Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ciencias Matemáticas
Unidad de Posgrado

Método del punto proximal inexacto con cuasi-distancia para optimización

TESIS

Para optar el Grado Académico de Magíster en Matemática Aplicada con mención en Matemática Computacional

AUTOR

José Luis HUAMÁN ÑAUPA

ASESOR

Dr. Erik Alex PAPA QUIROZ

Lima, Perú

2023



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Huamán, J. (2023). *Método del punto proximal inexacto con cuasi-distancia para optimización*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	José Luis Huamán Ñaupa
Tipo de documento de identidad	DNI
Número de documento de identidad	10069890
URL de ORCID	https://orcid.org/0000-0002-9832-3443
Datos de asesor	
Nombres y apellidos	Erik Alex Papa Quiroz
Tipo de documento de identidad	DNI
Número de documento de identidad	10451642
URL de ORCID	https://orcid.org/0000-0002-8678-6918
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Edinson Raúl Montoro Alegre
Tipo de documento	DNI
Número de documento de identidad	09627181
Miembro del jurado 1	
Nombres y apellidos	Marlo Carranza Purca
Tipo de documento	DNI
Número de documento de identidad	09965090
Miembro del jurado 2	
Nombres y apellidos	Felipe Antonio García Moreno
Tipo de documento	DNI
Número de documento de identidad	06804667
Datos de investigación	

Línea de investigación	Optimización Matemática Aplicada
Grupo de investigación	Optimización Matemática y Computacional - OPTIMACO
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	País: Perú Departamento: Ayacucho Provincia: Huamanga Distrito: Jesús Nazareno Latitud: -13.15113 Longitud: -74.22186
Año o rango de años en que se realizó la investigación	2019 – 2023
URL de disciplinas OCDE	Matemáticas aplicadas https://purl.org/pe-repo/ocde/ford#1.01.02

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS MATEMÁTICAS
VICEDECANATO DE INVESTIGACIÓN Y POSGRADOUNIDAD DE POSGRADO

ACTA DE SUSTENTACIÓN VIRTUAL DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGÍSTER EN MATEMÁTICA APLICADA CON MENCIÓN EN MATEMÁTICA COMPUTACIONAL.

Siendo las, 10:10 horas del día viernes treinta de junio de dos mil veintitrés, en la sala virtual meet.google.com/kbo-daur-itk, el Jurado de Tesis conformado por los siguientes docentes:

Presidente	:	Dr. Edinson Raúl Montoro Alegre
Miembro	:	Mg. Marlo Carranza Purca
Miembro Externo:		Dr. Felipe Antonio García Moreno
Asesor	:	Dr. Erik Alex Papa Quiroz

se reunieron para la sustentación de la tesis titulada: **«MÉTODO DEL PUNTO PROXIMAL INEXACTO CON CUASI-DISTANCIA PARA OPTIMIZACIÓN»** presentada por el Señor Bachiller **José Luis Huamán Ñaupa**, para optar el Grado Académico de Magíster en Matemática Aplicada con Mención en Matemática Computacional.

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, según tabla adjunta, habiendo obtenido el Señor Bachiller **José Luis Huamán Ñaupa**, el Calificativo de 18 **(MUY BUENO)**.

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de **Magíster en Matemática Aplicada con Mención en Matemática Computacional** al Bachiller **José Luis Huamán Ñaupa**.

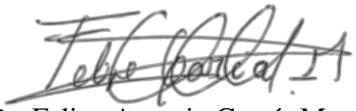
Siendo las 11:23. horas, se levantó la sesión, firmando para constancia la presente Acta:



Dr. Edinson Raúl Montoro Alegre
PRESIDENTE



Mg. Marlo Carranza Purca
MIEMBRO



Dr. Felipe Antonio García Moreno
MIEMBRO EXTERNO



Dr. Erik Alex Papa Quiroz
MIEMBRO ASESOR

Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

FACULTAD DE CIENCIAS MATEMÁTICAS

VICEDECANATO DE INVESTIGACIÓN Y POSGRADO

UNIDAD DE POSGRADO

Informe de Evaluación de Originalidad

Vicedecano de Investigación y Posgrado: Dr. Eugenio Cabanillas Lapa

Operador: Mag. José Luis Acuña Guillermo

Tesis: "MÉTODO DEL PUNTO PROXIMAL INEXACTO CON CUASI-DISTANCIA PARA OPTIMIZACIÓN"

Tesista: JOSE LUIS HUAMAN ÑAUPA

Recepción de documento: 04/01/2023

Aplicación del software: 30/01/2023

Software utilizado: Turnitin

Configuración del software:

- Excluir material bibliográfico
- Excluir material citado
- Excluir fuentes pequeñas N° palabras: 40
- Sin depósito

Resultado del análisis con el software: ocho por ciento (8%)

Fuentes de originales de las similitudes encontradas:

- Ver abajo

Documento cumple con criterios de originalidad.

- **El documento cumple los criterios de originalidad. <u>Sin observaciones</u>.**

Fecha del informe: 31/01/2023

UNMSM

Firmado digitalmente por CABANILLAS LAPA Eugenio FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 31.01.2023 19:57:55 -05:00

Dr. Eugenio Cabanillas Lapa

VICEDECANO DE INVESTIGACIÓN Y POSGRADO

DEDICATORIA

Este trabajo va dedicado a mi familia y en particular para mi esposa e hija, que me dieron su apoyo y paciencia en la elaboración de esta tesis.

AGRADECIMIENTOS

Gracias a Dios todopoderoso por darme salud para seguir adelante con mi proyecto de vida.

A toda la familia tanto de mi esposa como la mía por apoyarme en la elaboración de esta tesis, en especial para mi esposa e hija.

Un particular agradecimiento al Dr. Erik Alex Papa Quiroz,por su paciencia en la elaboración y la creación de este trabajo, por motivarme a seguir investigando y estudiando cada vez más.

A mis compañeros de clases, por motivarme y alentarme a seguir en la terminación del desarrollo de esta tesis.

