrow \infty} f(x^k).$$

Definición II.17 f es llamado **cerrado** si es sci.

Definición II.18 $X \subset \mathbb{R}^n, X \neq \emptyset$, definimos la función **indicatriz** de X como

$$\delta(x|X) = \begin{cases} 0 & \text{si } x \in X \\ +\infty & \text{si } x \notin X. \end{cases}$$

Definición II.19 El conjunto de nivel de una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, es dado por

$$L_f(\alpha) = \{x \in \mathbb{R}^n / f(x) \leq \alpha\}.$$

Definición II.20 El epígrafo de una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, es el conjunto

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

Teorema II.10 Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función, los siguientes enunciados son equivalentes:

- i) f es sci,
- ii) $\text{epi}(f)$ es cerrado,
- iii) $L_f(\alpha)$ es cerrado $\forall \alpha \in \mathbb{R}$.

Demostración. Véase [6], página 16. ■

Definición II.21 Sea f función convexa propia sci. Se llama **función asintótica** (o función de recesión), a una cierta función, denotada por f_∞ (o f_0^+ Rockafellar), tal que

$$\text{epi}(f_\infty) = (\text{epi}(f))_\infty.$$

Proposición II.2 Sea f una función convexa propia sci, entonces f_∞ es positivamente homogénea de grado 1, es decir,

$$f_\infty(kd) = kf_\infty(d), \quad \forall k > 0,$$

$$f_\infty(0) = 0.$$

Demostración. Véase [6], página 24. ■

Proposición II.3 Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ convexa propia sci y $a \in \text{dom}(f)$, entonces

$$\begin{aligned} f_\infty(d) &= \sup_{k>0} \frac{f(a+kd) - f(a)}{k} \\ &= \lim_{k \rightarrow \infty} \frac{f(a+kd) - f(a)}{k} \\ &= \lim_{k \rightarrow \infty} \frac{f(a+kd)}{k}. \end{aligned}$$

Demostración. Véase [6], página 24. ■

2.3.3. Conjugación y sub-diferencial

Definición II.22 Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, se llama conjugación (en el sentido de Fenchel) a la función siguiente:

$$f^*(x^*) = \sup_{x \in \text{dom}f} [\langle x, x^* \rangle - f(x)],$$

donde $x^* \in \mathbb{R}^n$.

Teorema II.11 Sea f una función convexa propia y sci, entonces, f^* es convexa propia y sci. Además $f^{**} = f$.

Demostración. Véase [6], página 36. ■

Definición II.23 La función conjugación de la función indicatriz de C , es llamada función soporte de C . Se tiene entonces

$$\delta^*(x^*|C) = \sup_{x \in C} [\langle x, x^* \rangle - \delta(x|C)] = \sup_{x \in C} \langle x, x^* \rangle.$$

Sea C un conjunto no vacío, se llama cono barrera de C al conjunto K definido por

$$K = \{x^* : \delta^*(x^*|C) < +\infty\} = \text{dom}(\delta^*(\cdot|C)).$$

Proposición II.4 Sea una función convexa propia y sci, entonces f_∞ es la función soporte del $\text{dom}(f^*)$, es decir,

$$f_\infty(d) = \delta^*(d|\text{dom}(f^*)).$$

Demostración. Véase [6], página 44. ■

Definición II.24 Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función convexa. Diremos que $y \in \mathbb{R}^n$ es un subgradiente de f en el punto $x \in \mathbb{R}^n$ si

$$f(z) \geq f(x) + \langle y, z - x \rangle, \forall z \in \mathbb{R}^n.$$

El conjunto de todos los subgradientes de f en x , se llama subdiferencial de f en x y lo denotamos por $\partial f(x)$.

Proposición II.5 Sea f una función convexa propia, y $a \in \text{dom}(f)$ entonces,

$$\partial f(a) \neq \emptyset, \text{ si y solo si, } a \in \text{int}(\text{dom}(f)).$$

Demostración. Véase [6], página 46. ■

Definición II.25 Sea $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ una función convexa, $\epsilon > 0$, diremos que $y \in \mathbb{R}^n$ es un ϵ -subgradiente de f en el punto $x \in \mathbb{R}^n$ cuando

$$f(z) \geq f(x) + \langle y, z - x \rangle - \epsilon, \forall z \in \mathbb{R}^n.$$

Al conjunto de todos los ϵ -subgradiente de f en x , denotado por $\partial_\epsilon f(x)$, se llama ϵ -subdiferencial de f en x .

Observación II.2 $0 \in \partial_\epsilon f(x)$, si y solo si, $f(x) \leq \inf_{z \in \mathbb{R}^n} f(z) + \epsilon$.

Definición II.26 Se denomina vector gradiente al vector de derivadas parciales y se le denota por

$$\nabla f(x)^T = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$$

Funciones convexas diferenciables

Cuando una función convexa es diferenciable, se puede caracterizar a través de la gradiente, como se muestra en el siguiente teorema.

Teorema II.12 Sea $C \subset \mathbb{R}^n$ un conjunto convexo y abierto $f : C \rightarrow \mathbb{R}$ una función diferenciable en C . Entonces las propiedades siguientes son equivalentes:

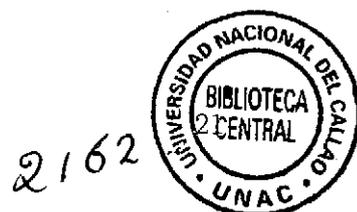
- (i) La función f es convexa en C .
- (ii) Para todo $x \in C$ y para todo $y \in C$, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.
- (iii) Para todo $x \in C$ y para todo $y \in C$, $\langle \nabla f(x) - \nabla f(y), y - x \rangle \geq 0$.

Cuando la función f es dos veces diferenciable en C , las propiedades arriba también son equivalentes a:

- (iv) La matriz Hessiana de f es semidefinida positiva para todo punto de C :

$$\langle f'' d, d \rangle \geq 0, \forall x \in C \forall d \in \mathbb{R}^n$$

Demostración. Véase [12]



2162

Funciones convexas no diferenciables

Subgradiente y subdiferencial de una función convexa

Definición II.27 Sea f una función convexa en \mathbb{R}^n . El elemento $x^* \in \mathbb{R}^n$ es llamado subgradiente de f en x^* si:

$$f(z) \geq f(x) + x^* \cdot (z - x), \forall z \in \mathbb{R}^n$$

Definición II.28 El conjunto de subgradientes de f es llamado subdiferencial de f en x y es denotado por $\partial f(x)$. Si $\partial f(x)$ es no vacío, decimos que f es subdiferenciable en x .

Proposición II.6 Una función convexa $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en el punto $x \in \mathbb{R}^n$ si, y solamente si, el conjunto $\partial f(x)$ contiene un solo elemento. En este caso $\partial f(x) = \{\nabla f(x)\}$.

Demostración. Véase [12]

Proposición II.7 Sean f_1, f_2, \dots, f_p son p funciones convexas de \mathbb{R}^n en $\mathbb{R} \cup \{+\infty\}$. Si $\bigcap_{i=1}^p \text{ri}(\text{dom}(f_i)) \neq \emptyset$ entonces $\forall x \in \mathbb{R}^n$,

$$\partial(f_1 + f_2 + \dots + f_p)(x) = \partial f_1(x) + \partial f_2(x) + \dots + \partial f_p(x)$$

Demostración. Véase[8]

Proposición II.8 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función convexa, entonces \bar{x} minimiza f en \mathbb{R}^n si, y solo si, $0 \in \partial f(\bar{x})$.

Demostración. Véase [12]

2.4. Optimización

La Optimización Matemática es una línea de investigación de la Matemática Aplicada que estudia el problema:

$$\begin{cases} \min & f(x) \\ \text{s. a.} & x \in X. \end{cases} \quad (2.4.1)$$

Donde $f : X \rightarrow \mathbb{R}$ es una función arbitraria y X es un subconjunto de \mathbb{R}^n .

Definición II.29 (*Mínimo local y global*)

Sea $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una función definida sobre X . Un punto $\bar{x} \in X$ es llamado un punto de mínimo local de f sobre X si existe $\epsilon > 0$ tal que

$$f(\bar{x}) \leq f(x), \forall x \in X \cap B(\bar{x}, \epsilon).$$

El punto $\bar{x} \in X$ es llamado punto de mínimo global de f sobre X si

$$f(\bar{x}) \leq f(x), \forall x \in X.$$

Definición II.30 (*Máximo local y global*)

Sea $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una función definida sobre X . Un punto $\bar{x} \in X$ es llamado un punto de máximo local de f sobre X si existe $\epsilon > 0$ tal que

$$f(\bar{x}) \geq f(x), \forall x \in X \cap B(\bar{x}, \epsilon).$$

El punto $\bar{x} \in X$ es llamado punto de máximo global de f sobre X si

$$f(\bar{x}) \geq f(x), \forall x \in X.$$

Observación II.3 *Todo problema de maximización puede ser expresado como un problema de minimización. En efecto, el problema*

$$\begin{cases} \max & g(x) \\ \text{s. a.} & x \in X \end{cases}$$

es equivalente a

$$\begin{cases} \min & -g(x) \\ \text{s. a.} & x \in X. \end{cases}$$

Teorema II.13 *Si $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es semicontinua inferior en un conjunto no vacío y compacto X entonces, existe un punto de mínimo global.*

Demostración. Véase [6], página 17. ■

Teorema II.14 (*Teorema de minimización convexa*). *Sean $X \subset \mathbb{R}^n$ un conjunto convexo y $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa en X . Entonces todo minimizador local del problema (1.1) es global. Además, el conjunto de minimizadores es convexo. Si f es estrictamente convexa, no puede tener más de un minimizador.*

Demostración. Véase [12], página 69. ■

Teorema II.15 *Sea $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa. Supongamos que exista un $\lambda \in \mathbb{R}$ tal que el conjunto de nivel $L_f(\lambda)$ es no vacío y limitado, entonces $L_f(\lambda)$ es limitado para todo $t \in \mathbb{R}$.*

Demostración Véase [12], página 140. ■

Teorema II.16 *(Condición de optimalidad para minimización de una función convexa en un conjunto convexo)*

Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa y $C \subset \mathbb{R}^n$ un conjunto convexo, entonces $\bar{x} \in \mathbb{R}^n$ es un minimizador de f en C si y solo si

$$\text{existe } y \in \partial f(\bar{x}) \text{ tal que } \langle y, x - \bar{x} \rangle \geq 0, \forall x \in C,$$

o equivalentemente,

$$0 \in \partial f(\bar{x}).$$

Demostración. Véase [12], página 168. ■

Proposición II.9 *Sea f convexa sci propia sobre \mathbb{R}^n , entonces el conjunto mínimo de f es no vacío y compacto, si y solo si,*

$$f_\infty(d) > 0, \forall d \neq 0.$$

Demostración.- Véase [17], página 265. ■

Proposición II.10 *Sea f una función convexa de valor real en \mathbb{R}^n entonces los conjuntos de nivel $L_{f, \mathbb{R}^n}(\alpha)$ no vacíos, tienen el mismo cono asintótico.*

Demostración.-

Sea α_1 y α_2 satisfaciendo $\alpha_1 \leq \alpha_2$, $L_{f, \mathbb{R}^n}(\alpha_1) \neq \emptyset \wedge L_{f, \mathbb{R}^n}(\alpha_2) \neq \emptyset$. Sean A_1 y A_2 denota los conos asintóticos de los conjuntos de nivel.

$L_{f, \mathbb{R}^n}(\alpha_1) \subset L_{f, \mathbb{R}^n}(\alpha_2)$, se sabe que $A_1 \subset A_2$.

Probaremos la otra inclusión:

Sea $y \in A_2$. Tomemos $x \in L_{f, \mathbb{R}^n}(\alpha_1) \subset L_{f, \mathbb{R}^n}(\alpha_2)$. Dado $\lambda > 0$, $f(x + \lambda y) \leq \alpha_2$.

Definamos la siguiente función $g : \mathbb{R}^+ \setminus \{0\} \rightarrow \mathbb{R}$ definido de la siguiente manera

$$\lambda \mapsto g(\lambda) = f(x + \lambda y).$$

Es no creciente. Sea $(\lambda_i)_{i \in \mathbb{N}}$ una sucesión de números positivos que tienden a cero. Por la continuidad de f , de tal manera $(g(\lambda_i)_{i \in \mathbb{N}}) = (f(x + \lambda_i y)_{i \in \mathbb{N}})$ convergen a $f(x)$ cuando i tiende a $+\infty$.

Por lo tanto $\forall \lambda > 0$, allí existe un N_λ algo semejante que para cualquier $i \geq N_\lambda$, $f(x + \lambda y) \leq f(x + \lambda_i y)$ y pasando límite $f(x + \lambda y) \leq f(x) \leq \alpha_1$.

Así $x + \lambda y \in L_{f, \mathbb{R}^n}(\alpha_1)$ como λ es arbitrario en $\mathbb{R}^+ \setminus \{0\}$ tenemos que $y \in A_1$. La prueba este completa. ■

La proposición previa nos permite dar la siguiente definición:

Una dirección asintótica de una función f es una dirección de un conjunto no vacío $L_{f, \mathbb{R}^n}(\alpha)$.

Decimos que la función f no tiene direcciones asintóticas si el cono asintótico de todos los conjuntos no vacíos de f es igual a $\{0\}$.

Proposición II.11 *Sea una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa, que no tiene direcciones asintóticas. Entonces el ínfimo de f es finito.*

Demostración.-

Si f es continua, es suficiente para saber que los conjuntos de nivel no vacíos de f son compactos.

Por hipótesis f no tiene direcciones asintóticas o equivalentemente, el cono asintótico común de los conjuntos de nivel no vacíos de f es igual a $\{0\}$. Sabemos que cada conjunto de nivel está acotada. Desde que los conjuntos de nivel son cerrados, la prueba está completa. ■

Proposición II.12 *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa, que no tiene direcciones asintóticas. Sea $(x_i)_{i \in \mathbb{N}}$, una sucesión de puntos de \mathbb{R}^n tal que $\lim_{i \rightarrow \infty} f(x_i) = \inf f$, entonces:*

(i) *La sucesión $(x_i)_{i \in \mathbb{N}}$ es acotada. Cualquier punto de acumulación de esta sucesión es un minimizador de f .*

(ii) *Si f tiene un único minimizador x , entonces la sucesión $(x_i)_{i \in \mathbb{N}}$ converge para x .*

Demostración.-

(i) Para algún $\epsilon > 0$ sea $\alpha = \inf f + \epsilon$

Para cualquier i bastante grande, $f(x_i) \leq \alpha$ o equivalentemente $x_i \in L_{f, \mathbb{R}^n}(\alpha)$. Sino, por la suposición, f no tiene direcciones asintóticas y consecuentemente, $L_{f, \mathbb{R}^n}(\alpha)$ esta acotado. Por lo tanto la sucesión $(x_i)_{i \in \mathbb{N}}$ esta acotada.

Si x^* es un punto de acumulación de esta sucesión entonces $x^* = \lim_{k \rightarrow \infty} x_{i_k}$ para una subsucesión $(x_{i_k})_{k \in \mathbb{N}}$. Por la continuidad,

$$f(x^*) = \lim_{k \rightarrow \infty} f(x_{i_k}) = \lim_{i \rightarrow \infty} f(x_i) = \inf f$$

(ii) La sucesión está acotada y este minimizador es el único punto de acumulación. Por eso la sucesión entera $(x_i)_{i \in \mathbb{N}}$ converge a este punto. La prueba esta completa.

■

Comentario Note que si f tiene un minimizador, entonces f no tiene direcciones asintóticas por que el conjunto de nivel es unitario, por lo tanto compacto. ■

Proposición II.13 Sea $f, \{f_i\}_{i \in I} : \mathbb{R}^n \rightarrow \mathbb{R}$ funciones convexas, donde I es un conjunto arbitrario de índices.