Índice general

1. INTRODUCCIÓN **2**
- 1.1. Situación problemática 2
- 1.2. Formulación del problema 4
  - 1.2.1. Problema general 4
  - 1.2.2. Problemas específicos 4
- 1.3. Justificación de la investigación 4
  - 1.3.1. Objetivos 5

2. MARCO TEÓRICO **6**
- 2.1. Elementos básicos de análisis en $\mathbb{R}^n$ 6
  - 2.1.1. Sucesiones en $\mathbb{R}^n$ 6
- 2.2. Elementos básicos de convexidad 8
- 2.3. Funciones Lipschizianas 13
- 2.4. Subdiferenciabilidad 13
  - 2.4.1. Subdiferenciabilidad Clásica 14
  - 2.4.2. Subdiferenciabilidad Generalizada 16
  - 2.4.3. Subdiferenciabilidad de Fréchet 18

3. MÉTODO DEL PUNTO PROXIMAL **21**
- 3.1. Método del punto proximal clásico 21
  - 3.1.1. Algoritmo de punto proximal Clásico 22
- 3.2. Análisis de la convergencia del algoritmo de punto proximal 24

4. CUASI-DISTANCIA **32**
- 4.1. Espacios con distancias no simétricas 32

5. ALGORITMO PROXIMAL INEXACTO **38**
- 5.1. Antecedentes del problema 38
- 5.2. El algoritmo 40
- 5.3. Algoritmo Proximal Inexacto (API) 40
  - 5.3.1. Existencia de los iterados 41
  - 5.3.2. Propiedades de convergencia para el caso general 41
  - 5.3.3. Propiedad de convergencia para funciones localmente Lipschitz 43
  - 5.3.4. Propiedad de convergencia para funciones convexas 45
  - 5.3.5. Convergencia para funciones que satisface la condición de Kurdyka Lojasiewicz 46

5.3.6. Algunos ejemplos 50

6. VELOCIDAD DE CONVERGENCIA PARA FUNCIONES QUE SATISFACEN LA CONDICIÓN DE K.L. 60
6.0.1. Aplicaciones conductuales 70

7. METODOLOGÍA 72
7.1. Tipo y Diseño de Investigación 72
7.2. Unidad de análisis 72
7.3. Población de estudio 72
7.4. Técnica para la recolección de datos 72
7.5. Análisis e interpretación de la información 73

8. RESULTADO Y DISCUSIÓN 74

9. CONCLUSIONES Y FUTURAS INVESTIGACIONES 76
9.1. Conclusiones 76
9.2. Futuras investigaciones 77

Bibliografía 80

RESUMEN

Autor :	Jose Luis Huaman Ñaupa.
Asesor :	Dr. Erik Alex Papa Quiroz.
Grado obtenido :	Magíster en matemática aplicada.
Mención :	Matemática computacional.

En esta tesis exponemos un algoritmo de punto proximal inexacto que utiliza cuasi-distancia para resolver problemas de minimización en espacios Euclidianos. Este algoritmo ha sido motivado por el método proximal introducido por Attouch, Bolte y Svaiter, sección 4 (Math Program, Ser A 137:91-129), pero aquí consideramos cuasi-distancia, para funciones objetivo arbitrarias (no necesariamente diferenciables), errores vectoriales en el residual del punto critico regularizado. Obtenemos bajo algunos supuestos adicionales la convergencia de esta sucesión producida por el algoritmo a un punto crítico del problema.

Palabras claves:

- Método del punto proximal.
- Cuasi-distancia.
- Método inexacto.
- Rapidez de convergencia.

Índice de figuras

2.1. (a) conjunto no convexo, (b) conjunto convexo 8
2.2. Gráfica de f 9
2.3. Gráfica del conjunto de nivel $L_f(\alpha)$ y el dominio efectivo dom f. . . . 11
2.4. Gráfica de $epi(f)$ 11
2.5. La función valor absoluto. 15
2.6. Gráfica de el subdiferencial de la función f. 15
2.7. La gráfica de la función $f(x) = x^2.sen(\frac{1}{x})$ 17
2.8. La derivada de la función $f(x) = x^2 sen(1/x)$ con $x \neq 0$ 20

6.1. La región resaltada tiene por base $\triangle_{k-1} - \triangle_k$ y altura $\Phi(\triangle_{k-1})$. . . . 62

NOTACIONES Y SÍMBOLOS

Notación	Explicación
$\mathbb{R}^n$	$\{y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n; y_i \in \mathbb{R}, \forall i = 1, \ldots, n\}$ indica el espacio Euclidiano n-dimensional.
$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{+\infty\}$.
$x^k \to \bar{x}$	$\lim_{k \to +\infty} x^k = \bar{x}$ en el espacio Euclidiano $\mathbb{R}^n$.
∇f	Gradiente de la función f.
$f'(x, v)$	La derivada direccional de la función f evaluado x en la direc ción del vector v.
$\partial_c f(x_0)$	El subdiferencial de la función convexa f evaluada en $x_0 \in \mathbb{R}$.
$\widehat{\partial} f(x_0)$	El subdiferencial de Fréchet de f evaluada en el punto $x_0 \in \mathbb{R}$.
$\partial f(x_0)$	El subdiferencial limitante de f evaluado en el punto $x_0 \in \mathbb{R}$.
$f^\circ(x, d)$	La derivada direccional de f en el punto x (en el sentido de Clarke) en la dirección del vector d.
$\partial^\circ f(x)$	El subdiferencial de f en x (en el sentido de Clarke)
$\langle x, y \rangle = x^t.y$	$\sum_{i=1}^n x_i y_i$ producto interior de $x, y \in \mathbb{R}^n$.
$\|\|x\|\| = \|\|x\|\|_2$	$\sqrt{\langle x, x \rangle}$ norma Euclideana.
$B(\bar{x}, \epsilon)$	$\{x \in \mathbb{R}^n; \|\|x - \bar{x}\|\| < \epsilon\}$ llamada bola abierta
$epi(f)$	$\{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}; f(x) \leq \alpha\}$ epigrafo de la función f.
$L_f(\alpha)$	$\{x \in \mathbb{R}^n; f(x) \leq \alpha\}$ conjunto de nivel inferior de f.

$\lim\limits_{k\to+\infty} \inf f(x^k)$	$\sup \inf\limits_{n\geq 0\ k\geq n} \{f(x^k)\}$ límite inferior de la sucesión.
$\lim\limits_{y\to x\ t\downarrow 0} \sup f(y)$	$\inf\limits_{\delta>0} \sup\limits_{y\in B(x,\delta), t\in(0,\delta)} f(y) = \inf\limits_{\delta>0} \sup\limits_{\|\|y-x\|\|+t<\delta} f(y)$ límite superior de una función localmente Lipschitz.
$z \in arg\min\limits_{y\in X}\{g(y)\}$	z es un punto de mínimo de la función g en X
int(X)	Conjunto de puntos interiores de X.

Capítulo 1

INTRODUCCIÓN

1.1. Situación problemática

La Optimización Matemática, es una de las ramas de la Matemática Aplicada, que ayuda a resolver diversos problemas en el área de la ciencia e ingeniería. Un problema de optimización se describe por

$$\text{mín}\{\bar{f}(x);\ \text{con } x \in M\} \equiv \min_{s.a:\ x\in M} \bar{f}(x), \tag{1.1}$$

donde la expresión mín significa minimizar la función $\bar{f} : I\!R^n \to I\!R$ y $M \subseteq I\!R^n$ contiene a los elementos que son restricciones del problema. Aquí en el presente tesis, se trabajará con funciones extendidas a valores reales, es decir, con funciones del tipo $\bar{f} : I\!R^n \to I\!R \cup \{+\infty\}$ (denotamos por $I\!R \cup \{+\infty\} = \overline{I\!R}$).

La dificultad que se presentan en algunos problemas de optimización, motivaron a varios de los científicos a indagar diversos métodos para resolverlos, entre ellos se encuentra algunos de los métodos que es el del punto proximal (MPP) incorporado por Martinet (1970); este método resuelve un problema de minimización convexa. Después Rockafellar (1976) utilizó el MPP para localizar ceros de operadores monótonos maximales. Tuvo una gran aceptación estos algoritmos, que más adelante fueron extendidos para resolver problemas de desigualdades variacionales, problemas de equilibrio con operadores monótonos y posteriormente operadores no monótonos, en particular para los problemas no convexos de minimización. La gran atención no solo se debió a la extensión de la teoría para funciones no convexas, si no también a las multiples aplicaciones, véase los trabajos de Attouch and Bolte (2009), Attouch et al. (2013), Chen and Pan (2008).

En las últimas décadas el método mencionado a sido extendido con la finalidad de resolver problemas (1.1) cuasi convexa esto se debe a diversos campos de aplicaciones, véase esto en localización, dos Santos Gromicho (2013) y en la teoría económica, ver por ejemplo Takayama and Akira (1985). mientras que Souza et al. (2010) utilizo una distancia generalizada y Papa Quiroz and Oliveira (2012) lo estudio sobre variedades de Hadamard.

El (MPP) es un proceso repetitivo que genera una sucesión $\{x^k\}$ dado un punto inicial x^0 tal que para todo $k \in I\!N$ se tiene que encontrar un punto x^{k+1} satisfaciendo

la siguiente condición:

$$x^{k+1} \in argmin\{\bar{f}(s) + \frac{1}{2\lambda_k}||s - x^k||^2 : s \in I\!R^n\} \tag{1.2}$$

donde $argmin$ es el conjunto de soluciones óptimas de la función $f(s)+\frac{1}{2\lambda_k}||s-x^k||^2$, $\{\lambda_k\}_{k\in I\!N}$ una sucesión positiva y $||\ \ || = \langle\cdot,\cdot\rangle^{1/2}$ es la norma Euclideana.

En los recientes años se han realizado muchos estudios para extender el MPP con el objetivo de solucionar problemas de optimizaciónque involucran funciones no convexa y no diferenciables. Uno de estos estudios propuso el uso de la distancia generalizada, es decir, funciones similares a la distancia del término cuadrático en (1.2), por ejemplo la distancia de Bregman, dada en Kiwiel (1997) en lugar de la Euclideana de (1.2). Este tipo de distancia preserva algunas de las propiedades de la norma Euclideana como, la convexidad, continuidad y coercitividad. Esta situación se complica cuando distancias similares no tienen algunas de estas propiedades mencionadas.

El concepto de espacio métrico no es suficiente para capturar todos los fenómenos interesantes que ocurren en la naturaleza. Esto nos lleva a estudiar bajo hipótesis mas débil a la distancia, sin embargo la mayoría de estos espacios son cuasi metrizables, es decir, espacios que cumplen con algunas propiedades de un espacio métrico. Las cuasi métrica o también llamada la cuasi-distancia fueron extensivamente estudiados en el contexto topológico por Gerla et al. (2006), Künzi (2001), Stojmirovic (2003). Ellos generalizan la distancia en el sentido de que éstas cuasi-distancias no son simétricas. La cuasi-distancia es aplicable en teoría computacional veáse Brattka (2003), Künzi et al. (2006) también son aplicables en Economía,vea Romaguera and Sanchis (2003) tal como función de utilidad y elección de un consumidor.

El algoritmo del punto proximal con cuasi-distancia (MPPQ) es una generalización de (1.2) expresado por

$$x^{k+1} \in argmin\{f(u) + \frac{\lambda_k}{2}q^2(u, x^k); u \in I\!R^n\} \tag{1.3}$$

con q la cuasi distancia, λ_k es una sucesión acotada de números positivos y f una función propia semicontinua inferior. La convergencia del MPPQ de (1.3) fue estudiada y desarrollada por Moreno et al. (2012) en la versión exacta.

En el estudio del algoritmo del punto proximal se observó que en cada subproblema dado por (1.2) se obtiene computacionalmente una solución aproximada y esto motivó a los investigadores a extender su estudio a algoritmos en la versión inexacta, siguiendo estas ideas se tiene la extensión para funciones cuasi-convexas vea Papa Quiroz et al. (2015), funciones semicontinua inferior Attouch et al. (2013) y con término de perturbación generalizada utilizando cuasi-distancia Bento and Soubeyran (2015b) y para el caso multiobjetico el trabajo de Rocha (2016).

Motivado por el reciente trabajo de Papa Quiroz (2019b) en el sentido inexacto (Método del punto proximal inexacto MPPI) con funciones multiobjetivo, quien, nos proporciona un criterio de error en la subdiferencial dada en el espacio Euclideano,

que bajo ciertas condiciones nos asegura la convergencia de la sucesión hacia un punto crítico.
Un error en la subdiferencial del algoritmo (1.3) es dado como

$$e^k = g^k + \lambda_k \partial(q(x^k, x^{k+1})) \tag{1.4}$$

donde e^k es error y $g^k \in \partial f$.
Fue Bento and Soubeyran (2015b) quien trabajó con el Método de Punto Proximal Inexacto con cuasi-distancia (MPPIQ), cuyo error en la subdiferencial de (1.3) es dado por (1.4) que implica la condición débil

$$||g^{k+1}|| \leq b\Gamma'(q(x^k, x^{k+1}))||v^{k+1}||$$

$g^k \in \partial f(x^k)$, ∂f es la subdiferencial de Frechet, Γ es una función que satisface ciertas condiciones, $b > 0$ y $v^k \in \partial q(x^k, .)$. Es con ésta condición débil que va ha trabajar en su algoritmo.

Aquí estamos interesados en estudiar el MPPIQ considerando el error (1.4) para funciones que satisfacen la desigualdad de Kurdyka Lojasiewicz. Como no existen trabajos existentes relacionados con este tipo de error nos proponemos a realizar dicho trabajo.