Asumamos que alli existe $x_0 \in \mathbb{R}^n$ que satisfaga $f_i(x_0) \leq 0, \forall i \in I$. Asumamos que no hay direcciones asintóticas comunes entre las funciones f y $\{f_i\}_{i \in I}$. Entonces el ínfimo de f bajo las restricciones $f_i(x_0) \leq 0, \forall i \in I$ es alcanzado.

Demostración.-

Sea $C = \{x \in \mathbb{R}^n / f_i(x) \leq 0, \forall i \in I\}$.

El conjunto C es convexo y cerrado. Es no vacío por la hipótesis.

Sea $A = \{x \in \mathbb{R}^n / x \in C \wedge f(x) \leq f(x_0)\}$.

Note que A es convexo, cerrado y no vacío ($x_0 \in A$) una dirección asintótica de A es una dirección asintótica común de f y además de los $\{f_i\}_{i \in I}$. Por la hipótesis, esta dirección asintótica común es $\{0\}$.

Por lo tanto A está acotado. La prueba está completa. ■

2.5. Método de punto proximal

Consideremos problema:

$$(P) \begin{cases} \min & f(x) \\ \text{s. a.} & x \in \mathbb{R}^n. \end{cases}$$

Donde $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función propia, semicontinua inferior y \mathbb{R}^n es el espacio euclidiano con norma $\|\cdot\|$. El método del punto proximal (MPP) fue introducido por Martinet en 1970 para problemas de optimización convexa y posteriormente por Rockafellar para encontrar ceros de operadores monótonos maximales. El algoritmo punto proximal genera la sucesión $x^k \subset \mathbb{R}^n$ de la siguiente manera:

$$x^0 \in \mathbb{R}^n \tag{2.5.1}$$

$$x^{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) + \lambda_k \|x - x^k\|^2\} \tag{2.5.2}$$

donde λ_k es un número real que cumple

$$0 < \lambda \leq \bar{\lambda} \tag{2.5.3}$$

para algunos $\bar{\lambda} > 0$ (incluye los casos de λ_k constante), y $\|\cdot\|$. Se puede mostrar que la sucesión generada por 2.5.1 y 2.5.2 converge para un minimizador de f . seguimos aquí un enfoque basado en la noción de la convergencia Fejér

Definición II.31 una sucesión $\{y^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$, es llamado Fejér convergente a un conjunto $U \subseteq \mathbb{R}^n$, respecto a la norma euclidiana si:

$$\|y^{k+1} - u\| \leq \|y^k - u\|; \forall u \in U, \forall k \geq 0.$$

Definición II.32 una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ se dice que es coerciva o coercitiva si

$$\lim_{\|x\| \rightarrow +\infty} g(x) = +\infty$$

Proposición II.14 Si la sucesión $\{y^k\}_{k \in \mathbb{N}}$ es Fejér convergente en un conjunto $U \neq \emptyset$, entonces $\{y^k\}_{k \in \mathbb{N}}$ es acotada. Si y es un punto en la frontera de la sucesión a lo largo de U entonces,

$$\lim_{k \rightarrow +\infty} y^k = \bar{y}.$$

Demostración.-

como $\{y^k\}_{k \in \mathbb{N}}$ es Fejér convergente, por la definición se tiene que:

$$\|y^{k+1} - u\| \leq \|y^k - u\|, \forall k \geq 0, \forall u \in U,$$

esto implica que:

$$\|y^{k+1} - u\| \leq \|y^0 - u\|, \forall k \geq 0, \forall u \in U.$$

Luego, se tiene que $\{y^k\}_{k \in \mathbb{N}}$ está contenida en una bola de centro u y radio $\|y^0 - u\|$, así $\{y^k\}_{k \in \mathbb{N}}$ esta acotada. Para la segunda parte, como y está en la frontera, entonces existe un subsucesión $\{y^{f_k}\}_{f_k \in \mathbb{N}}$ de $\{y^k\}_{k \in \mathbb{N}}$ que converge a y . Por otra parte, como $\{y^k\}_{k \in \mathbb{N}}$ es Fejér convergente y como y está en U entonces $\{\|y^k - y\|\}$ es decreciente y no negativa, así se tiene que la sucesión $\{\|y_k^f - y\|\}$ converge a cero. Entonces el conjunto de subsucesiones converge a cero. Así $\lim_{k \rightarrow \infty} \|y^k - y\| = 0$ concluyendo que $\lim_{k \rightarrow \infty} y^k = y$. ■

Teorema II.17 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa y continuamente diferenciable. Supongamos que el conjunto de minimizadores de f en \mathbb{R}^n es no vacío. Entonces la sucesión $\{x^k\}$ generada por:

$$x^0 \in \mathbb{R}^n \tag{2.5.4}$$

$$x^{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) + \lambda_k \|x - x^k\|^2\} \tag{2.5.5}$$

converge a un punto $x^* \in U$.

Donde λ_k son números reales que satisfacen:

$$0 < \lambda_k \leq \bar{\lambda} \forall k \in \mathbb{N} \tag{2.5.6}$$

para algún $\bar{\lambda} > 0$ y $\|\cdot\|$ es la norma euclidiana.

Demostración.-

La demostración se divide en 4 pasos. El paso 1 se probará que $\{x^k\}$ está bien definida. En segunda instancia veremos que $\{x^k\}$ es Fejér convergente a U . En el tercer paso se establecerá que el factor $\lim_{k \rightarrow \infty} (x^{k+1} - x^k) = 0$, la cual se usará en el paso 4, donde se probará que el conjunto de puntos de la sucesión $\{x^k\}$ pertenece a U . De esta manera con los pasos desde el 2 hasta el 4, junto con la proposición

II.14, se obtiene la tesis del teorema.

Paso 1.- La sucesión esta bien definida.

En efecto, considere la recurrencia $f_k = f(x) + \lambda_k \|x - x^k\|^2$. Así f alcanza un mínimo, pues esta es acotada inferiormente, además que $\lim_{\|x\| \rightarrow \infty} f_k(x) = \infty$. Luego como f es continua y la minimización en 2.5.5 reduce el problema a un conjunto compacto, se garantiza así la solución y por lo tanto f tiene un mínimo. Por otra parte, como f es convexa además con $\lambda_k \|x - x^k\|^2$ es estrictamente convexa, entonces f_k es estrictamente convexa de esta manera se garantiza que el mínimo es único y que $\{x^{k+1}\}$ son únicos, y por lo tanto, la sucesión está bien definida.

Paso 2.- Veamos que $\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2$, $\forall k \geq 0$ y todo $\bar{x} \in U$.

$$\begin{aligned} \|x^k - \bar{x}\|^2 &= \|x^k - x^{k+1} + x^{k+1} - \bar{x}\|^2 \\ &= \|x^k - x^{k+1}\|^2 + \|x^{k+1} - \bar{x}\|^2 + 2 \langle x^k - x^{k+1}, x^{k+1} - \bar{x} \rangle \end{aligned} \quad (2.5.7)$$

Luego, como x^{k+1} resuelve 2.5.5, se tiene:

$$\begin{aligned} 0 &= \nabla f_k(x^{k+1}) \\ &= \nabla f(x^{k+1}) + 2\lambda_k(x^{k+1} - x^k), \text{ de donde} \\ 2(x^{k+1} - x^k) &= -\frac{1}{\lambda_k} \nabla f(x^{k+1}). \end{aligned} \quad (2.5.8)$$

Así de 2.5.7 y 2.5.8, se tiene que:

$$\begin{aligned} \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2 - \|x^{k+1} - \bar{x}\|^2 &= 2 \langle x^k - x^{k+1}, x^{k+1} - \bar{x} \rangle \\ &= \frac{1}{\lambda_k} \langle \nabla f(x^{k+1}), x^{k+1} - \bar{x} \rangle \\ &\geq \frac{1}{\lambda} [f(x^{k+1}) - f(\bar{x})] \geq 0. \end{aligned} \quad (2.5.9)$$

Dado que \bar{x} es el minimizador de f . Por lo tanto, de la última desigualdad 2.5.9, obtenemos lo deseado.

Paso 3.- Veamos ahora que:

$$\lim_{k \rightarrow \infty} (x^{k+1} - x^k) = 0. \quad (2.5.10)$$

Del caso anterior

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2$$

$$\|x^k - \bar{x}\|^2,$$

Así, $\{\|x^k - \bar{x}\|\}$ es decreciente, y es acotada pues al ser decreciente

$$0 \leq \|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 \leq \dots \leq \|x^0 - \bar{x}\|^2,$$

por tanto tenemos que converge.

para cada $k \geq 0$

$$\|x^{k+1} - x^k\|^2 = \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2,$$

y así

$$\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\|^2 = \lim_{k \rightarrow +\infty} (\|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2) = 0.$$

Luego,

$$\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$$

Paso 4.- $\{x^k\}$ tiene puntos frontera y todo ellos pertenecen a U .

La existencia de los puntos fronteras están garantizando el paso 2, siendo más explícitos, se puede observar que $\{x^k\}$ es Fejér convergente a un conjunto U y por la primera parte de la proposición II.14 tenemos que $\{x^k\}$ es acotada y así queda garantizada la existencia de puntos frontera de la sucesión. Sea \bar{x} un punto clausura de la sucesión $\{x^k\}$ y tomemos $\{x^{f_k}\}$ una subsucesión de $\{x^k\}$, tal que $\lim_{k \rightarrow \infty} x^{f_k} = \bar{x}$. por 2.5.8

$$\nabla f(x^{f_{k+1}}) = 2\lambda_{j_k}(x^{f_k} - x^{f_{k+1}}) \quad (2.5.11)$$

Luego por el caso 3 $\lim_{k \rightarrow \infty} x^{f_{k+1}} = \lim_{k \rightarrow \infty} x^{f_k} = \bar{x}$. Así, tomando límite en la igualdad 2.5.11 cuando $k \rightarrow \infty$, usando $\lambda_k \leq \bar{\lambda}$ y además que f es continuamente diferenciable, tendríamos $\nabla f(\bar{x}) = 0$, por convexidad de f , $\bar{x} \in U$.

De esta forma por los casos 2; 4 y la proposición II.14 segunda parte, existe $x^* \in U$ tal que

$$x^* = \lim_{k \rightarrow \infty} x^k. \quad \blacksquare$$

Este teorema es la versión clásica de la convergencia del Método de Punto Proximal. Inicialmente, en 1965, Moreau [18] ofrece esta regularización, concretamente

propone la siguiente función:

$$F(y) = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{1}{2} \|x - y\|^2 \right\},$$

que luego es modificada por Yosida, agregando el parámetro λ convirtiéndose en la regularización Moreau -Yosida, conocida por la expresión siguiente:

$$F_\lambda(y) = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{1}{2} \lambda \|x - y\|^2 \right\}.$$

Observe que el Método de punto proximal clásico sustituye el parámetro λ por una sucesión $\{\lambda_k\}$ acotadas y de números positivos.

Nuestro trabajo de investigación tiene el siguiente problema de interés:

$$(P) \begin{cases} \min & f(x) \\ \text{s. a.} & x \geq 0. \end{cases}$$

Donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $x \in \mathbb{R}_{++}^n$. Daremos una motivación de problemas que pueden ser formuladas en la forma (P)

2.5.1. Motivación

Problemas de la forma (P) se pueden obtener de diversas aplicaciones. Presentaremos un ejemplo relacionado al problema de control óptimo.

Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto y acotado. Considere el siguiente problema de optimización, denominado “el problema de control restringido”.

$$(P1) \begin{cases} \min & J(u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u - u_d\|_{L^2(\Omega)}^2 \\ \text{s. a.} & u \in F, \end{cases}$$

donde, $\Psi, y_d, u_d \in L^2(\Omega)$, son funciones dadas, $F = \{u \in L^2(\Omega) : u(\bar{x}) \leq \Psi(\bar{x}), \bar{x} \in \Omega\}$, y $y = y(u) \in H_0^1(\Omega)$, es la solución débil de la ecuación de Poisson

$$-\Delta y = u.$$

La función y se denomina “estado” y u es la “variable de control”. El objetivo es minimizar la distancia al estado de control deseado y_d (por ello el subíndice d), sujeto a algunas restricciones sobre la variable de control. El término de penalidad

en la función objetivo, $\frac{\alpha}{2} \|u - u_d\|_{L^2(\Omega)}^2$, es una regularización estándar donde $\alpha > 0$, es el parámetro de regularización. Tomaremos para el caso bidimensional $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$, en el problema de minimización (P1). Así tenemos

$$(P2) \begin{cases} \min & J(u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u - u_d\|_{L^2(\Omega)}^2 \\ \text{s.a} & -\Delta y = u \\ & u \leq \Psi. \end{cases}$$

Aproximaremos la solución de la ecuación de Poisson mediante el método de diferencias finitas usando el esquema de los 5-puntos para la aproximación del laplaciano.

Discretizaremos el dominio $\Omega = (0, 1) \times (0, 1)$, para esto subdividimos a Ω en cuadrados de lados iguales con puntos $\{(x_i, y_j) = (i_h, j_h) : i, j = 0, 1, 2, \dots, N, N \in \mathbb{N}\}$ con $h = \frac{1}{N}$, para algún $N \in \mathbb{N}$.

Donde se tendrá el lado de cada cuadrado

$$\Delta x = \Delta y = \frac{1 - 0}{N}.$$

La aproximación de 5-puntos del laplaciano es:

$$\Delta y(x_i, y_j) \approx \frac{y(x_i, y_{j-1}) + y(x_{i-1}, y_j) - 4y(x_i, y_j) + y(x_{i+1}, y_j) + y(x_i, y_{j+1}))}{\Delta x \cdot \Delta y}.$$

Denotemos, $y(x_i, y_j) = y_{ij}$ y $u(x_i, y_j) = u_{ij}$

Reemplazando en la ecuación de Poisson obtenemos siguiente sistema:

$$y_{i,j-1} + y_{i-1,j} - 4y_{ij} + y_{i+1,j} + y_{i,j+1} = \frac{-1}{N^2} u_{ij}.$$

Expresando matricialmente:

$$Ay = -\frac{1}{N^2}u - g, \tag{2.5.12}$$

donde:

$$A = \begin{pmatrix} B & I & 0 & \cdots & 0 \\ I & B & I & & \vdots \\ \vdots & \ddots & \ddots & \ddots & I \\ 0 & \cdots & 0 & I & B \end{pmatrix}_{(N-1)^2 \times (N-1)^2} \quad B = \begin{pmatrix} -4 & 1 & 0 & \cdots & 0 \\ 1 & -4 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & -4 \end{pmatrix}_{(N-1) \times (N-1)}$$

$y = (y_{ij})$, $u = (u_{ij})$, $g \in M^{(N-1)^2 \times 1}$, es la matriz columna cuyos elementos son los puntos de la función y , evaluados en la frontera de Ω . Por hipótesis, $y \in H_0^1(\Omega)$, es decir, $y(\bar{x}) = 0$, $\forall \bar{x} \in \partial(\Omega)$ de esto sigue que $g = 0$ (matriz columna nula). Luego en (1) tenemos $Ay = -\frac{1}{N^2}u$, esto es:

$$N^2 Ay = -u. \quad (2.5.13)$$

Pero se sabe que $\Psi(\bar{x}) \geq u(\bar{x})$, $\forall \bar{x} \in \Omega$. Así

$$\Psi(x_i, y_j) \geq u(x_i, y_j) \quad \forall i, j = 1 \dots n.$$

Denotando $\Psi(x_i, y_j) = \Psi_{ij}$, se tiene $\Psi_{ij} \geq u_{ij}$. Sea $\Psi \in \mathbb{M}^{(n-1)^2 \times 1}$ cuyas componentes son Ψ_{ij} entonces, $\Psi - u \geq 0$ y reemplazando (2) tenemos finalmente la ecuación de Poisson discretizada

$$\Psi + N^2 Ay \geq 0. \quad (2.5.14)$$

Veamos un caso particular donde

$$y_d = \frac{1}{6} \text{sen}(2\pi x) \text{sen}(2\pi y) e^{2x}$$

$$u_d = 0$$

$$\Psi = 0$$

$$\alpha = 10^{-2}.$$

Reemplazando en (P2) obtenemos

$$(P3) \begin{cases} \min & J(u) = \frac{1}{2} \int_{\Omega} (y(x, y) - \frac{1}{6} \text{sen}(2\pi x) \text{sen}(2\pi y) e^{2x})^2 dx dy + \frac{\alpha}{2} \int_{\Omega} u(x, y)^2 dx dy \\ \text{s. a:} & -\Delta y = u \\ & u \leq 0. \end{cases}$$

Aproximando las integrales dobles mediante sumatorias se tiene,

$$\int_{\Omega} y(\bar{x}) d\bar{x} = \int_0^1 \int_0^1 y(x, y) dx dy \approx \sum_{i=1}^n \sum_{j=1}^n y_{ij} \Delta x \Delta y.$$

De la discretización de la ecuación de Poisson, $z = -u = N^2 Ay$, se tiene $y = zN^{-2}A^{-1}$, y reemplazando la aproximación de la integral por sumatorias en (P3), obtenemos el siguiente problema de optimización.

$$(P5) \begin{cases} \min & \sum_{i=1}^n \sum_{j=1}^n [\frac{1}{N^2}(a_{ij})z_{ij} - (y_{ij})_d]^2 \Delta x \Delta y + \frac{\alpha}{2} \sum_{i=1}^n \sum_{j=1}^n z_{ij}^2 \Delta x \Delta y \\ \text{s. a :} & z \geq 0, \end{cases}$$

donde esta última función objetivo depende únicamente de z , lo cual se puede escribir como el problema de optimización deseado, esto es:

$$(P) \begin{cases} \min & f(z) \\ \text{s. a :} & z \geq 0. \end{cases}$$

2.5.2. φ - divergencia

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

($\varphi 1$) φ es dos veces continuamente diferenciable en $\text{int}(\text{dom}\varphi) = (0, +\infty)$.