1.2. Formulación del problema

1.2.1. Problema general

¿ Es posible ampliar la versión de tipo inexacta del método de punto proximal usando cuasi-distancia en vez de la distancia E uclideana ?.

1.2.2. Problemas específicos

1. ¿Qué restricciones deben comprender los errores de aproximación con la finalidad que el MPPI con cuasi distancia sea extendido?

2. ¿Cuáles serán las condiciones de convergencia para el MPPI con cuasi distancia?.

3. ¿Cómo resolver un caso particular con el MPPI con cuasi distancia?.

1.3. Justificación de la investigación

El avance de la siguiente tesis se apoya en lo siguiente:

1. Por reportes de algunos trabajos que afirman la efectividad de las iteraciones inexactas y son muy favorables en dar soluciones aproximadas a problemas de minimización. Entre los informes destacados cabe enfatizar los trabajos del paper Papa Quiroz et al. (2015) donde se estudió la convergencia del MPP inexacto

para funciones cuasi-convexas. Otros reportes dan fe acerca de las funciones cuasi convexas que son diferenciables y hacen uso de las iteraciones inexactas las cuales serán muy útiles para el propósito de este proyecto, también se dan en Bento and Neto (2014), Pan and Chen (2007).

2. El paper de Moreno et al. (2012) nos da una motivación para el uso del método de punto proximal inexacto con cuasi-distancia, pues en ella se da una aplicación (en el campo de la economía) utilizando cuasi-distancia que viene a ser una generalización de MPP.

3. En las recientes décadas se ha empezado a tomar un gran interés en las extensiones de versión exacta de los algoritmos de punto proximal a una de versión inexacta. Esto ocurre no solo por que presenta una extensión simple de la teoria de convergencia, sino más bien a la diversidad de áreas que se le puede aplicar. Ver por ejemplo los trabajos de Attouch and Bolte (2009), Attouch et al. (2013), Kaplan and Tichatschke (1998), Mas Colell et al. (1995).

   Un caso particular son los recientes reportes que se enfocan en un problema de minimización para funciones objetivos cuasiconvexas. Véase las aplicaciones en localización dos Santos Gromicho (2013), programación fraccional, y un especial interés en economia, ver Takayama and Akira (1985) y Mas Colell et al. (1995).

4. Por último esta investigación permitira trabajos futuros sobre algoritmos de optimización mucho más eficientes, esto es uno de los motivos que se intenta desarrollar en el presente proyecto.

1.3.1. Objetivos

Objetivo general

Ampliar el algoritmo a su versión inexacta usando cuasi-distancias para funciones convexas y no convexas en un problema de minimización.

Objetivos específicos

- Determinar las condiciones que debe tener el error e^{k+1} para que el algoritmo inexacto converja.
- Extender las propiedades de convergencia de un algoritmo de punto proximal, para minimización de funciones convexas y no convexas.
- Presentar una aplicación a un problema de decisión en economía.

Capítulo 2

MARCO TEÓRICO

2.1. Elementos básicos de análisis en $\mathbb{R}^n$

2.1.1. Sucesiones en $\mathbb{R}^n$

El conjunto $\mathbb{R}^n$ con $n \in \mathbb{N}$ está formada por las n-adas coordenadas

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) | \text{ donde } x_i \in \mathbb{R} \wedge i = 1, 2, \cdots, n\}$$

en este conjunto se define dos operaciones; la suma y el producto por un escalar (en nuestro caso consideramos un escalar, un número real) que hacen de $\mathbb{R}^n$ un espacio vectorial.

Aquí en este capítulo vamos a considerar la norma Euclideana, cuya notación es dado. $||.||$.

Definición 2.1.1. *Una sucesión en el espacio vectorial $\mathbb{R}^n$ es una función de la forma $x : \mathbb{N} \to \mathbb{R}^n$, donde asocia a cada $k \in \mathbb{N}$ un único vector $x^k \in \mathbb{R}^n$ el cual lo llamaremos el k-ésimo término.*
Una sucesión será indicada por $\{x^1, x^2, \ldots, x^k, \ldots\}$, $\{x^k\}_{k\in\mathbb{N}}$ o simplemente $\{x^k\}$.

Ejemplo 2.1.2. *Considerando el espacio $\mathbb{R}^2$ sea la sucesión $\{x^k\}$ dada por $x^k = (k, \frac{1}{k})$ cuyos elementos podemos enlistar como sigue*

$$\left\{(1,1), (2, \frac{1}{2}), (3, \frac{1}{3}), \ldots\right\}.$$

Definición 2.1.3. *Una subsucesión es la limitación de algunos elementos de la sucesión $\{x^k\} \subset \mathbb{R}^n$ a un subconjunto infinito $\mathbb{N}^* = \{k_1 < k_2 < \ldots < k_i < \ldots\} \subset \mathbb{N}$. La subsucesión será indicada por $\{x^k\}_{k\in\mathbb{N}^*}$ o $\{x^{k_i}\}_{i\in\mathbb{N}}$ o sencillamente $\{x^{k_i}\}$.*

Definición 2.1.4. *Una sucesión $\{x^k\} \subset \mathbb{R}^n$ se manifiesta que está acotada si existe un $K > 0$;*

$$||x^k|| \leq K, \text{ para todo } k \in \mathbb{N}.$$

Definición 2.1.5. *(Límite de una sucesión) Dado una sucesión $\{x^k\} \subset \mathbb{R}^n$. Un elemento $x \in \mathbb{R}^n$ es limite de la sucesión $\{x^k\}$ cuando para cada $\epsilon > 0$ existe un $N \in \mathbb{N}$*

tal que $||x^k - x|| < \epsilon$*, siempre que* $k \geq N$.
Si se verifica lo anterior lo denotamos por

$$\lim_{k\to\infty} x^k = x$$

Si la sucesión $\{x^k\} \subset \mathbb{R}^n$ *satisface la definición, diremos que la sucesión es convergente en caso contrario diremos que es divergente.*

Proposición 2.1.6. *Dado la sucesión* $\{x^k\} \subset \mathbb{R}^n$*. Tenemos* $\lim_{k\to\infty} x^k = 0$ *si y solo si* $\lim_{k\to\infty} ||x^k|| = 0$

Demostración. Véase en Lima (2000), página 14. □

Ejemplo 2.1.7. *dado las siguientes sucesiones:*

1. $\{x^k\} = \{k\}$
2. $\{x^k\} = \{\frac{1}{k}\}$
3. $\{x^k\} = \{(-1)^k\}$.

La sucesión 1) no esta acotada pues el conjunto de los números naturales $\mathbb{N}$ *no lo está, para los otros casos las sucesiones son acotadas pues se tiene* $0 \leq x^k \leq 1$*,* $-1 \leq x^k \leq 1$ *en* $k \in \mathbb{N}$ *para 2) y 3). La única sucesión convergente es* $\{x^k\} = \{\frac{1}{k}\}$ *la cual converge a cero cuando* k *tiende al infinito.*

Teorema 2.1.8. *Toda sucesión convergente* $\{x^k\}$ *se encuentra acotada.*

Demostración. Véase en Lima (2000), página 14. □

Teorema 2.1.9. *Dado la sucesión* $\{x^k\} \subseteq \mathbb{R}^n$*. Si* $\lim_{k\to\infty} x^k = a$*, entonces toda subsucesión de* $\{x^k\}$ *converge hacia* a*.*

Demostración. Véase en Benazic (2000), página 22. □

Teorema 2.1.10. *(Bolzano - Weierstrass) Dado cualquier sucesión acotada en* $\mathbb{R}^n$ *esta admite una subsucesión que es convergente.*

Demostración. Véase en Benazic (2000), página 22. □

Teorema 2.1.11. *Toda sucesión de números reales monótona y acotada es convergente*

Demostración. Véase en Hasser et al. (1970), página 467. □

Definición 2.1.12. *Sea* $a \in \mathbb{R}^n$*. Se dice que* a *es un* **punto de acumulación** *de la sucesión* $\{x^k\} \subset \mathbb{R}^n$ *si para* $\{x^{k_i}\} \subset \{x^k\}$ *se tiene* $\lim_{i\to\infty} x^{k_i} = a$*.*

Ejemplo 2.1.13. *Del Ejemplo (2.1.7) la sucesión dada por 1) no tiene puntos de acumulación, En 2) se observa que la sucesión* $\{x^k\} = \{\frac{1}{k}\}$ *es convergente y su único punto de acumulación es el cero. Por último la sucesión* $\{x^k\} = \{(-1)^k\}$ *tiene dos puntos de acumulación que son 1 y -1.*

Teorema 2.1.14. *Sea $\{x^k\} \subset I\!R^n$ una sucesión. Las siguientes afirmaciones son equivalentes*

a) La sucesión $\{x^k\}$ es convergente.

b) La sucesión $\{x^k\}$ es acotada y tiene a lo más un punto de acumulación.

Demostración. Véase en Benazic (2000), página 23. □

Lema 2.1.15. *Considerese las sucesiones $\{\alpha^k\}, \{\lambda^k\}$ y $\{\rho^k\}$ de números reales no negativos que cumplen*

$$\alpha^k \leq (1 + \lambda^k)\alpha^{k-1} + \rho^k$$

con $\sum_{k=1}^{+\infty} \rho^k < +\infty$, $\sum_{k=1}^{+\infty} \lambda^k < +\infty$. Implica que $\{\alpha^k\}$ converge.

Demostración. Véase en Polyak (1987), Lema 2, página 44. □

2.2. Elementos básicos de convexidad

Definición 2.2.1. *Sea $C \subset I\!R^n$ un conjunto no vacío. Se afirma que C es un conjunto convexo si para cualquier $x, y \in C$ se tiene*

$$[x.y] = \{\alpha x + (1 - \alpha)y;\ \alpha \in [0, 1]\} \subset C$$

llamaremos $[x, y]$ el segmento de extremos x e y.

Ejemplo 2.2.2. *Dada la siguiente figura, se observa que (a) no es un conjunto convexo pues el segmento de recta [x, y] no esta contenido en el conjunto, mientras (b) si es un conjunto convexo.*

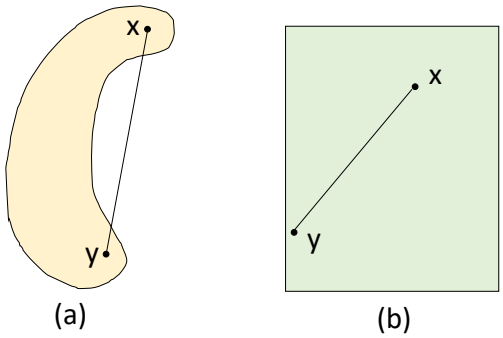


Figura 2.1: (a) conjunto no convexo, (b) conjunto convexo

Definición 2.2.3. *Sea $C \subset \mathbb{R}^n$ no vacío y convexo. Se afirma que $f : C \to \mathbb{R}$ es una función convexa si para $x, y \in C$ y $\lambda \in [0, 1]$ se tiene*

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

Observación: Si para $\lambda \in \langle 0, 1\rangle$ la desigualdad de arriba es estricta entonces se dice que f es una función estrictamente convexa.

Definición 2.2.4. *Una combinación convexa de elementos $x_o, x_1, \ldots, x_p$ de $\mathbb{R}^n$ es una combinación lineal $\sum_{i=0}^{p} \lambda_i x_i$ con λ_i no negativo y $\sum_{i=0}^{p} \lambda_i = 1$.*

Vamos a considerar de aquí en adelante que $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ es la recta real extendida.

Definición 2.2.5. *Dado la función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, Llamaremos dominio efectivo de f, al conjunto*

$$\{x \in \mathbb{R}^n; f(x) < +\infty\}.$$

lo cual es indicado por $domf$.

Definición 2.2.6. *Se dice que $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es una función propia si*

1) $domf \neq \emptyset$,

2) $\forall x \in domf$, se tiene $f(x) > -\infty$.

Las funciones propias $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ son obtenidas tomando un conjunto $C \subset \mathbb{R}^n$ no vacío y una función $f : C \to \overline{\mathbb{R}}$ y poniendo $f(x) = +\infty$ para todo $x \notin C$.

Ejemplo 2.2.7. *Dada la función $f : [-2, +\infty > \to \mathbb{R}$ que esta definida como $f(x) = x^2$. A esta función se le puede extender su dominio a todo $\mathbb{R}$.*

En este caso $\bar{f} : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ definido por

$$\bar{f}(x) = \begin{cases} x^2 & x \geq -2 \\ +\infty & x < -2 \end{cases}$$

Definición 2.2.8. *Se dice que la función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es semicontinua inferior (sci) en $\bar{x} \in domf$ si cualquier sucesión $\{x^k\}$ de $domf$ convergente a $\bar{x}$ se tiene que*

$$\liminf_{k \to \infty} f(x^k) \geq f(\bar{x})$$

donde

$$\sup_{n \in \mathbb{N}} \inf_{k \geq n} f(x^k) = \liminf_{k \to \infty} f(x^k) :$$

Si se tiene que para todo $x \in domf$ la función f es semicontinua inferior entonces diremos que f es semicontinua inferior en $domf$.