($\varphi 2$) φ es estrictamente convexa en su dominio.

($\varphi 3$) $\lim_{t \rightarrow 0^+} \varphi'(t) = -\infty$.

($\varphi 4$) $\varphi(1) = \varphi'(1) = 0$ y $\varphi''(1) > 0^+$.

Denotamos por Φ las clases de funciones que satisfacen ($\varphi 1$) – ($\varphi 4$).

Proposición II.15 sea $\varphi \in \Phi$

1. $\varphi(t) \geq 0$ y $\varphi(t) = 0$ si y solo si $t = 1$
2. $\varphi(t)$ es decreciente en $(0, 1)$ y creciente en $(1, +\infty)$
3. $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$

Demostración.-

Sea $\varphi \in \Phi$ cualquiera, por $(\varphi 4)$ se tiene que $\varphi(1) = 0$ es un mínimo local de φ y junto con $(\varphi 2)$ tenemos que $\varphi(1)$ es un mínimo global de φ , así de 1 y 2 se cumplen pues $\text{dom}(\varphi) = (0, +\infty]$. Por 2 y por $(\varphi 2)$ se obtiene 3. ■

Introduciremos dos sub clases de núcleo φ :

$$\Phi_1 := \{\varphi \in \Phi, \text{ existe } M > 0 \text{ tal que } \varphi''(t) \leq M, \forall t \geq 1\}. \quad (2.5.15)$$

$$\Phi_2 := \left\{ \varphi \in \Phi; \varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t - 1), \forall t > 0 \right\}. \quad (2.5.16)$$

Veamos tres funciones para estas clases de núcleos (que pertenezcan a $\Phi_1 \cap \Phi_2$.)

Ejemplo II.1 Sea

$$\varphi_1(t) = -\log t + t - 1, \quad \text{dom}\varphi = (0, +\infty).$$

véase la figura

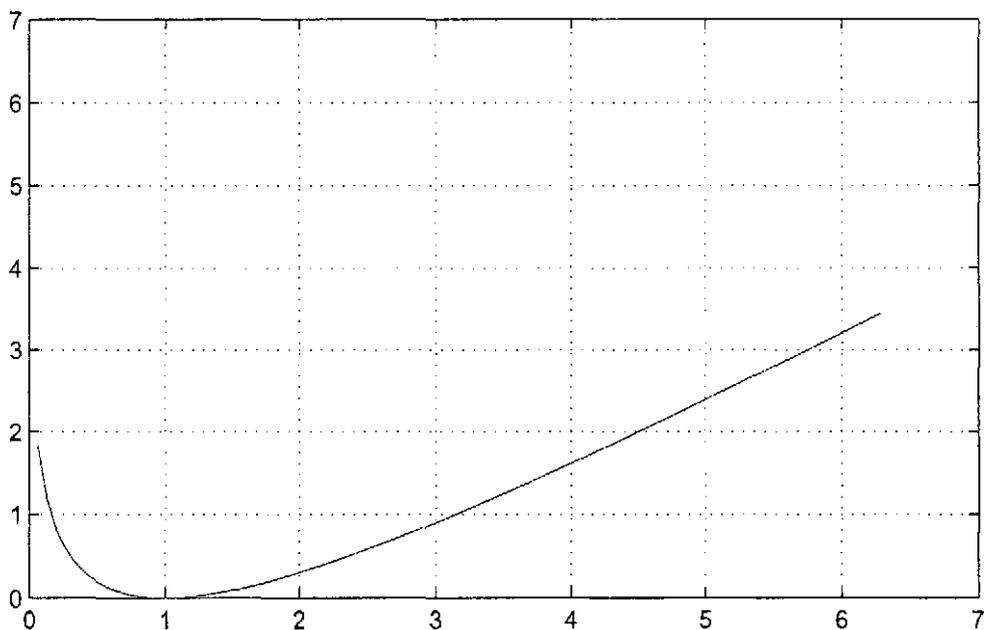


Figura II.1: Figura 1

Se verifica que $\varphi_1 \in \Phi$. En efecto,

$$\begin{aligned}\varphi_1'(t) &= 1 - \frac{1}{t} \\ \varphi_1''(t) &= \frac{1}{t^2}.\end{aligned}$$

Así tenemos que

- i φ es dos veces continuamente diferenciable en $(0, +\infty)$.
- ii $\varphi_1''(t) > 0$, $\forall t \in \text{dom}\varphi_1$, φ_1 es estrictamente convexa.
- iii $\lim_{t \rightarrow 0} \varphi_1'(t) = \lim_{t \rightarrow 0} (1 - \frac{1}{t}) = -\infty$.
- iv $\varphi_1(t) = 0 = \varphi_1'(t)$, y $\varphi_1''(t) > 0$.

Por lo tanto, $\varphi_1 \in \Phi$.

Veamos ahora que $\varphi_1 \in \Phi_1 \cap \Phi_2$.

En efecto,

$$\begin{aligned}t &\geq 1 \\ \frac{1}{t^2} &\leq 1 \\ \varphi_1''(t) &\leq 1, \text{ de esto sigue:}\end{aligned}$$

$\varphi \in \Phi$, existe $M = 1 > 0$, tal que $\varphi_1''(t) \leq 1$, $\forall t \geq 1$. Esto es $\varphi_1 \in \Phi_1$.

Desde que $\varphi_1''(1) = 1$, tenemos

$$\varphi_1''(1)(1 - \frac{1}{t}) \leq (1 - \frac{1}{t}), \quad \forall t > 0. \quad (2.5.17a)$$

Así mismo para todo $t > 0$ tomemos dos casos,

primer caso

$$\begin{aligned}t \leq 1 &\Rightarrow (t - 1) \leq 0 \\ 1 &\leq \frac{1}{t}.\end{aligned}$$

De ambas desigualdades tenemos,

$(t - 1)1 \geq \frac{1}{t}(t - 1)$ de esto queda:

$$1 - \frac{1}{t} \leq t - 1.$$

Segundo caso $t > 1$. de manera similar:

$$\left(1 - \frac{1}{t}\right) \leq \varphi_1''(1)(t - 1), \quad \forall t > 0. \quad (2.5.17b)$$

De 2.5.17a, 2.5.17b y $\varphi_1 \in \Phi$. Finalmente tenemos:

$$\varphi_1 \in \Phi, \quad \varphi_1''(1)\left(1 - \frac{1}{t}\right) \leq \left(1 - \frac{1}{t}\right) \leq \varphi_1''(1)(t - 1), \quad \forall t > 0.$$

Así $\varphi \in \Phi_2$. de esto se sigue que:

$$\varphi_1 \in \Phi_1 \cap \Phi_2.$$

Ejemplo II.2 considere

$$\varphi_2(t) = t \log t - t + 1, \quad \text{dom} \varphi_2 = [0, +\infty).$$

véase la figura

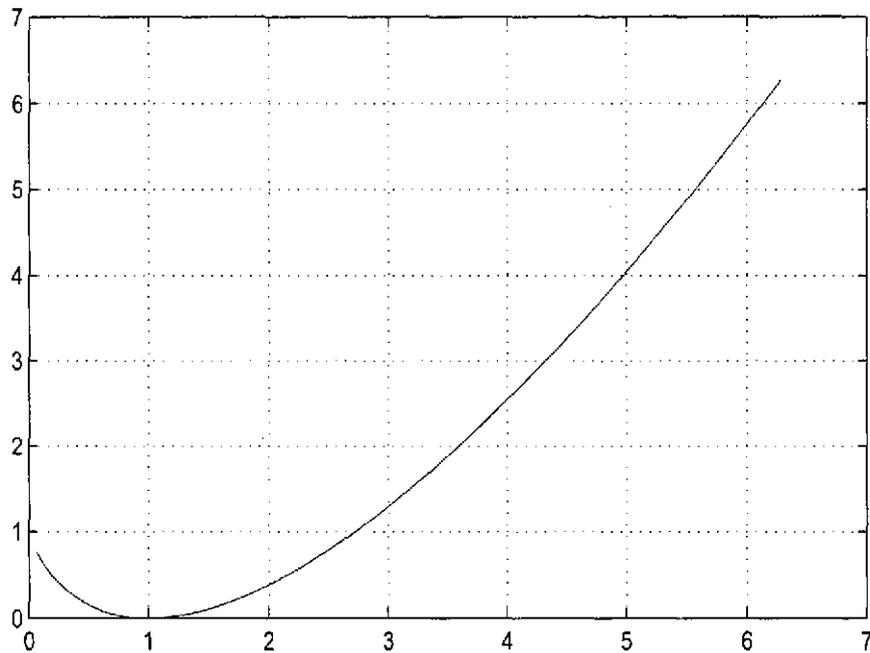


Figura II.2: Figura 2

Se verifica que $\varphi_2(t) \in \Phi$. En efecto,

(i) $\varphi_2'(t) = \log t$ y $\varphi_2''(t) = \frac{1}{t}$. Así φ_2 es dos veces continuamente diferenciable en $(0, +\infty)$.

(ii) $\varphi_2''(t) > 0$, $\forall t \in \text{int}(\text{dom}\varphi_2)$; φ_2 es estrictamente convexa en $(0, +\infty)$.

Probaremos que φ_2 es estrictamente convexa en $t_1 = 0$ y para $t_2 \neq 0$, es decir $\varphi_2(\lambda t_2) < (1 - \lambda)\varphi_2(0) + \lambda\varphi_2(t_2)$ con $0 < \lambda < 1$.

$$\begin{aligned} \lambda < 1 &\implies \lambda t_2 < t_2, \\ (\lambda t_2) \log(\lambda t_2) - (\lambda t_2) + 1 &< \lambda t_2 \log t_2 - \lambda t_2 + 1 + \lambda - \lambda, \\ \varphi_2(\lambda t_2) &< \lambda(t_2 \log t_2 - t_2 + 1) + (1 - \lambda), \\ \varphi_2(\lambda t_2) &< \lambda\varphi_2(t_2) + (1 - \lambda)\varphi_2(0). \end{aligned}$$

Análogamente podemos probar que, φ es estrictamente convexa para $t_2 = 0$ y $t_1 \neq 0$. Así se tiene que, φ_2 es estrictamente convexa en su dominio.

(iii) $\lim_{t \rightarrow 0} \varphi_2'(t) = \lim_{t \rightarrow 0^+} (\log(t)) = -\infty$.

(iv) $\varphi_2(1) = 0 = \varphi_2'(1)$ y $\varphi_2''(1) > 0$.

De esto sigue que: $\varphi_2 \in \Phi$.

Ahora probaremos que $\varphi_2 \in \Phi_1 \cap \Phi_2$.

En efecto, para todo $t \geq 1$ se tiene $\varphi_2''(t) = \frac{1}{t} \leq 1 = M$, de esto se sigue que: $\varphi_2 \in \Phi_1$.

Veamos ahora que $\varphi_2 \in \Phi_2$, $\forall t > 0$. Se presentará los siguientes casos: cuando $0 < t \leq 1$ y para $1 \geq t$

(a) Para $0 < t \leq 1$, se tiene que $1 \leq \frac{1}{t}$ entonces por Teorema del Valor Medio, existe una constante M tal que

$$\log \frac{1}{t} = \frac{1}{M} \left(\frac{1}{t} - 1 \right) \text{ con, } 1 < M < \frac{1}{t}.$$

De esto sigue, $\log t = \frac{1}{M} \left(1 - \frac{1}{t} \right)$. Como $1 < M < \frac{1}{t}$, entonces de $1 - \frac{1}{t} < 0$, se tiene

$$\left(1 - \frac{1}{t} \right) > M \left(1 - \frac{1}{t} \right),$$

en consecuencia

$$\log t > \left(1 - \frac{1}{t} \right). \tag{2.5.18a}$$

De manera similar, para $t - 1 < 0$ sigue

$$(t - 1)M > \frac{1}{t}(t - 1),$$

de donde se tiene

$$(t - 1) > \log t. \quad (2.5.18b)$$

De 2.5.18a y 2.5.18b se sigue $(1 - \frac{1}{t}) < \log(t) < (t - 1)$, $0 < t \leq 1$. (b) Para $t > 1$, por Teorema del Valor Medio $\exists M > 0$ tal que $\log t = \frac{1}{M}(t - 1)$ con $1 < M < t$. Para $1 - \frac{1}{t} > 0$, se tiene $\log t > (1 - \frac{1}{t})$, También, con $t - 1 > 0$ se tiene $(t - 1) > \log t$. De estas dos últimas desigualdades:

$$(1 - \frac{1}{t}) < \log t < (t - 1), \quad t > 1.$$

Entonces de (a) y (b) se sigue que, $\varphi_2 \in \Phi_2$.

Así se tiene que $\varphi_2 \in \Phi_1 \cap \Phi_2$.