Ejemplo 2.2.9. *La siguiente función dada $f : \mathbb{R}^n \to \overline{\mathbb{R}}$*

$$f(x) = \begin{cases} |x-1|+1 & x < 2 \\ 1 & x = 2 \\ +\infty & \textit{en los otros casos} \end{cases}$$

es una función semicontinua inferior.

Como se observa en la gráfica, la función no necesariamente tiene que ser continua en dicho punto.

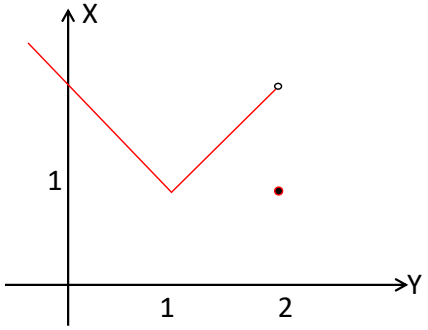


Figura 2.2: Gráfica de f

Definición 2.2.10. *Se dice que la función $f : I\!R^n \to \overline{I\!R}$ es semicontinua superior (scs) en $\bar{x} \in domf$ si cualquier sucesión $\{x^k\} \subset domf$ convergente a $\bar{x}$ se tiene que*

$$\limsup_{k\to\infty} f(x^k) \leq f(\bar{x})$$

donde

$$\limsup_{k\to\infty} f(x^k) := \inf_{n\in\mathbb{N}} \sup_{k\geq n} f(x^k).$$

Definición 2.2.11. *Dada la función $f : I\!R^n \to I\!R$, se afirma lo siguiente::*

- *f es una función positivamente homogenea si $f(\lambda x) = \lambda f(x)$ para todo $\lambda > 0$.*
- *f es subadiditava si $f(x + y) \leq f(x) + f(y)$ para todo $x, y \in I\!R^n$.*

Teorema 2.2.12. *Si se tiene que $f : X \subset I\!R^n \to I\!R$ es continua en $a \in X$. Entonces f es semicontinua inferior en a.*

Demostración. Véase en Rockafellar and Wets (2009), página 13. □

Definición 2.2.13. *$X \subseteq I\!R^n, X \neq \emptyset$, indicamos como función indicatriz de X a*

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

Definición 2.2.14. *Dado $f : I\!R^n \to I\!R$ y $\alpha \in I\!R$. El conjunto*

$$L_f(\alpha) = \{x \in I\!R^n; f(x) \leq \alpha\}.$$

es llamado conjunto de nivel inferior de f.

Definición 2.2.15. *La sucesión $\{x^k\}$ en $X \subset I\!R^n$ es llamada crítica (respecto al conjunto X) si:*

$$\lim_{k\to\infty} x^k = x \in \overline{X} - X \qquad \vee \qquad \lim_{k\to\infty} \| x^k \| = +\infty$$

Se dice que $f : X \subset I\!R^n \to I\!R$ es una función coerciva en X si para todo $\{x^k\}$ sucesión crítica cumple

$$\lim_{k\to\infty} supf(x^k) = +\infty,$$

donde

$$\lim_{k\to\infty} supf(x^k) = \inf_{n\in\mathbb{N}} \sup_{k\geq n} f(x^k)$$

Corolario 2.2.16. *Sea una función $f : C \subset I\!R^n \to I\!R$. Si f es coerciva y semicontinua inferior sobre $C \neq \emptyset$, entonces existe un elemento $c \in C$; $f(c) \leq f(x), \forall x \in C$.*

Demostración. Véase en Papa Quiroz (2019a), página 38. □

Definición 2.2.17. *El epígrafo de $f : I\!R^n \to I\!R$, está formada por el conjunto*

$$epi(f) = \{(x, \lambda) \in I\!R^n \times I\!R; f(x) \leq \lambda\}.$$

Proposición 2.2.18. *Dado un conjunto $C \subset \mathbb{R}^n$ convexo. Una función $f : C \to \mathbb{R}$ es convexa en C, si, y solamente si, $epi(f)$ es un subconjunto convexo e $\mathbb{R}^n \times \mathbb{R}$.*

Demostración. Véase en Papa Quiroz (2019a), página 126. □

Ejemplo 2.2.19. *En la siguiente figura se observa los siguientes conjuntos: el dom f, $L_f(\alpha)$ y el epi(f) de una función extendida $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.*

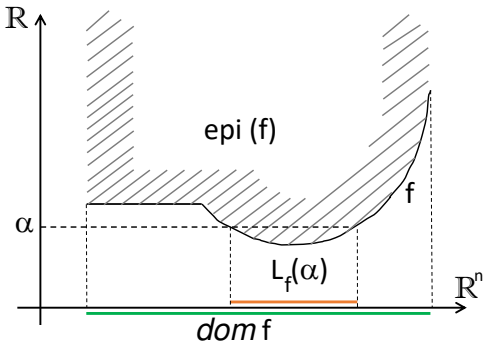


Figura 2.3: Gráfica del conjunto de nivel $L_f(\alpha)$ y el dominio efectivo dom f.

Teorema 2.2.20. *Dada una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, las siguientes proposiciones son equivalentes*

i) La función f es semicontinua inferior sobre $\mathbb{R}^n$

ii) El conjunto $epi(f)$ es cerrado en $\mathbb{R}^n \times \mathbb{R}$

iii) Para cualquier $\alpha \in \mathbb{R}$ el conjunto $L_f(\alpha)$ es cerrado en $\mathbb{R}^n$

Demostración. Véase en Rockafellar and Wets (2009), página 9. □

Ejemplo 2.2.21. *Dada la función*

$$f(x) = \begin{cases} \frac{1}{x-x_0} & x > x_0 \\ +\infty & \textit{en otros casos} \end{cases}$$

esta es una función semicontinua inferior cuyo $dom(f) = (x_0, +\infty)$

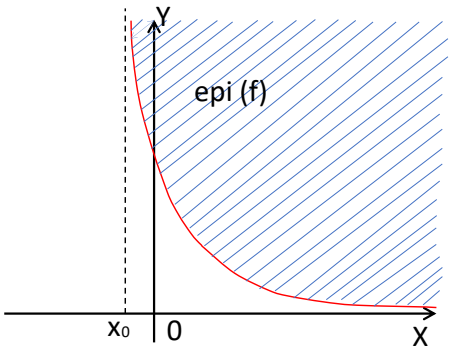


Figura 2.4: Gráfica de $epi(f)$

Proposición 2.2.22. *Sea una función $f : C \subset \mathbb{R}^n \to \mathbb{R}$ convexa y $g : C \subset \mathbb{R}^n \to \mathbb{R}$ una función estrictamente convexa, entonces la suma $f + \lambda g$ es estrictamente convexa, con $\lambda > 0$.*