Ejemplo II.3 sea $\varphi_3(t) = 2(\sqrt{t} - 1)^2$, $\text{dom}\varphi = [0, +\infty)$. véase la figura

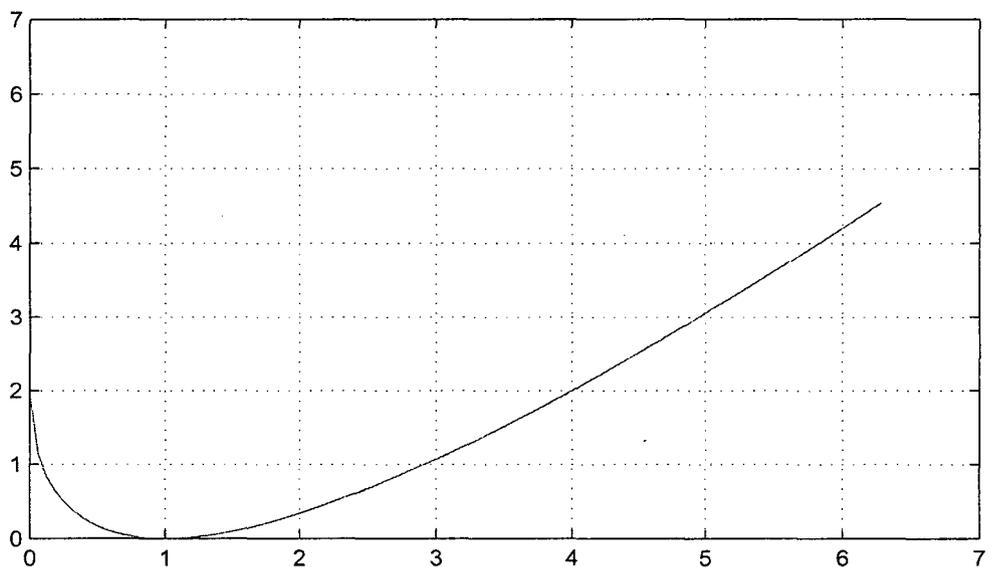


Figura II.3: Figura 3

1. $\varphi_3'(t) = 2\left(1 - \frac{1}{\sqrt{t}}\right)$ y $\varphi_3''(t) = \frac{1}{t^{3/2}}$. Así φ_3 es dos veces continuamente diferenciable en $(0, \infty)$.

2. $\varphi_3''(t) > 0$, $\forall t \in \text{int}(\text{dom}\varphi_3)$, φ es estrictamente convexa en $(0, +\infty)$. Probaremos que φ es estrictamente convexa en $t_1 = 0$ y para $t_2 \neq 0$, es decir

$$\varphi(\lambda t_2) < (1 - \lambda)\varphi(0) + \lambda\varphi(t_2), \text{ con } 0 < \lambda < 1.$$

$$\lambda < 1 \Rightarrow \lambda t_2 < 1t_2,$$

$$\sqrt{\lambda t_2} < \sqrt{t_2},$$

$$1 - \lambda + \lambda t_2 - 2\lambda\sqrt{t_2} + \lambda > 1 + \lambda t_2 - 2\sqrt{\lambda t_2},$$

$$2(1 - \lambda) + 2\lambda(\sqrt{t_2} - 1)^2 > 2(\sqrt{\lambda t_2} - 1)^2,$$

$$(1 - \lambda)\varphi(0) + \lambda\varphi(t_2) > \varphi(\lambda t_2).$$

Análogamente, φ es estrictamente convexa en $t_2 = 0$ y $t_1 \neq 0$. Así φ_3 es estrictamente convexa en su dominio.

$$3. \lim_{t \rightarrow 0} \varphi_3'(t) = \lim_{t \rightarrow 0^+} \left(1 - \frac{1}{\sqrt{t}}\right) = -\infty.$$

$$4. \varphi_3(t) = 0 = \varphi_3'(t) \text{ y } \varphi_3''(t) > 0.$$

Tenemos $\varphi_3 \in \Phi$. Probaremos que $\varphi_3 \in \Phi_1 \cap \Phi_2$. En efecto, para todo $t \geq 1$, se puede notar que $\varphi_3 \in \Phi_1$.

Vamos a demostrar que $\varphi_3 \in \Phi_2$. Dado $\varphi_3''(1) = 1$, tenemos que probar $\forall t > 0$; $(1 - 1/t) \leq \varphi_3'(t) \leq (t - 1)$.

(a) Encontraremos la 1ra desigualdad, es decir,

$$\forall t > 0; (1 - 1/t) \leq \varphi_3'(t).$$

$$0 < t \leq 1, 0 < \sqrt{t} \leq 1, \implies 1 - \sqrt{t} \leq 0,$$

$$2(1 - 1/\sqrt{t}) \geq (1 + 1/\sqrt{t})(1 - 1/\sqrt{t}),$$

$$2(1 - 1/\sqrt{t}) \geq (1 - 1/t).$$

Cuando

$$t \geq 1, 1 \geq 1/\sqrt{t}, 1 - 1/\sqrt{t} \geq 0,$$

$$2 > 1 + 1/\sqrt{t},$$

$$2(1 - 1/\sqrt{t}) \geq (1 - 1/t).$$

Entonces, $\forall t > 0$, $2(1 - 1/\sqrt{t}) \geq (1 - 1/t)$.

(b) Ahora la 2da desigualdad es decir, $\forall t > 0$; $(t - 1) \geq \varphi_3'(t)$

$$\begin{aligned}
 0 < t \leq 1, \quad \sqrt{t} \leq 1 \quad \sqrt{t} \leq 1 \\
 \implies \quad 1 - \sqrt{t} \leq 0 \\
 \sqrt{t} - 1 \geq (1 - 1/\sqrt{t}) \\
 \text{usando } \sqrt{t} - 1 \leq 0 \text{ y } \sqrt{t} + 1 \leq 2 \\
 t - 1 \geq (1 - 1/\sqrt{t}).
 \end{aligned}$$

Entonces, $\forall t > 0$, $2(1 - 1/\sqrt{t}) \leq (t - 1)$.

Con lo cual concluye la prueba de los ejemplos.

Definición II.33 Si $\varphi \in \Phi$, entonces la aplicación $d : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$, definida por

$$d_\varphi(x, y) = \sum_{j=1}^n y_j \varphi\left(\frac{x_j}{y_j}\right),$$

se dice que es una φ -divergencia

Proposición II.16 Son ciertas las afirmaciones:

1. $d_\varphi(x, y) \geq 0 \quad \forall (x, y) \in \mathbb{R}_{++}^n$
2. $d_\varphi(x, y) = 0$ si y solo si $x = y$
3. El conjunto de nivel de $d_\varphi(\cdot, y)$ están acotados para todo $y \in \mathbb{R}_{++}^n$
4. El conjunto de nivel de $d_\varphi(x, \cdot)$ están acotados para todo $x \in \mathbb{R}_{++}^n$
5. $d_\varphi(x, y)$ es conjuntamente convexa en x, y y estrictamente convexa en x
6. $\lim_{k \rightarrow \infty} d_\varphi(y, y^k) = 0$ si y solo si $\lim_{k \rightarrow \infty} y^k = y$.

De los ejemplos anteriores se obtiene las distancias respectivas

$$\varphi_1(t) = t - \log t - 1, \text{ entonces } d_{\varphi_1}(x, y) = d_{\varphi_1}(x, y)$$

$$\varphi_2(t) = t \log t - t + 1, \text{ entonces } d_{\varphi_2}(x, y) = \sum_{j=1}^n \left(x_j \log \frac{x_j}{y_j} + y_j - x_j \right)$$

$$\varphi_3(t) = (\sqrt{t} - 1)^2, \text{ entonces } d_{\varphi_3}(x, y) = \sum_{j=1}^n (\sqrt{x_j} - \sqrt{y_j})^2$$

2.5.3. Método de punto proximal con φ -divergencia

El problema de nuestro interés en

$$(P) \begin{cases} \min & f(x) \\ \text{s. a :} & x \geq 0. \end{cases} \quad (2.5.19)$$

con $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa, el método de punto proximal con distancia φ divergencia para problemas 2.5.19 genera la sucesión $x^k \subset \mathbb{R}^n$ dado por

$$x^0 \quad (2.5.20)$$

$$x^{k+1} = \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) + \lambda_k d_\varphi(x, x^k)\} \quad (2.5.21)$$

con λ_k satisface $0 < \lambda_k \leq \bar{\lambda}$ para algunos $\bar{\lambda} > 0$. La condición de optimalidad para 2.5.21, afirmamos que $u_k \in \partial f(x^{k+1})$ con $u_j^k = -\lambda_k \varphi' \left(\frac{x_j^{k+1}}{x_j^k} \right)$. Si f es diferenciable, x^{k+1} es la solución x del sistema

$$\nabla f(x)_j + \lambda \varphi' \left(\frac{x_j}{x_j^k} \right) = 0 \quad (2.5.22)$$

2.5.22 es un sistema no lineal de n ecuaciones con n variables x_1, \dots, x_n . cuando 2.5.19 tiene solución, el problema de minimización 2.5.21 se reduce a subconjuntos compactos ($\lim_{t \rightarrow \infty} \varphi(t) = \infty$) por las condiciones impuestas sobre φ y así se garantiza la existencia de x^{k+1} . La unicidad de x^{k+1} , sigue de la convexidad de f y la proposición (II.16)5, lo que permite probar que está bien definida la sucesión, sin embargo, la prueba de convergencia es mas dura Véase [9], debido a que la sucesión no es Fejér convergente, extendemos la noción de Fejér convergente a casi-Fejér convergente.

Definición II.34 la sucesión $\{y^k\} \subset \mathbb{R}_{++}^n$, es casi-Fejér convergente al conjunto $U \subset \mathbb{R}_{++}^n$, con respecto a la φ divergencia d_φ , si para cada $u \in U$; existe una sucesión $\{\epsilon_k\} \subset \mathbb{R}_{++}$ tal que: $\sum_{k=0}^{+\infty} \epsilon_k < +\infty$ y para todo $k \geq 0$,

$$d_\varphi(u, y^{k+1}) \leq d_\varphi(u, y^k) + \epsilon_k \quad (2.5.23)$$

Proposición II.17 si $\{y^k\} \subset \mathbb{R}_{++}^n$ es casi-Fejér convergente a $U \subset \mathbb{R}_{++}^n$, con respecto a la φ -divergencia d_φ , entonces $\{y^k\}$ es limitada. Si un punto clausura \bar{y} de $\{y^k\}$ pertenece a U , entonces $\bar{y} = \lim_{k \rightarrow \infty} y^k$

los resultados del método punto proximal y φ divergencia se pueden encontrar en [10]

Proposición II.18 *La sucesión generada por el algoritmo del punto proximal con φ divergencias está bien definida y contenida en \mathbb{R}_{++}^n . Más aún $u^k \in \partial f(x^{k+1})$ para todo $k \geq 0$.*

Proposición II.19 *Para todo $k \geq 0$ se tiene que: $0 \leq \lambda_k d_\varphi(x^{k+1}, x^k) \leq \lambda_k d_{\widehat{\varphi}}(x^k, x^k) \leq f(x^k) - f(x^{k+1})$ donde $\widehat{\varphi}(t) = (1-t)\varphi(t)$*

Corolario II.2 1. *la sucesión $\{f(x^k)\}$ es decreciente y convergente.*

$$2. \sum_{k=0}^{\infty} \lambda_k d_{\widehat{\varphi}}(x^{k+1}, x^k) < \infty$$

$$3. \lim_{k \rightarrow +\infty} x_j^k \varphi \left(\frac{x_j^{k+1}}{x_j^k} \right) = 0, \text{ para todo } j = 1, \dots, n$$

Proposición II.20 *Si la sucesión $\{x^k\}$ converge a x^* y si se cumple: (i) f es $C^1(\mathbb{R}_+^n)$ o (ii) $\exists z \in T$ tal que $I(x^*) \subset I(z)$ entonces $x^* \in T$; donde $T = \{x / x \text{ es solución óptima}\}$ y $I(x) = \{j \in \{1, \dots, n\} : x_j = 0\}$.*

Teorema II.18 *Si la sucesión $\{y^k\}$ es casi-Fejér convergente a un conjunto no vacío V , entonces $\{y^k\}$ es acotada y $\{d(v, y^k)\}$ es acotado para todo $v \in V$.*

El resultado final en [10] que demuestra la convergencia del método es el siguiente:

Teorema II.19 *si se cumple cualquiera de las siguientes condiciones:*

1. *la sucesión $\{x^k\}$ tiene puntos límite.*
2. *el conjunto T es acotado.*
3. *Existe solución z del problema primal 2.5.19 tal que $\{x_j^k\}$ es acotada, para aquellas componentes tales que $z_j > 0$.*
4. *$\varphi'(t) \leq \varphi''(1) \log t$ para todo $t > 0$.*
5. *x^0 es suficiente cercano a la solución z^* del problema 2.5.19.*

Entonces la sucesión $\{x^k\}$ generada por 2.5.21 converge. Más aún bajo cualquiera de las condiciones 3, 4, ó 5 o si f es $C^1(\mathbb{R}_+^n)$, cualquier punto límite de la sucesión es una solución del problema 2.5.19

2.5.4. Método punto proximal con distancia homogénea de orden dos

Ahora tomamos un $\varphi \in \Phi$ y definimos d_φ para $x, y \in \mathbb{R}_{++}^p$ por:

$$d_\varphi(x, y) = \sum_{j=1}^p y_j^2 \varphi\left(\frac{x_j}{y_j}\right). \quad (2.5.24)$$

La función d_φ satisface las siguientes propiedades

d_φ es una función homogénea de orden 2

$\forall (x, y) \in \mathbb{R}_{++}^p \times \mathbb{R}_{++}^p$ tenemos

$$\begin{aligned} d_\varphi(x, y) &\geq 0, \\ d_\varphi(x, y) &= 0 \quad \text{si y solo si} \quad x = y \end{aligned} \quad (2.5.25)$$

Esquema de Iteración Básica (EIB) Dado $\varphi \in \Phi$, $x^0 \in \mathbb{R}_{++}^p$, $\epsilon \geq 0$, y $\lambda_k > 0$.

El esquema genera la sucesión $\{x^k\} \in \mathbb{R}_{++}^p$ que satisface:

$$g^k \in \partial_{\epsilon_k} f(x^k), \quad (2.5.26)$$

$$\lambda_k g^k + \Phi'(x^{k-1}, x^k) = 0, \quad (2.5.27)$$

donde $\Phi' := (a_1 \varphi(b_1/a_1), \dots, a_p \varphi(b_p/a_p))^T$, $\forall a, b \in \mathbb{R}_{++}^p$.

$$(2.5.28)$$

Obtendremos tres relaciones importantes las que serán útiles para nuestro análisis de convergencia estos son.

$$\lambda_k (f(x^k) - f(x)) \leq \left\langle x - x^k, \Phi'(x^{k-1}, x^k) \right\rangle + \lambda_k \epsilon_k. \quad (2.5.29)$$

$$f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k, \quad \forall x \in \mathbb{R}_{++}^p. \quad (2.5.30)$$

$$f(x^{k-1}) - f(x^k) \geq \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k. \quad (2.5.31)$$

$$f(x^{k-1}) - f(x^k) \geq \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k.$$

Definición II.35

$$\epsilon - \operatorname{argmin} F(x) = \{z : F(z) \leq \inf F + \epsilon\}.$$

Donde F es conocida y $\epsilon \geq 0$ y sea

$$F_k(x) = f(x) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}). \quad (2.5.32)$$

De esto se sigue que el Esquema de Iteración Básica, no es otra cosa que una versión aproximada del método proximal, en el sentido que:

$$x^k \in \epsilon_k - \operatorname{argmin} F_k(x).$$

En efecto, sea $\{x^k\}$ generada por el EIB, probaremos que $x^k \in \operatorname{argmin} F_k(x)$.

De (2.5.30) tenemos

$$f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k,$$

$$f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k,$$

$$F_k(x) \geq F_k(x^k) - \epsilon_k$$

$$\inf F_k(x) \geq F_k(x^k) - \epsilon_k,$$

entonces

$$F_k(x^k) \leq \inf F_k(x) + \epsilon_k.$$

Así, se tiene

$$\{x^k\} \in \epsilon_k - \operatorname{argmin} F_k(x). \quad \blacksquare$$

En adelante escogeremos una opción específica para la función φ .

Dado $\mu > 0$, $\nu \geq 0$, parámetros fijos, $h \in \Phi$ se define:

$$\varphi(t) := \mu h(t) + \frac{\nu}{2} (t-1)^2. \quad (2.5.33)$$

Nuestra primera clase de algoritmo es:

Método Proximal Interior(MPI)

Tomamos $\mu = 1$, $\nu = 0$ en (2.5.33)

$\varphi \equiv h \in \Phi$, y $\{x^k\}$ generado por el EIB.

y la segunda es la siguiente:

Método Proximal Interior Regularizado(MPIR)

Tomamos $h \in \Phi_2$ en (2.5.33) con

$$\nu \geq \mu h''(1) > 0$$

y $\{x^k\}$ generado por el EIB.

EL MPI incluye el algoritmo estudiado por Tseng y Bertsekas(1993) y Ben-tal y Zibulevsky (1997). Pero también permite la minimización inexacta. El MPIR aparece completamente nuevo y se mostrará que posee propiedades más fuertes. La motivación para el (MPIR) es que el núcleo h es usado para forzar que las iteraciones, dadas en el MPI, permanezcan en el interior de la región viable. Mientras el uso del núcleo cuadrático $(t - 1)^2$ en d_φ , nos lleva a un término de la forma $\|x - y\|^2$, siendo este el usual término usado en "regularización".

En el MPIR los parámetros μ, ν siempre se asumirá que se cumple la relación:

$$\nu \geq \mu h''(1) > 0.$$

Ahora estudiaremos el análisis de convergencia

Lema II.1 *Sea $\{\nu_k\}$, y $\{\beta_k\}$ sucesiones de números reales no negativos, que cumplen:*

1. $\nu_{k+1} \leq \nu_k + \beta_k$.

2. $\sum_{k=1}^{\infty} \beta_k < \infty$.

Entonces, la sucesión $\{\nu_k\}$ es convergente.

Prueba. Como $\sum_{k=1}^{\infty} \beta_k < \infty$ entonces $\beta_k \rightarrow 0$, es decir,

$$\forall \epsilon > 0, \exists N_0 \in \mathbb{N}, \text{ tal que } \forall k > N_0 \text{ se tiene } \beta_k < \epsilon.$$

Ahora como $\nu_{k+1} \leq \nu_k + \beta_k$ tenemos que $\forall k \geq N_0$,

$$\nu_{k+1} \leq \nu_k + \beta_k < \nu_k + \epsilon.$$

Así, $\forall k \geq N_0, \nu_{k+1} \leq \nu_k$. Definimos

$$\nu = \inf \{\nu_k; k \geq N_0\}.$$

Se tiene que existe k_0 : $\nu_{k_0} < \nu + \epsilon$, $k_0 \geq N_0$ como $k \geq k_0 \geq N_0$
 $\Rightarrow \nu_k \leq \nu_{k_0} < \nu + \epsilon$, además de $\nu \leq \nu_k$ sigue $\nu - \epsilon < \nu \leq \nu_k$
de eso sigue $\nu - \epsilon < \nu_k < \nu + \epsilon$. Por lo tanto

$$\forall \epsilon > 0, \exists N_0 \in \mathbb{N} / \forall k \geq N_0; -\epsilon \leq \nu_k - \nu \leq \epsilon.$$

Es decir $\{\nu_k\}$ es convergente. ■

Lema II.2 *Dados $\{\lambda_k\}$ una sucesión de números positivos, $\{a_n\}$ una sucesión de números reales y $b_n := \sigma_n^{-1} \sum_{k=1}^n \lambda_k a_k$, donde $\sigma_n := \sum_{k=1}^n \lambda_k$. Si $\sigma_n \rightarrow \infty$ entonces*

1. $\liminf a_n \leq \liminf b_n \leq \limsup b_n \leq \limsup a_n$.

2. Si $\lim a_n = a < \infty$, entonces $\lim b_n = a$.

Demostración. Para (1) se tiene

$$b_n = \frac{\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n}{\lambda_1 + \lambda_2 + \dots + \lambda_n}.$$

Definamos

$$a = \sup a_n.$$

Entonces se tiene:

$$\lambda_i a_i \leq \lambda_i a.$$

Sumando de $i = 1$ hasta n y como a es supremo sigue $b_n \leq a$. Además Por propiedad de \limsup ,

$$\limsup b_n \leq \limsup a_n. \tag{2.5.34}$$

Ahora definamos $a' = \inf a_n$. Entonces se tiene:

$$\lambda_i a' \leq \lambda_i a_i.$$

Sumando de $i = 1$ hasta n y como a' es ínfimo sigue $a_n \leq b_n$. Por propiedad de \liminf se tiene:

$$\liminf a_n \leq \liminf b_n. \tag{2.5.35}$$

Se sabe que

$$\liminf b_n \leq \limsup b_n \tag{2.5.36}$$

De 2.5.34, 2.5.35 y 2.5.36 obtenemos el resultado.

$$\liminf a_n \leq \liminf b_n \leq \limsup b_n \leq \limsup a_n.$$

Para (2), se tiene que $\lim a_n = a < \infty$ entonces

$$\liminf a_n = \limsup a_n = a,$$

usando la prueba anterior se obtiene el resultado

$$\lim a_n = \lim b_n = a. \quad \blacksquare$$

A lo largo de esta sección asumiremos la siguiente hipótesis del problema (P):

$$\text{H1} \quad \text{dom} f \cap \mathbb{R}_{++}^p \neq \emptyset.$$

Lema II.3 Sea $x^{k-1} \in \mathbb{R}_{++}^p$, $\lambda > 0$ y F_k definida en (2.5.32), bajo una de las siguientes hipótesis:

- (i) El conjunto óptimo X_* es no vacío y compacto,
- (ii) φ es cofinito, es decir, $\varphi_\infty(d) = +\infty$, para todo $d \neq 0$ (aquí φ_∞ denota la función de resección de φ),

existe un único $y^k \in \mathbb{R}_{++}^p$ tal que

$$y^k = \operatorname{argmin}_x F_k(x), \quad (2.5.37)$$

$$\lambda_k g^k + \Phi'(x^{k-1}, y^k) = 0, \quad g^k \in \partial f(y^k). \quad (2.5.38)$$

Teorema II.20 Sea $\{x^k\}$ una sucesión generada por MPI. Donde, $\bar{\lambda} \geq \lambda_i \geq \lambda >$

0 ,
 $\sum_{k=1}^{\infty} \epsilon_k < \infty$, y el conjunto óptimo es no vacío y compacto, entonces

1. Existe un único $y^k \in \mathbb{R}_{++}^p$ tal que

$$y^k = \operatorname{argmin}_x F_k(x),$$

$$\lambda_k g^k + \Phi'(x^{k-1}, y^k) = 0, \quad g^k \in \partial f(y^k).$$

2. La sucesión es limitada, minimizante y todo punto límite de la sucesión es una solución óptima.

Además tenemos:

$$\liminf_{k \rightarrow \infty} g_i^k \geq 0, \quad \lim_{k \rightarrow \infty} g_i^k x_i^k = 0, \quad \forall i = 1 \dots p. \quad (2.5.39)$$

Para nuestro análisis de convergencia de MPIR comenzaremos con el siguiente lema.

Lema II.4 Sea $h \in \Phi_2$, $\varphi(t) = \mu h(t) + (\nu/2)(t-1)^2$, con $\nu \geq \mu h''(1) > 0$ y

$$\theta := (\nu + \mu h''(1)/2). \quad (2.5.40)$$

Para algún $a, b \in \mathbb{R}_{++}^p$ y $c \in \mathbb{R}_+^p$, tenemos

$$\langle c - b, \Phi'(a, b) \rangle \leq \theta(\|c - a\|^2 - \|c - b\|^2), \quad (2.5.41)$$

donde $\Phi'(a, b)$ es definido en (2.5.28).

Demostración. Sea $t_j = b_j/a_j$. Desde que $h \in \Phi_2$, tenemos

$$h'(t_j) \leq h''(1)(t_j - 1),$$

luego multiplicando por $a_j c_j \geq 0$, $\forall j = 1 \dots p$, se tiene:

$$c_j a_j h'(b_j/a_j) \leq h''(1) c_j (b_j - a_j) \quad (2.5.42)$$

Así mismo, se tiene

$$-h(t_j) \leq -h''(1)(1 - \frac{1}{t_j}), \quad \forall t_j > 0, \quad j = 1 \dots p.$$

Multiplicando la expresión anterior por, $a_j b_j \geq 0$, y de 2.5.42 se tiene:

$$c_j a_j h'(\frac{b_j}{a_j}) - a_j b_j h'(\frac{b_j}{a_j}) \leq h''(1) c_j [b_j - a_j] + h''(1) a_j [a_j - b_j].$$

Multiplicando por μ , luego sumando $c_j b_j \nu (\frac{b_j}{a_j} - 1) - a_j b_j \nu (\frac{b_j}{a_j} - 1)$ adecuadamente, sigue:

$$c_j a_j (\mu h'(\frac{b_j}{a_j}) + \nu (\frac{b_j}{a_j} - 1)) - a_j b_j (\mu h'(\frac{b_j}{a_j}) + \nu (\frac{b_j}{a_j} - 1)) \leq \mu h''(1) c_j [b_j - a_j] + \mu h''(1) a_j [(a_j - b_j)] +$$

$$+c_j a_j \nu \left(\frac{b_j}{a_j} - 1 \right) - b_j a_j \nu \left(\frac{b_j}{a_j} - 1 \right).$$

Ahora desde que $\varphi(t) = \mu h(t) + (\nu/2)(t-1)^2$, tenemos

$$\varphi'(t) = \mu h'(t) + \nu(t-1).$$

Reemplazando $\varphi'(t)$ en la desigualdad anterior y sumando de $j = 1$ hasta hasta p

Se tiene,

$$\langle c-b, \Phi'(a, b) \rangle \leq \mu h''(1) \langle b-a, c-a \rangle + \nu \langle b-a, c-b \rangle.$$

Donde Φ' , esta definido en (2.5.28). Usando las siguientes identidades

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

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

Se obtiene el resultado,

$$\langle c-b, \Phi'(a, b) \rangle \leq \theta (\|c-a\|^2 - \|c-b\|^2). \quad \blacksquare$$

Teorema II.21 Sean $\{\lambda_k\}$ una sucesión arbitraria de números positivos, $\sigma_n := \sum_{k=1}^n \lambda_k$ y $\{x^k\}$ una sucesión generada por MPIR, entonces

(i) existe un único $y^k \in \mathbb{R}_{++}^p$ tal que se cumple (3.1) y (3.2).

(ii) $f(x^n) - f(x) \leq \theta \sigma_n^{-1} \|x - x^0\| + \sigma_n^{-1} \sum_{k=1}^n \sigma_k \epsilon_k$, $\forall x \in \mathbb{R}_+^p$.

(iii) Si $\sigma_n \rightarrow \infty$ y $\epsilon_k \rightarrow 0$ entonces $\liminf_{n \rightarrow \infty} f(x_n) = f_*$. Además si $\sum_{k=1}^n \epsilon_k < \infty$, entonces $f(x^n) \rightarrow f_* = \inf f(x)$, $x \in \mathbb{R}_+^p$ si f_* es finito.

(iv) Si $X^* \neq \emptyset$, $\sigma_n \rightarrow \infty$, y $\sum_{k=1}^{\infty} \lambda_k \epsilon_k < \infty$, entonces la sucesión $\{x^n\}$ converge para la solución óptima de (P). Además, si $\lambda_k \geq \lambda > 0$ entonces, la relación (3.5) cumple.

Para ver las pruebas de estos lemas y teoremas y para más detalles sobre la distancia homogénea de orden dos, se puede ver [1] y [4]

2.6. Métodos proximales basados en núcleos homogéneos de orden r

Considere

$$(P) \begin{cases} \min & f(x) \\ \text{s. a:} & x \geq 0. \end{cases}$$

Donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $x \in \mathbb{R}_+^n$.

Nuestro trabajo de tesis consiste en analizar el método de punto proximal con distancia homogénea de orden r , para resolver el problema (P) , cabe resaltar que en esta parte del trabajo tomaremos y analizaremos las distintas propiedades y teoremas dados en el método de punto proximal de orden dos, para el caso de MPI, pero esta vez induciendo la nueva distancia r -homogénea y de este modo concluir con la convergencia de una solución óptima del problema (P)

2.6.1. Métodos proximales con distancia r -homogéneas

1) Motivación del Algoritmo

$$x^k = \operatorname{argmin} \left\{ f(x) + \frac{1}{\lambda_k} d_\varphi(x, x^{k-1}) : x \geq 0 \right\}.$$

Donde

$$d_\varphi(x, y) = \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right), r \geq 1.$$

Sea: $\varphi : \mathbb{R} \rightarrow (-\infty, +\infty]$ una función convexa propia y cerrada, tal que, $\operatorname{dom}\varphi \subseteq [0, +\infty)$, y que cumple las siguientes propiedades:

- (i) $\varphi \in \mathcal{C}^2(0, +\infty)$.
- (ii) φ es estrictamente convexa en su dominio.
- (iii) $\lim_{t \rightarrow 0} \varphi'(t) = -\infty$.
- (iv) $\varphi(1) = \varphi'(1) = 0$ y $\varphi''(1) > 0^+$.

Denotamos con $\Phi = \{\varphi : \varphi \text{ satisface i - iv}\}$.

Si

$r=1$ entonces se tendrá MÉTODO DE PUNTO PROXIMAL φ -DIVERGENCIA (Teboulle, 1992)

$r=2$ entonces se tendrá MÉTODOS DE PUNTO PROXIMAL DE SEGUNDO ORDEN (Auslender, Teboulle, Ben-Tiba, 1999)

2) Propiedades de la función d_φ .

Propiedad II.1 d_φ es una función homogénea de orden r , es decir,

$$d_\varphi(\alpha x, \alpha y) = \alpha^r d_\varphi(x, y), \quad \forall x, y \in \mathbb{R}_{++}^n, \quad r \geq 1.$$

Demostración En efecto, sean $x, y \in \mathbb{R}_{++}^n$ $x = (x_1, x_2, x_3, \dots, x_n)$, $y = (y_1, y_2, y_3, \dots, y_n)$ $\forall \alpha > 0$, se tiene $\alpha x, \alpha y \in \mathbb{R}_{++}^n$, por definición se tiene,

$$\begin{aligned} d_\varphi(\alpha x, \alpha y) &= \sum_{j=1}^n (\alpha y)_j^r \varphi\left(\frac{\alpha x_j}{\alpha y_j}\right), \quad \forall \alpha > 0, \\ &= \alpha^r \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right), \quad \forall \alpha > 0, \\ d_\varphi(\alpha x, \alpha y) &= \alpha^r d_\varphi(x, y) \quad \forall \alpha > 0. \end{aligned}$$

Propiedad II.2 $\forall (x, y) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ tenemos

$$\begin{aligned} d_\varphi(x, y) &\geq 0, \\ d_\varphi(x, y) &= 0 \quad \text{si y solo si} \quad x = y \end{aligned} \tag{2.6.1}$$

Demostración En efecto, por (iv) tenemos que 1 es punto mínimo de φ y por (ii) φ es estrictamente convexa en todo su dominio, entonces 1 es único punto mínimo. sigue que $\varphi(t) \geq 0$. junto con $d_\varphi(x, y) \geq 0$.

Probaremos la primera implicancia:

$$d_\varphi(x, y) = 0 \quad \implies \quad x = y.$$

$$d_\varphi(x, y) = \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right) = 0 \text{ como } y \in \mathbb{R}_{++}^n \text{ implica } y_j^r > 0 \text{ por lo tanto } \varphi(x_j/y_j) =$$

0 de (ii) y (iv) sigue $\frac{x_j}{y_j} = 1$ por lo tanto $x = y$.