Demostración. Véase en Papa Quiroz (2019a), ejercicio propuesto página 152. □

Teorema 2.2.23. *Dado $C \subset \mathbb{R}^n$ convexo y f una función convexa en C. Entonces el conjunto de minimizadores de f sobre C es convexo.*

Demostración. Véase en Rockafellar and Wets (2009), Teorema 2.6. □

Teorema 2.2.24. *Sea $C \subset \mathbb{R}^n$ convexo y f una función estrictamente convexa en C. Se afirma que no existe más de un minimizador global de f sobre C.*

Demostración. Véase en Bazaraa et al. (2013), página 126. □

Teorema 2.2.25. *Sea $S \subset \mathbb{R}^n$ convexo, no vacío y $f : S \to \mathbb{R}$ una función convexa. Entonces para cada $\bar{x} \in intS$, existe un vector ξ tal que*

$$f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para cada } x \in S;$$

ξ es llamado subgradiente de f en $\bar{x}$.

Demostración. Véase la prueba en Bazaraa et al. (2013), página 107. □

Teorema 2.2.26. *Sea $C \subset \mathbb{R}^n$ un conjunto abierto convexo y $f : C \to \mathbb{R}$ una función diferenciable en C.*
Las siguientes propiedades son equivalentes

a) La función f es convexa en C.

b) $f(y) \geq f(x) + \langle \bigtriangledown f(x), y - x \rangle$, para todo $x, y \in C$.

c) $\langle \bigtriangledown f(y) - \bigtriangledown f(x), y - x \rangle \geq 0$, $\forall x, y \in C$.

Demostración. Véase en Barbolla et al. (2001), página 22. □

Corolario 2.2.27. *Sea $C \subset \mathbb{R}^n$ un conjunto abierto convexo y $f : C \to \mathbb{R}$ una función diferenciable en C.*
Las siguientes propiedades son equivalentes

a) La función f es estrictamente convexa en C.

b) $f(y) > f(x) + \langle \bigtriangledown f(x), y - x \rangle$, para todo $x, y \in C$ tal que $x \neq y$

c) $\langle \bigtriangledown f(y) - \bigtriangledown f(x), y - x \rangle > 0 \quad \forall x, y \in C$ tal que $x \neq y$

Demostración. Véase en Barbolla et al. (2001), página 22. □

Teorema 2.2.28. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa y diferenciable en $\bar{x}$. Si $\bigtriangledown f(\bar{x}) = 0$ entonces $\bar{x}$ es un mínimo global de f en $\mathbb{R}^n$*

Demostración. Véase en Papa Quiroz (2019a), página 155. □

Teorema 2.2.29. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en $\bar{x}$. Si $\bar{x}$ es minimizador local de f. Entonces $\bigtriangledown f(\bar{x}) = 0$.*

Demostración. Véase en Papa Quiroz (2019a), página 50. □

2.3. Funciones Lipschizianas

Definición 2.3.1. *Sea $X \neq \emptyset$ un subconjunto de $\mathbb{R}^n$. Una función $f : X \to \mathbb{R}$ satisface la condición de Lipschitz sobre X si existe un número positivo $K > 0$ tal que*

$$|f(x) - f(y)| \leq K||x - y||, \quad \forall x, y \in X.$$

Definición 2.3.2. *Dado $\emptyset \neq X \subset \mathbb{R}^n$. La función $f : X \to \mathbb{R}$ se dice localmente lipschiziana en $\bar{x} \in X$ si existe un $\epsilon > 0$ y $K_{\bar{x}} > 0$ tal que*

$$|f(x) - f(y)| \leq K_{\bar{x}}||x - y||, \quad \forall x, y \in B_\epsilon(\bar{x})$$

Teorema 2.3.3. *Dado $X \subset \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$ una función convexa. Si $x^0 \in int(X)$, entonces f es localmente Lipschitziana en x^0.*

Demostración. Véase en Papa Quiroz (2019a), página 135. □

Corolario 2.3.4. *Dado $X \subset \mathbb{R}^n$ un conjunto convexo y $f : X \to \mathbb{R}$ una función convexa. La función f es continua en $\bar{x}$ siempre que $\bar{x} \in int(X)$.*

Demostración. Véase en Papa Quiroz (2019a), página 136. □

Corolario 2.3.5. *Dado $f : \mathbb{R}^n \to \mathbb{R}$ convexa, entonces la función f es continua en $\mathbb{R}^n$.*

Demostración. Véase en Papa Quiroz (2019a), página 136. □

Definición 2.3.6. *Dado $G \subset \mathbb{R}^n$ no vacío. La función distancia $d_G : \mathbb{R}^n \to \mathbb{R}$ para el conjunto G es definido por la siguiente expresión*

$$d_G(x) = inf\{||x - c||; c \in G\}.$$

Teorema 2.3.7. *La función d_G es Lipschitz con constante $K = 1$, en otras palabras*

$$|d_G(x) - d_G(y)| \leq ||x - y|| \text{ para todo } x, y \in \mathbb{R}^n.$$

Si el conjunto G es convexo entonces la función d_G es convexa.

Demostración. Véase en Makela and Neittaanmaki (1992), página 23. □

2.4. Subdiferenciabilidad

Los problemas que surgen al optimizar una función que está determinada sobre un dominio fue estudiada por muchos investigadores desde el siglo pasado, esta inquietud dio desarrollo al cálculo diferencial quien dio ciertas caracteristicas a las funciones diferenciables que bajo ciertas condiciones se podía determinar la solución de un problema de minimización. Pero con el desarrollo de la ciencia y la tecnología se dieron cuenta que algunos fenómenos naturales tienen en su modelo matemático funciones que no son diferenciables esto motivo a reducir algunas exigencias matemáticas sobre el punto donde la función alcanza el mínimo.

Las subderivadas y las subdiferenciables nació como una necesidad para solucionar problemas de minimización cuyas funciones no son diferenciables ni continuas.