Recíprocamente

$$\text{si } x = y \quad \implies \quad d_\varphi(x, y) = 0.$$

De la hipótesis $x_j = y_j, \forall j = 1, \dots, n$ es decir $\frac{x_j}{y_j} = 1$. Reemplazando en la

definición de d_φ tenemos $d_\varphi(x, y) = \sum_{j=1}^n y_j^r \varphi(1) = \sum_{j=1}^n y_j^r(0) = 0$ concluimos

$$d_\varphi(x, y) = 0.$$

3) Hipótesis

H1) $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$, es una función propia convexa y cerrada.

H2) $\text{dom} f \cap \mathbb{R}_{++}^n \neq \emptyset$ **Esquema general del algoritmo** Dado $\varphi \in \Phi, x^0 \in \mathbb{R}_{++}^n, \epsilon_k \geq 0, y \lambda_k > 0$.

El esquema genera la sucesión $\{x^k\} \subset \mathbb{R}_{++}^n$ que satisface:

$$g^k \in \partial_{\epsilon_k} f(x^k), \quad (2.6.2)$$

$$\lambda_k g^k + \Phi' \left(\frac{x^k}{x^{k-1}} \right) = 0, r \geq 1 \quad (2.6.3)$$

$$\text{donde } \Phi' \left(\frac{x^k}{x^{k-1}} \right) := \left((x_1^{k-1})^{r-1} \varphi' \left(\frac{x_1^k}{x_1^{k-1}} \right), \dots, (x_n^{k-1})^{r-1} \varphi' \left(\frac{x_n^k}{x_n^{k-1}} \right) \right). \quad (2.6.4)$$

$$x^k = \text{argmin} \left\{ f(x) + \frac{1}{\lambda_k} \sum_{j=1}^n (x_j^{k-1})^r \varphi' \left(\frac{x^k}{x_j^{k-1}} \right) \right\}$$

$\partial_\epsilon f(x) = \{s \in \mathbb{R}^n : f(y) \geq f(x) + \langle s, y - x \rangle - \epsilon\}$ es la subdiferencial.

Propiedad II.3

$$\lambda_k (f(x^k) - f(x)) \leq \left\langle \Phi' \left(\frac{x^k}{x^{k-1}} \right), x - x^k \right\rangle + \lambda_k \epsilon_k. \quad (2.6.5)$$

Demostración En efecto, $g^k \in \partial_{\epsilon_k} f(x^k)$, entonces

$$f(x) \geq f(x^k) + \langle g^k, x - x^k \rangle - \epsilon_k, \forall x \in \mathbb{R}_{++}^n \cap \text{dom} f.$$

Esto es,

$$f(x) - f(x^k) \geq \left\langle -\lambda_k^{-1} \Phi' \left(\frac{x^k}{x^{k-1}} \right), x - x^k \right\rangle - \epsilon_k$$

$$\lambda_k (f(x^k) - f(x)) \leq \left\langle \Phi' \left(\frac{x^k}{x^{k-1}} \right), x - x^k \right\rangle + \lambda_k \epsilon_k.$$

Propiedad II.4

$$f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k, \forall x \in \mathbb{R}_{++}^n \cap \text{dom} f. \quad (2.6.6)$$

Demostración En efecto, de (P1)

$$\begin{aligned}\lambda_k(f(x^k) - f(x)) &\leq \sum_{j=1}^n (x_j^{k-1})^{r-1} \varphi' \left(\frac{x_j^k}{x_j^{k-1}} \right) (x_j - x_j^k) + \lambda_k \epsilon_k, \\ &= \sum_{j=1}^n (x_j^{k-1})^r \varphi' \left(\frac{x_j^k}{x_j^{k-1}} \right) \left(\frac{x_j}{x_j^{k-1}} - \frac{x_j^k}{x_j^{k-1}} \right) + \lambda_k \epsilon_k,\end{aligned}$$

Como φ es convexa y diferenciable en $\text{int}(\text{dom}\varphi)$ se cumple:

$$\varphi(r) - \varphi(s) \geq \varphi'(s)(r - s), \quad \forall r, s \in \text{int}(\text{dom}\varphi).$$

Tomando $s = \frac{x_j^k}{x_j^{k-1}} \wedge r = x_j^{k-1}$

$$\begin{aligned}\lambda_k(f(x^k) - f(x)) &\leq \sum_{j=1}^n (x_j^{k-1})^r \left(\varphi \left(\frac{x_j}{x_j^{k-1}} \right) - \varphi \left(\frac{x_j^k}{x_j^{k-1}} \right) \right) + \lambda_k \epsilon_k, \\ &= d_\varphi(x, x^{k-1}) - d_\varphi(x^k, x^{k-1}) + \lambda_k \epsilon_k.\end{aligned}$$

Uniendo los extremos obtenemos:

$$\lambda_k(f(x^k) - f(x)) \leq d_\varphi(x, x^{k-1}) - d_\varphi(x^k, x^{k-1}) + \lambda_k \epsilon_k.$$

Donde se tiene:

$$f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k, \quad \forall x \in \mathbb{R}_{++}^p.$$

Propiedad II.5

$$f(x^{k-1}) - f(x^k) \geq \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k. \quad (2.6.7)$$

Demostración En efecto, tomando $x = x^{k-1}$ en (2.6.6) tenemos

$$f(x^{k-1}) + \lambda_k^{-1} d_\varphi(x^{k-1}, x^{k-1}) \geq f(x^k) + \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k.$$

Donde obtenemos el resultado

$$f(x^{k-1}) - f(x^k) \geq \lambda_k^{-1} d_\varphi(x^k, x^{k-1}) - \epsilon_k. \quad \blacksquare$$

2.6.2. Análisis de convergencia

En este capítulo vamos a demostrar que la sucesión generada por el algoritmo del MPI con la distancia homogénea de orden r , es acotada y, que converge a una solución óptima del problema (P). Para esto estudiaremos la segunda sub clase de Φ $\Phi_1 := \{\varphi \in \Phi : \exists M > 0 \text{ tal que } \varphi''(t) \leq M, \forall t \geq 1 \lim_{t \rightarrow 0} \varphi'(t) = -\infty \text{ o } \lambda_k \text{ es tal que } x^k \in \mathbb{R}_{++}^n, \forall k \geq 0\}$.

Daremos nuestro primer resultado.

Lema II.5 *Sea $x^{k-1} \in \mathbb{R}_{++}^p$, $\lambda_k > 0$ y F_k definida en (2.5.32), bajo una de las siguientes hipótesis:*

- (i) *El conjunto óptimo X_* es no vacío y compacto,*
- (ii) *φ es cofinito, es decir, $\varphi_\infty(d) = +\infty$, para todo $d \neq 0$ (aquí φ_∞ denota la función de recesión de φ),*

entonces existe un único $y^k \in \mathbb{R}_{++}^p$ tal que

$$y^k = \operatorname{argmin}_x F_k(x), \quad (2.6.1)$$

$$\lambda_k g^k + \Phi'(x^{k-1}, y^k) = 0, \quad g^k \in \partial f(y^k). \quad (2.6.2)$$

Demostración. Afirmamos que F_k es estrictamente convexa.

En efecto, de la estricta convexidad de φ , sigue

$$\varphi\left((1-t)\left(\frac{x_j}{y_j}\right) + t\left(\frac{x_j}{y_j}\right)\right) < (1-t)\varphi\left(\frac{x_j}{y_j}\right) + t\varphi\left(\frac{x_j}{y_j}\right), \quad x, y, z \in \mathbb{R}_{++}^n.$$

Multiplicando por y_j^r y luego Sumando de $j = 1$ hasta p

$$\sum_{j=1}^n y_j^r \varphi\left((1-t)\frac{x_j}{y_j} + t\frac{x_j}{y_j}\right) < (1-t) \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right) + t \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right), \quad x, y, z \in \mathbb{R}_{++}^n.$$

Multiplicando por λ^{-1} ,

$$\lambda^{-1} d_\varphi((1-t)x + tz, y) < (1-t)\lambda^{-1} d_\varphi(x, y) + t\lambda^{-1} d_\varphi(z, y).$$

Además desde que f es convexa y de la última desigualdad se tiene:

$$f((1-t)x + ty) + \lambda^{-1}d_\varphi((1-t)x + ty, x^{k-1}) < \\ (1-t)(f(x) + \lambda^{-1}d_\varphi(x, x^{k-1})) + t(f(y) + \lambda^{-1}d_\varphi(y, x^{k-1})), \text{ y así,}$$

$$F_k((1-t)x + ty) < (1-t)F_k(x) + tF_k(y).$$

Por lo tanto F_k es estrictamente convexa. Esto implica la unicidad del conjunto óptimo

$$S_k = \operatorname{argmin}_x F_k.$$

Si S_k es no vacío tenemos que $y^k \in \mathbb{R}_{++}^n$. En efecto, desde que $\varphi \in \Phi$ y de la propiedad $\lim_{t \rightarrow 0^+} \varphi'(t) = -\infty$, se concluye que $\operatorname{dom} \varphi \subset \mathbb{R}_{++}^n$, es decir si, $z \in \operatorname{argmin} F_k$ tendremos

$$z \in \mathbb{R}_{++}^n.$$

Aplicando el criterio de optimalidad para F_k (dado que tiene un minimizador) se obtiene

$$\lambda_k g^k + \Phi'(x^{k-1}, y^k) = 0, \quad g^k \in \partial f(y^k),$$

desde que

$$y^k = \operatorname{argmin} F_k \implies 0 \in \partial F_k, \\ 0 \in \partial(f(y^k) + \frac{1}{\lambda_k}d_\varphi(y^k, x^{k-1})), \\ 0 = g_k + \nabla_1 \frac{1}{\lambda_k}d_\varphi(y^k, x^{k-1}).$$

Además

$$\nabla_1 d_\varphi(y_1^k, \dots, y_n^k, x^{k-1}) = \left(\frac{\partial d_\varphi}{\partial y_1^k}, \dots, \frac{\partial d_\varphi}{\partial y_n^k} \right), \\ = \left(x_1^{k-1} \varphi' \left(\frac{y_1^k}{x_1^{k-1}} \right), \dots, x_n^{k-1} \varphi' \left(\frac{y_n^k}{x_n^{k-1}} \right) \right), \\ = \Phi'(x^{k-1}, y^k).$$

Por lo tanto,

$$0 = g_k + \lambda_k^{-1} \Phi'(x^{k-1}, y^k).$$

Para ambos casos hemos asumido que el conjunto óptimo S_k es no vacío y compacto. Probaremos dicha afirmación, por Rockafellar (teorema 27.1d) tenemos S_k es no vacío y compacto si:

$$(F_k)_\infty(d) > 0 \quad \forall d \neq 0. \quad (2.6.3)$$

De (2.5.32) $F_k = f(x) + d_\varphi(x, x^{k-1})$, tenemos

$$(F_k)_\infty(d) = f_\infty(d) + \delta(d|\mathbb{R}_+^n) + d_\varphi(\cdot, x^{k-1})_\infty(d). \quad (2.6.4)$$

Analizaremos para cada hipótesis.

(i) Si X_* es no vacío y compacto, por [17] (Teorema 27.1d) tenemos

$$f_\infty(d) + \delta(d|\mathbb{R}_+^n) > 0, \quad \forall d \neq 0.$$

Desde que, $d_\varphi(\cdot, x^{k-1}) \geq 0$, sigue que

$$d_\varphi(\cdot, x^{k-1})_\infty(d) = \lim_{t \rightarrow \infty} \frac{d_\varphi(x^{k-1} + td, x^{k-1})}{t} \geq 0 \quad \forall t > 0.$$

De estas dos últimas desigualdades tenemos que se cumple, $(F_k)_\infty(d) > 0$ entonces S_k es no vacío y compacto.

(ii) Si φ es cofinito. Así de (2.6.4) tenemos

$$(F_k)_\infty(d) = f_\infty(d) + \delta(d|\mathbb{R}_+^n) + \infty = +\infty \quad \forall d \neq 0.$$

De esto sigue $(F_k)_\infty(d) = +\infty, \forall d \neq 0$, por lo tanto S_k es no vacío y compacto. ■

Teorema II.22 Sea $\{x^k\}$ una sucesión generada por el Método Proximal interior.

Donde, $\bar{\lambda} \geq \lambda_i \geq \lambda > 0$,

$\sum_{k=1}^{\infty} \epsilon_k < \infty$, y el conjunto óptimo es no vacío y compacto, entonces

1. Existe un único $y^k \in \mathbb{R}_{++}^n$ tal que

$$y^k = \operatorname{argmin}_x F_k(x),$$

$$\lambda_k g^k + \Phi'(x^{k-1}, y^k) = 0, \quad g^k \in \partial f(y^k).$$

2. La sucesión es limitada, minimizante y todo punto límite de la sucesión es una solución óptima.

Además tenemos:

$$\liminf_{k \rightarrow \infty} g_i^k \geq 0, \quad \lim_{k \rightarrow \infty} g_i^k x_i^k = 0, \quad \forall i = 1 \dots n. \quad (2.6.5)$$

Demostración.

(1) Sigue del Lema II.5

(2) Se tiene que $d_\varphi \geq 0$, y de (2.6.7)

$$f(x^{k-1}) + \epsilon_k \geq f(x^k), \quad (2.6.6)$$

tomando sumatoria para $k = 1, \dots, n$ se tiene

$$f(x^0) + \sum_{i=1}^n \epsilon_i \geq f(x^n).$$

Entonces $f(x^n) \leq \sum_{i=1}^n \epsilon_i + f(x^0) = \alpha$.

Se tiene así que el conjunto óptimo es no vacío y compacto, f tiene un conjunto de nivel limitado para α entonces este es limitado para cualquier α .

Ahora como $x^n \in L_f(\alpha)$, entonces $\{x^n\}$ es acotado, así queda probado la primera parte.

De (2.6.6) y del Lema 3.1 $\{f(x^k)\}$ es convergente.

$$\lim_{k \rightarrow \infty} f(x^k) = l_* \in \mathbb{R}, \quad (2.6.7)$$

utilizando (2.6.7) implica que

$$\lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) = 0. \quad (2.6.8)$$

En efecto, de (2.6.7) y $\bar{\lambda} \geq \lambda_k$, tenemos,

$$\bar{\lambda} [f(x^{k-1}) - f(x^k) + \epsilon_k] > \lambda_k [f(x^{k-1}) - f(x^k) + \epsilon_k] \geq d_\varphi(x^k, x^{k-1}),$$

tomando límite para los extremos, se tiene

$$\begin{aligned} \bar{\lambda} [\lim_{k \rightarrow \infty} ((f(x^k) - f(x)) + \epsilon_k)] &\geq \lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) \geq 0, \\ \bar{\lambda} (0) &\geq \lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) \geq 0, \\ 0 &\geq \lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) \geq 0, \end{aligned}$$

usando el teorema del sandwich se obtiene:

$$\lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) = 0.$$

Sea ahora \bar{x} un punto límite de la sucesión $\{x^k\}$, entonces existe una subsucesión $\{x^{k(l)-1}\}$, convergiendo para \bar{x} es decir:

$$\bar{x} = \lim_{l \rightarrow \infty} x^{k(l)-1}, \text{ con } \lim_{l \rightarrow \infty} k(l) = +\infty. \quad (2.6.9)$$

Definiremos los siguientes conjuntos:

$$I^+ = \{i : \bar{x}_i > 0\},$$

$$I^0 = \{i : \bar{x}_i = 0\},$$

y $\forall i \in I^0$ se define:

$$J_+^i = \left\{ k(l) : \frac{x_i^{k(l)}}{x_i^{k(l)-1}} > 1 \right\}$$

$$J_-^i = \left\{ k(l) : \frac{x_i^{k(l)}}{x_i^{k(l)-1}} \leq 1 \right\}.$$

Dado que $\varphi \in \Phi$, se tiene $\varphi(t) \geq 0$, para todo $t > 0$ y $\varphi(1) = 0$, de (2.6.8) tenemos que:

$$1 = \lim_{l \rightarrow \infty} \frac{x_i^{k(l)}}{x_i^{k(l)-1}} \quad \forall i \in I^+.$$

Además

$$\bar{x}_i = \lim_{l \rightarrow \infty} x_i^{k(l)} \quad \forall i \in I^+. \quad (2.6.10)$$

En efecto, como

$$\begin{aligned} \lim_{k \rightarrow \infty} d_\varphi(x^k, x^{k-1}) &= 0, \\ \lim_{k \rightarrow \infty} \sum_{i=1}^n (x_i^{k(l)-1})^r \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) &= 0, \\ \lim_{k \rightarrow \infty} (x_i^{k(l)-1})^r \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) &= 0. \end{aligned}$$

Simbólicamente se tiene

$$\forall \delta > 0, \exists l_0 \in \mathbb{N} / l > l_0 \Rightarrow \left| (x_i^{k(l)-1})^r \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \right| < \delta, \quad \forall i \in I^+.$$

De esto sigue, $\forall i \in I^+$, entonces, $\lim_{k \rightarrow \infty} (x_i^{k(l)-1})^r > 0$.

Afirmamos que,

$$\varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \rightarrow 0,$$

$$\forall \epsilon > 0, \exists l'_0 \in \mathbb{N}/l' > l'_0 \Rightarrow \left| \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \right| < \epsilon, \forall i \in I^+.$$

Supongamos lo contrario:

$$\left| \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \right| \geq \epsilon, \text{ multiplicando por } |(x_i^{k(l)-1})^r|,$$

obtenemos:

$$|(x_i^{k(l)-1})^r| \left| \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \right| \geq |(x_i^{k(l)-1})^r| \epsilon > 0,$$

del primer límite $\delta \geq |(x_i^{k(l)-1})^r| \left| \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) \right| \geq |(x_i^{k(l)-1})^r| \epsilon > 0,$

esto ocurre para todo $\delta > 0$ en particular para $\delta = |(x_i^{k(l)-1})^r| \epsilon > 0,$

con lo cual tenemos la contradicción, por lo tanto

$$\lim_{l \rightarrow +\infty} \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) = 0.$$

Dado que φ es continua

$$\lim_{l \rightarrow +\infty} \varphi\left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}}\right) = 0 = \varphi(1),$$

entonces toda sucesión de puntos $\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \rightarrow 1, \forall i \in I^+.$

Además se tiene,

$$\lim_{l \rightarrow \infty} x_i^{k(l)-1} = \bar{x}_i \forall i \in I^+,$$

como ambos límites existen se sigue:

$$\lim_{l \rightarrow \infty} \frac{x_i^{k(l)}}{x_i^{k(l)-1}} \cdot x_i^{k(l)-1} = 1 \cdot \bar{x}_i \forall i \in I^+,$$

por lo tanto

$$\bar{x}_i = \lim_{l \rightarrow \infty} x_i^{k(l)} \quad \forall i \in I^+,$$

luego usando (2.6.3) $\lambda_k g^k + \Phi'(x^{k-1}, x^k) = 0$, y recordando que: $\lambda_k \geq \lambda > 0$, tenemos:

$$\lim_{l \rightarrow \infty} g_i^{k(l)} = 0 \quad \forall i \in I^+. \quad (2.6.11)$$

Veamos:

$$\begin{aligned} 0 &= \lambda_k g^k + \Phi'(x^{k-1}, x^k) \geq \lambda g^k + \Phi'(x^{k-1}, x^k), \\ 0 &\geq \lambda g^k + \Phi'(x^{k-1}, x^k) > \Phi'(x^{k-1}, x^k), \\ 0 &\geq \lambda g_i^k + (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) > (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right), \end{aligned}$$

pero tenemos por la continuidad de φ :

$$\lim_{t \rightarrow 1} \varphi'(t) = \varphi'(t) = 1,$$

$$\text{además, } \lim_{l \rightarrow \infty} \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) = 1,$$

$$\text{con esto tenemos, } \lim_{l \rightarrow \infty} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) = 0,$$

$$\text{también sabemos } \lim_{l \rightarrow \infty} (x_i^{k(l)-1})^{r-1} = \bar{x}_i^{r-1},$$

y como ambos límites existen concluimos

$$\lim_{l \rightarrow \infty} (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) = \bar{x}_i^{r-1} 0 = 0, \quad (2.6.12a)$$

entonces tomando límite a la desigualdad anterior:

$$\lim_{l \rightarrow \infty} 0 \geq \lim_{l \rightarrow \infty} \left(\lambda g_i^{k(l)} + ((x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right)) \right) \geq \lim_{l \rightarrow \infty} (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right),$$

por teorema del sandwich

$$\lim_{l \rightarrow \infty} \left(\lambda g_i^{k(l)} + (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) \right) = 0, \quad (2.6.12b)$$

dado que los límites 2.6.12a y 2.6.12b existen, obtenemos:

$$\lim_{l \rightarrow \infty} \left(\lambda g_i^{k(l)} + (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) - (x_i^{k(l)-1})^{r-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) \right) = 0$$

$$\lim_{l \rightarrow \infty} \lambda g_i^{k(l)} = 0, \quad \forall i \in I^+.$$

por lo tanto

$$\lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \quad \forall i \in I^+.$$

Se puede obtener:

$$\lim_{l \rightarrow \infty} x_i^{k(l)-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) = 0, \quad \forall i \in I^0, k(l) \in J_+^i.$$

además de (2.6.3) tenemos

$$\lim_{l \rightarrow \infty, k(l) \in J_+^i} g_i^{k(l)} = 0 \quad \forall i \in I^0. \quad (2.6.13)$$

En efecto, dado, $\lambda_k \geq \lambda > 0$ de (2.6.3) sigue:

$$\begin{aligned} \lambda_k g^{k(l)} + \Phi'(x^{k(l)-1}, x^{k(l)}) &\geq \lambda g^{k(l)} + \Phi'(x^{k(l)-1}, x^{k(l)}) > \Phi'(x^{k(l)-1}, x^{k(l)}), \\ 0 &\geq \lambda g^{k(l)} + \Phi'(x^{k(l)-1}, x^{k(l)}) > \Phi'(x^{k(l)-1}, x^{k(l)}), \end{aligned}$$

tomando límite

$$0 \geq \lim_{l \rightarrow \infty} (\lambda g^{k(l)} + \Phi'(x^{k(l)-1}, x^{k(l)})) \geq \lim_{l \rightarrow \infty} \Phi'(x^{k(l)-1}, x^{k(l)}),$$

por teorema de sandwich: $\lim_{l \rightarrow \infty} (\lambda g^{k(l)} + \Phi'(x^{k(l)-1}, x^{k(l)})) = 0$, y dado que

$\lim_{l \rightarrow \infty} x_i^{k(l)-1} \varphi' \left(\frac{x_i^{k(l)}}{x_i^{k(l)-1}} \right) = 0$ obtenemos el resultado:

$$\lim_{l \rightarrow \infty, k(l) \in J_+^i} g_i^{k(l)} = 0, \quad \forall i \in I^0.$$

Además desde que $\varphi \in \Phi$ y por la propiedad (i) y (iv) (es decir φ es estrictamente en su dominio y $\varphi(1) = \varphi'(1) = 0$, $\varphi''(1) > 0$), implica que $\varphi'(t) \leq 0$, $\forall t$ $0 < t \leq 1$.

En efecto, dado que φ es estrictamente convexa, entonces el único mínimo es 1, en $0 < t \leq 1$ se tendrá la gráfica decreciente, es decir su pendiente es negativa, por lo tanto tenemos $\varphi'(t) \leq 0$.

Además se tiene:

$$g_i^{k(l)} \geq 0, \quad \forall k(l) \in J_-^i, \forall i \in I^0.$$

En efecto, de (2.6.3), $\Phi'(x^{k(l)-1}, x^{k(l)})$ con $k(l) \in J_-^i$ y la observación anterior $\varphi'(t) \leq 0$, $0 < t \leq 1$ se tiene $\lambda_k g^k \geq 0$, además desde que $\lambda_k \geq \lambda > 0$, sigue el siguiente resultado:

$$g_i^{k(l)} \geq 0, \quad \forall k(l) \in J_-^i, \forall i \in I^0.$$

De esta desigualdad junto con (2.6.11) y (2.6.13) tenemos

$$\liminf_{l \rightarrow \infty} g_i^{k(l)} \geq 0, \quad \forall i = 1 \dots n. \quad (2.6.14)$$

En efecto:

$$\begin{aligned} & \text{como} \quad \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \quad \forall i \in I^+, \\ \Rightarrow & \liminf_{l \rightarrow \infty} g_i^{k(l)} = 0 \quad \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \end{aligned}$$

además,

$$\begin{aligned} & \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \quad \forall i \in I^0, k(l) \in J_+^i, \\ \Rightarrow & \liminf_{l \rightarrow \infty} g_i^{k(l)} = 0 \quad \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \\ & g_i^{k(l)} \geq 0 \quad \forall i \in J_-^i \Rightarrow \liminf_{l \rightarrow \infty} g_i^{k(l)} \geq 0, \end{aligned}$$

por lo tanto

$$\liminf_{l \rightarrow \infty} g_i^{k(l)} \geq 0, \quad \forall i = 1 \dots n.$$

Ahora tendremos que probar:

$$\lim_{l \rightarrow \infty} g_i^{k(l)} \cdot x_i^{k(l)} = 0, \quad \forall i = 1, \dots, n. \quad (2.6.15)$$

De (2.6.11) y (2.6.13),

$$\begin{aligned} & \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \quad \forall i \in I^+ \\ & \lim_{l \rightarrow \infty} g_i^{k(l)} = 0, \quad \forall i \in I^0, k(l) \in J_+^i, \end{aligned}$$

además dado que la sucesión $\{x^k\}$ es acotada tenemos por propiedad de límites,

$$\lim_{l \rightarrow \infty} g_i^{k(l)} x_i^{k(l)} = 0, \quad \forall i \in I^+, i \in I^0, k(l) \in J_+^i.$$

Por lo tanto, solo faltaría probar que

$$\lim_{l \rightarrow \infty} g_i^{k(l)} x_i^{k(l)} = 0, \quad \forall i \in I^0, k(l) \in J_-^i,$$

desde que $k(l) \in J_-^i$ tenemos:

$$0 < x_i^{k(l)} \leq x_i^{k(l)-1}, \quad \forall i \in I^0,$$

por (2.6.9) $\lim_{l \rightarrow \infty} x_i^{k(l)-1} = \bar{x}_i = 0, \forall i \in I^0$, aplicando límites tenemos

$$\lim_{l \rightarrow \infty} 0 \leq \lim_{l \rightarrow \infty} x_i^{k(l)} \leq \lim_{l \rightarrow \infty} x_i^{k(l)-1},$$

donde:

$$\lim_{l \rightarrow \infty} x_i^{k(l)} = 0 \quad \forall i \in I^0,$$

usando (2.6.14) tenemos que la sub sucesión, $\{g_i^{k(l)}\}_{k(l) \in J_-^i}$, es limitada inferior.

Solo restaría probar que dicha sub sucesión es limitada superiormente.

Tomando $\epsilon > 0$ y $x \in \mathbb{R}_{++}^n \cap \text{dom} f$ y con $x_i > \epsilon, \forall i$. Entonces para $k(l) \in J_-^i$ suficientemente grande, tenemos:

$$x_i - x_i^{k(l)} \geq \frac{\epsilon}{2}, \quad \forall i \in I^0. \quad (2.6.16)$$

En efecto, dado que

$$\lim_{l \rightarrow \infty} x_i^{k(l)} = 0 \equiv \forall \epsilon/2 > 0 \exists l_0 \in \mathbb{N}/l \geq l_0 \Rightarrow |x_i^{k(l)}| \leq \epsilon/2,$$

entonces sumando $x_i \geq \epsilon$ y $-x_i^{k(l)} \geq -\frac{\epsilon}{2}$, se obtiene el resultado

$$x_i - x_i^{k(l)} \geq \frac{\epsilon}{2}, \quad \forall i \in I^0,$$

usando (2.5.26)

$$f(x) \geq f(x^{k(l)}) + \langle g_i^{k(l)}, x_i - x_i^{k(l)} \rangle - \epsilon_{k(l)},$$

se tiene que

$$f(x) \geq f(x^{k(l)}) + \sum_{i=1}^n g_i^{k(l)}(x_i - x_i^{k(l)}) - \epsilon_{k(l)}, \quad (2.6.17)$$

que la sub sucesión, $\{g_i^{k(l)}\}_{k(l) \in J_-^i}$, es limitada para $i \in I^0$.

Supongamos lo contrario. Entonces existirá $i_0 \in I^0$ y una subsucesión $\{g_{i_0}^{k(l(j))}\}_{k(l(j)) \in J_-^{i_0}}$ tal que:

$$\lim_{j \rightarrow \infty} g_{i_0}^{k(l(j))} = +\infty, \quad g_{i_0}^{k(l(j))} \geq 0.$$

Además, desde que la sucesión $\{x^k\}$ es limitada, usando (2.6.14), existe $a \in \mathbb{R}$, tal que para j suficientemente grande:

$$\sum_{i \neq i_0} g^{k(l(j))}(x_i - x^{k(l(j))}) - \epsilon_{k(l(j))} \geq a.$$

En efecto, tenemos que $\liminf g_i^k = \sup \inf \{G^k\} \geq 0$,

donde: $G^k = \{g_i^k, g_i^{k+1}, g_i^{k+2}, \dots\}$

$$\liminf g_i^k \geq 0 \Rightarrow \sup \inf \{G^k\} \geq 0,$$

entonces

$$\exists k_0 \in \mathbb{N} \text{ tal que } \inf \{G^k\} \geq 0, \forall k \geq 0, g_i^k \geq 0,$$

como $\{x_i^k\}$, es limitado entonces, $x_i^k \leq M$ con $M > 0$, además tenemos que

$$\sum_{k=1}^n \epsilon_k < +\infty, \text{ es decir } \epsilon_k \rightarrow 0 \equiv \forall \delta > 0, \exists k'_0 \in \mathbb{N} / \epsilon < \delta,$$

de esto sigue: $g_i^k(x_i - x_i^k) \geq (x_i - M)$, tomando

$$\bar{k}_0 = \max \{k_0, k'_0\} \Rightarrow g_0^k(x_i - x_i^k) - \epsilon_k \geq (x_i - M) - \delta = a,$$

así se tiene

$$g_0^k(x_i - x_i^k) - \epsilon_k \geq a, a \in \mathbb{R}.$$

Así de 2.6.15, 2.6.2 y para un j suficiente grande:

$$g_{i_0}^{k(l(j))}(x_i - x_i^{k(l(j))}) \geq \frac{\epsilon}{2} g_{i_0}^{k(l(j))},$$

$$\sum_{i \neq i_0} g_i^{k(l(j))}(x_i - x_i^{k(l(j))}) - \epsilon_k \geq a.$$

Sumando ambas desigualdades obtenemos:

$$\sum_{i=1}^n g_{i_0}^{k(l(j))}(x_i - x_i^{k(l(j))}) - \epsilon \geq \frac{\epsilon}{2} g_{i_0}^{k(l(j))} + a,$$

entonces

$$f(x) \geq f(x^{k(l(j))}) + \sum_{i=1}^n g_{i_0}^k(x_i - x_i^{k(l)}) - \epsilon \geq f(x^{k(l(j))}) + \frac{\epsilon}{2} g_{i_0}^{k(l(j))} + a,$$

$$f(x) \geq f(x^{k(l(j))}) + \frac{\epsilon}{2} g_{i_0}^{k(l(j))} + a,$$

y desde que $\lim_{j \rightarrow \infty} f(x^{k(l(j))}) = l_*$, tomando límite en la última desigualdad obtenemos:

$$\lim_{j \rightarrow \infty} f(x) \geq \lim_{j \rightarrow \infty} f(x^{k(l(j))}) + \frac{\epsilon}{2} \lim_{j \rightarrow \infty} g_{i_0}^{k(l(j))} + \lim_{j \rightarrow \infty} a,$$

$$f(x) \geq l_* + \infty + a,$$

$$f(x) \geq \infty.$$

Se da la contradicción desde que $x \in \text{dom}(f)$, tenemos que la subsucesión $\left\{g_i^{k(l)}\right\}_{k(l) \in J_-^i}$ es limitada superior para $i \in I^0$.

$$\lim_{l \rightarrow \infty} g_i^{k(l)} x_i^{k(l)} = 0, \quad \forall i \in I^0, \quad k(l) \in J_-^i,$$

se tiene el resultado:

$$\lim_{l \rightarrow \infty} g_i^{k(l)} x_i^{k(l)} = 0, \quad \forall i = 1, \dots, n.$$

Ahora, sea $x \in \mathbb{R}_{++}^n$ arbitrario, tomando limite a (2.6.17)

$$\lim_{l \rightarrow \infty} f(x) \geq \lim_{l \rightarrow \infty} f(x^{k(l)}) + \sum_{i=1}^n \lim_{l \rightarrow \infty} g_i^{k(l)} (x_i - x_i^{k(l)}) - \lim_{l \rightarrow \infty} \epsilon_{k(l)},$$

$$f(x) \geq l_* + \sum_{i=1}^n \left(\lim_{l \rightarrow \infty} g_i^{k(l)} x_i - \lim_{l \rightarrow \infty} g_i^{k(l)} x_i^{k(l)} \right) - 0,$$

$$f(x) \geq l_* + \sum_{i=1}^n \left(\lim_{l \rightarrow \infty} g_i^{k(l)} x_i \right).$$

Por (2.6.14), $f(x) \geq l_* + \sum_{i=1}^n \left(\lim_{l \rightarrow \infty} g_i^{k(l)} x_i \right) \geq l_*$ obtenemos finalmente:

$$f(x) \geq l_* = \lim_{l \rightarrow \infty} f(x^k), \quad \forall x \in \mathbb{R}_{++}^n,$$

con lo cual tenemos que la sucesión $\{x_k\}$ es minimizante, así el teorema queda demostrado. ■

CAPÍTULO III

VARIABLES E HIPÓTESIS

3.1. Variables de la investigación

Variable dependiente: convergencia en optimización convexa

Variable independiente: método de punto proximal con distancia homogénea de orden r .

3.2. Operacionalización de variables

VARIABLES	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIÓN	INDICADORES
Método de punto proximal con distancia homogénea de orden r .	Método que permite determinar un punto óptimo para determinados problemas.	El algoritmo de punto proximal genera una sucesión de puntos que nos llevará a la solución del problema.	\mathbb{R}^n	Convergencia del método.
Convergencia en optimización convexa.	Seleccionar un elemento, de un conjunto de elementos posibles.	Maximizar una función en \mathbb{R}^n .	\mathbb{R}^n	Descubrimiento de un mejor valor.

3.3. Hipótesis

Si el problema de minimización convexa

$$\begin{cases} \min & f(x) \\ \text{s. a :} & x \in \mathbb{R}_+^n, \end{cases}$$

donde $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, es una función propia, convexa, cerrada y $\mathbb{R}_+^n = \{x \in \mathbb{R}^n; x_i \geq 0\}$, es dado por el siguiente algoritmo $x^{k+1} \stackrel{-}{=} \operatorname{argmin} \left\{ f(x) + \lambda_k^{-1} \sum_{j=1}^n y_j^r \varphi\left(\frac{x_j}{y_j}\right); x \in \mathbb{R}_{++}^n \right\}$, para $r = 3; 4; 5; \dots$ tal que $\varphi \in \Phi$, entonces el método de punto proximal es convergente.

CAPÍTULO IV

METODOLOGÍA

4.1. Tipo de investigación

La metodología usada durante la ejecución del proyecto consiste en un enfoque inductivo-deductivo de las definiciones teoremas y corolarios, así como también resultados de recientes investigaciones. luego realizamos un minucioso estudio de cada material obtenido, con la finalidad de adaptarlo a nuestro resultado.

4.2. Diseño de la investigación

El presente proyecto de tesis está dirigido a mostrar de una manera clara la convergencia del método del punto proximal usando la nueva distancia r homogénea, para esto primero empezamos con el método clásico de punto proximal, este genera una sucesión $\{x^k\}$ por el esquema iterativo: comenzar con un punto inicial $x^0 \in \mathbb{R}^n$ y resolver

$$x^k = \operatorname{argmin} \left\{ f(x) + \frac{1}{2} \lambda_k \|x - x^{k-1}\|^2 : x \in X \right\},$$

luego analizamos el método de punto proximal φ -divergencia, también llamada distancia homogénea de primer orden generando la sucesión $\{x^k\}$ por

$$x^k = \operatorname{argmin} \left\{ f(x) + \lambda_k^{-1} \tilde{d}_\varphi(x, x^{k-1}) : x \in \mathbb{R}_+^n \right\},$$

donde $\{\lambda_k\}$ es una sucesión de números positivos y

$$\tilde{d}_\varphi(x, y) := \sum_{j=1}^n y_j \varphi(y_j^{-1} x_j).$$

luego analizamos la llamada distancia homogénea de segundo orden propuesta por Auslender, Teboulle y Ben-tiba, Véase [1], que reemplaza la función \tilde{d}_φ por

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^2 \varphi(y_j^{-1} x_j),$$

finalmente para analizar la convergencia del método de punto proximal con distancia homogénea de orden r dado por

$$d_\varphi(x, y) := \sum_{j=1}^n y_j^r \varphi(y_j^{-1} x_j),$$

demostramos algunas propiedades así como también la homogeneidad.

4.3. Población y muestra

Por ser nuestro trabajo netamente abstracto, no existe población que estudiar, sin embargo nuestro estudio se encuentra inmerso dentro de \mathbb{R}^n y optimización matemática.

4.4. Técnicas e instrumentación de recolección de datos

Para la realización de este trabajo de tesis se revisó una bibliografía especializada y recopilación de información de artículos de tesis y de internet relacionada al tema de interés.

4.5. Procedimiento de recolección de datos

Por ser nuestro trabajo netamente abstracto, no se necesitó procedimiento de recolección más que la revisión bibliográfica (libros, páginas web, artículos, etc).

4.6. Procesamiento estadístico y análisis de datos

No hubo procesamiento alguno.

CAPÍTULO V

RESULTADOS

Los resultados mas relevantes de esta tesis son:

1. Se ha introducido una nueva distancia homogénea de orden r , que cumple las condiciones de la distancia φ -divergencia, tanto de primer orden como de segundo orden.
2. Presentamos el método del punto proximal con distancia homogénea de orden r para minimizar funciones convexas. Se obtiene que la sucesión $\{x^k\}$ generada por nuestro algoritmo converge a una solución óptima para el método de punto proximal.

CAPÍTULO VI

DISCUSIÓN DE RESULTADOS

1. Se ha considerado los resultados de convergencia del método para minimizar funciones convexas. Un estudio importante sería para funciones cuasi-convexas, el cual es un problema de optimización de mucho interés en la Economía y la Administración.
2. Se ha considerado solo el análisis de convergencia, el estudio de la velocidad de convergencia es un método importante desde el punto de vista teórico, pues nos da un mejor panorama de la eficiencia del método.
3. En la nueva distancia homogénea de orden r presentado es necesario analizar en que intervalo los parámetros λ_k aceleran la velocidad de convergencia.
4. No se ha logrado la implementación del algoritmo computacionalmente quedando abierta para futuros trabajos, dado que todo algoritmo debe ser implementado y comparado.

CAPÍTULO VII

CONCLUSIONES

1. En esta tesis hemos recopilado resultados de la convergencia del método del punto proximal con distancia homogénea de orden dos presentado por Auslender, Ben tiba y Teboulle en 1999 y la adaptamos para distancia homogénea de orden r observando que las propiedades de homogeneidad se verifican al igual que la convergencia hacia un punto óptimo. Esta tesis puede mejorarse el algoritmo usado en el sentido computacional, como se puede corroboran en distintas investigaciones en la actualidad.
2. Espero que este trabajo de tesis sirva como referencia para futuros trabajos de investigación que se puedan realizar, pues hay otros detalles por explorar, tanto en el aspecto teórico como computacional, dado su importancia en los distintos campos de la Ingeniería y Economía.

CAPÍTULO VIII

RECOMENDACIONES

1. Es recomendable este trabajo de tesis para aquellos estudiantes e investigadores que opten por la línea de Matemática Aplicada.
2. Este trabajo recopila el método de punto proximal de segundo orden presentado por Auslender alfred, Marc Teobulle, Sami ben tiba en 1999.
3. También es recomendable para futuras tesis, por ejemplo se podría hacer una tesis considerando la velocidad de convergencia.

REFERENCIAS BIBLIOGRÁFICAS

- [1] AUSLENDER ALFRED, MARC TEOBULLE, SAMI BEN TIBA *Interior Proximal And Multiplier Methods Based On Second Order Homogeneous Kernels*, Mathematics Of Operations Research 24. N° 3, Print in USA, (1999)
- [2] BARTLE ROBERT G., *The Elements of Real Analysis*. 2da. edition, John Wiley & Sons. Inc. New York, (1992)
- [3] BEN-TAL AHARON, AND., ZIBULEVSKY MICHAEL *Penalty -Barrier Methods for Convex Programming Problems*. SIAM J. optim 7 347-346,(1997)
- [4] BORDA MARCATINCO D., *Métodos Multiplicativos y Proximales en el Ortante Positivo* práctica pre profesional universidad nacional del callao(2008)
- [5] CENSOR, Y., ZENIOS, S., *The Proximal Minimization Algorithm with D-functions*. journal of optimization theory and applications 73(1992) 451-464
- [6] CROUZEIX J.P., OCAÑA E. Y SOSA W., *Análisis convexo*, IMCA monografía N°33, (2003).
- [7] EGGERMONT P. B., *Multilicative iterative algorithms for convex programming*, Linear Algebra Appl. vol. 130 (1990), pp. 403-419
- [8] FLORENZIANO MONIQUE., *Finite Dimensional Convexity and Optimization*. Paris-Francia: MSE-Université de Paris 1, 2001.
- [9] IUSEM ALFREDO, *Métodos de Ponto Proximal em Otimização*. 20° colóquio brasileiro de matemática. IMPA 24-28 JULHO. (1995)

- [10] A. IUSEM , M. TEBoulLE AND B. SVAITER, *Entropy-like proximal methods in convex programming*, Mathematics of Operations Research, 19(4), pp. 790-814, 1994.
- [11] IUSEM ALFREDO, *Métodos de Ponto Proximal em Otimização*. 2da. edition, John Wiley & Sons. Inc. New York, (1992)
- [12] IZMAILOV ALEXEY Y SOLODOV MICHAEL, *Optimização volume1 Condições de otimalidade, Elementos de Análise Convexa e de Dualidade*, IMPA,(2005).
- [13] LIMA ELON L., *Análise Real, Vol 1* (Sétima Edição), Coleção Matemática Universitária, IMPA, (2004).
- [14] TEBoulLE MARC *Entropic proximal mappings with applications to nonlinear programming*. Math. Oper. Res., 17, 670-690, (1992).
- [15] TEBoulLE MARC *Convergence of Proximal like algorithms*. SIAM J. optim 7 1069-1083,(1997).
- [16] OSMAN GÜLER, *On the Convergence of the Proximal Point Algorithm for Convex Minimization*. SIAM J. Control Optim. 29, pp. 403-419 (17 pages)
- [17] ROCKAFELLAR,R.T., *Convex Analysis*, Princenton Univerity Press, Princenton, NJ. (1970)
- [18] MOREAU J., *Proximité et dualité dans un espace Hilbertien*, Bull. Soc. Math. france 93 pp 97-116, 1.976.

ANEXO

ANEXO A

Matriz de Consistencia

PROBLEMA	OBJETIVO	HIPÓTESIS	METODOLOGÍA	POBLACIÓN
<p>El algoritmo clásico de punto proximal, genera una sucesión $x^k = \operatorname{argmin}\{f(x) + \frac{1}{2}\lambda_k \ x - x^{k-1}\ ^2\}$. En 1992, Véase [15], reemplaza la distancia cuadrática por la cuasi-distancia, conocida con el nombre de φ-divergencia, $\tilde{d}_\varphi(x, y) := \sum_{j=1}^n y_j \varphi(y_j^{-1} x_j)$. En 1999, Véase [1], reemplazan la función \tilde{d}_φ por $d_\varphi(x, y) := \sum_{j=1}^n y_j^2 \varphi(y_j^{-1} x_j)$, llamada distancia homogénea de segundo orden. Una idea que surge naturalmente es explorar las propiedades de convergencia, considerando la función: $d_\varphi(x, y) := \sum_{j=1}^p y_j^r \varphi(y_j^{-1} x_j)$, con $r = 1; 2; 3; \dots$ la que llamamos distancia homogénea de orden r.</p> <p>Planteamiento del problema ¿Será posible obtener la convergencia del método del punto proximal que usan las distancias homogéneas d_φ de orden r?</p>	<p>Objetivo General.- Analizar la convergencia del método de punto proximal cuya sucesión de puntos $x^k = \operatorname{argmin}\{f(x) + \lambda_k^{-1} d_\varphi(x, x^{k-1})\}$ para la distancia homogénea de orden r, donde $d_\varphi(x, y) := \sum_{j=1}^p y_j^r \varphi(\frac{x_j}{y_j})$, con $r = 3; 4; 5; \dots$, en problemas de optimización de convexa.</p> <p>Objetivo Específico.- Estudiar las propiedades de convergencia del método proximal usando las distancias homogéneas d_φ de orden r, para $r = 1$ y $r = 2$. Demostrar las propiedades de las distancias homogéneas d_φ de orden r.</p>	<p>Si el problema de minimización convexa $\begin{cases} \min & f(x) \\ \text{s. a:} & x \in \mathbb{R}_+^n, \end{cases}$ donde $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, es una función propia, convexa, cerrada y $\mathbb{R}_+^n = \{x \in \mathbb{R}^n; x_i \geq 0\}$, es dado por el siguiente algoritmo $x^{k+1} = \operatorname{argmin}\{f(x) + \lambda_k^{-1} \sum_{j=1}^n y_j^r \varphi(\frac{x_j}{y_j})\}$ para $r = 3; 4; 5; \dots$ tal que $\varphi \in \Phi$, entonces el método de punto proximal es convergente.</p>	<p>Tipo de investigación La investigación es de tipo científico-teórico.</p> <p>Método La metodología usada durante la ejecución del proyecto consiste en un enfoque inductivo-ductivo de las definiciones teoremas y corolarios.</p> <p>Diseño de la investigación El presente proyecto de tesis está dirigido a mostrar de una manera clara la convergencia del método del punto proximal, usando la nueva distancia r homogénea, para esto primero empezamos con el método clásico de punto proximal, luego analizamos el método proximal de primer orden y de segundo orden, con los cual tenemos un enfoque para cada caso, finalmente buscamos obtener la homogeneidad de la nueva distancia junto con el análisis de convergencia.</p>	<p>Por ser nuestro trabajo netamente abstracto, no existe población que estudiar, sin embargo, nuestro estudio se encuentra inmerso dentro de \mathbb{R}^n y la optimización matemática.</p>

ANEXO B

Mapa conceptual del trabajo