2.4.1. Subdiferenciabilidad Clásica

Definición 2.4.1. *Dado $f : \mathbb{R}^n \to \mathbb{R}$. La derivada direccional de la función f en x en la dirección $v \in \mathbb{R}^n$, se define por*

$$f'(x, v) = \lim_{t \to 0^+} \frac{f(x + tv) - f(x)}{t}$$

Cuando f es diferenciable en x, se tiene que la derivada direccional existe en todas las direcciones $v \in \mathbb{R}^n$. Así $f'(x, v)$ es una funcional lineal de v, y se tiene

$$f'(x, v) = \langle \bigtriangledown f(x), v \rangle.$$

Observación 2.4.2.

1. *La Definición 2.4.1 difiere de la definición de la derivada direccional clásica, pues en este caso se observa que t se aproxima a 0 por derecha, esta definición se considera por que vamos hacer nuestro estudio sobre funciones que no necesariamente son diferenciables.*
2. *La expresión $t \to 0^+$ nos afirma que $\frac{f(x + tv) - f(x)}{t}$ es evaluado para los números t positivos que se aproximan a cero.*

Definición 2.4.3. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa, el subdiferencial de f en $x \in \mathbb{R}^n$ es el conjunto*

$$\partial_c f(x) = \{\xi \in \mathbb{R}^n; f(x') \geq f(x) + \xi^T(x' - x), \forall x' \in \mathbb{R}^n\}.$$

A continuación mencionamos la relación que se presenta en el subdiferencial y la derivada direccional ordinaria.

Teorema 2.4.4. *Dado $f : \mathbb{R}^n \to \mathbb{R}$ convexa con $K > 0$ constante de Lipschitz, entonces se verifica la siguientes condiciones*

a) $f'(x, v)$ existe y satisface

$$f'(x, v) = \inf_{t \to 0^+} \frac{f(x + tv) - f(x)}{t}.$$

b) La función $v \to f'(x, v)$ es positivamente homogenea y subaditiva sobre $\mathbb{R}^n$ con

$$|f'(x, v)| \leq K||v||$$

c) $f'(x, v)$ es semicontinua superior como función de (x, v) y Lipschitziana con constante K como una función de v sobre $\mathbb{R}^n$.

d) $f(x, v) \geq -f'(x, -v)$.

Demostración. Véase en Makela and Neittaanmaki (1992), página 10. □

Teorema 2.4.5. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ función convexa. Para cualquier x en $\mathbb{R}^n$ se tiene*

i. $f'(x,v) = \max\{\xi^T.v; \xi \in \partial_c f(x)\}$. *Para todo* $v \in \mathbb{R}^n$.

ii. $\partial_c f(x) = \{\xi \in \mathbb{R}^n; f'(x,v) \geq \xi^T.v$ *para todo* $v \in \mathbb{R}^n\}$

iii. $\partial_c f(x)$ *es un conjunto no vacío, convexo y compacto, tal que* $\partial_c f(x,v) \subset B(0,K)$ *con* K *constante de Lipschitz de* f *en* x.

iv. La aplicación conjunto $\partial_c f : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ *es semicontinua superior, osea, si* $y_i \to y$ *y* $\varepsilon_i \in \partial_c f(y_i)$ *para cada* $i \in \mathbb{N}$, *entonces cada punto de acumulación* ε *de* $\{\varepsilon_i\}$ *está en* $\partial_c f(y)$.

Aquí $\mathcal{P}(A)$ *es el conjunto de todos los subconjunto de A.*

Demostración. Véase en Makela and Neittaanmaki (1992), página 14. □

Proposición 2.4.6. *Siendo la función* $f : \mathbb{R}^n \to \mathbb{R}$ *convexa y diferenciable en* $x \in \mathbb{R}^n$ *se tiene*

$$\partial_c f(x) = \{\bigtriangledown f(x)\}.$$

Demostración. Véase en Makela and Neittaanmaki (1992), página 16. □

Ejemplo 2.4.7. *Dado una función* $f(x) = |x|$ *definida sobre* $\mathbb{R}$, *se obtendrá el* $\partial_c f(x)$
Esta función es definida de la forma equivalente

$$f(x) = \begin{cases} x & ; x \geq 0 \\ -x & ; x < 0 \end{cases}$$

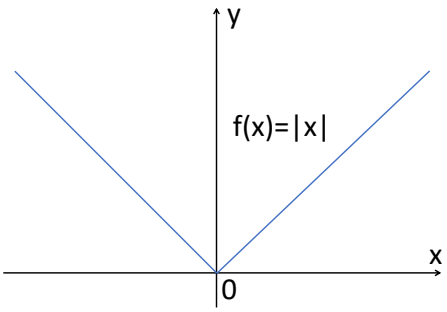


Figura 2.5: La función valor absoluto.

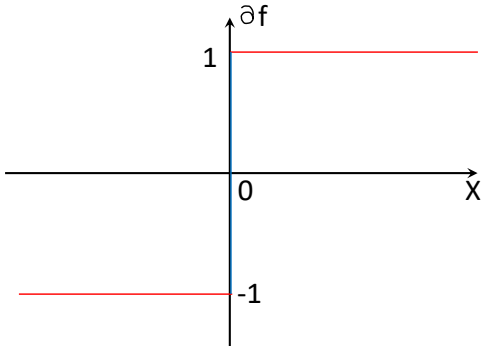


Figura 2.6: Gráfica de el subdiferencial de la función f.

El desarrollo se puede ver en Illesca Cangalaya (2016) página 28, donde el subdiferencial es dado por

$$\partial_c f(x) = \begin{cases} 1 & ; x > 0 \\ 0 & ; [-1,1] \\ -1 & ; x < 0 \end{cases}$$

2.4.2. Subdiferenciabilidad Generalizada

A continuación vamos a definir la derivada direccional generalizada en el sentido de Clarke.

Definición 2.4.8. *Dado una función $f : \mathbb{R}^n \to \mathbb{R}$ localmente Lipschitz en un punto $x \in \mathbb{R}^n$. Se define la derivada direccional de f en x en la dirección $w \in \mathbb{R}^n$ como*

$$f^\circ(x, w) = \lim_{\substack{y \to x \\ t \to 0^+}} \sup \frac{f(y + tw) - f(y)}{t}$$

Este es un límite que existe pues la expresión $\dfrac{f(y + tw) - f(y)}{t}$ está acotada superiormente por $K||w||$ (para un y suficientemente cercano a x y t cercano a cero) esto es por la condición de Lipschitz.

Teorema 2.4.9. *Dado $f : \mathbb{R}^n \to \mathbb{R}$ localmente Lipschitz en x y $K > 0$. Entonces*

i. $v \to f^\circ(x, v)$ *es positivamente homogenea y subaditiva en $\mathbb{R}^n$ con*

$$|f^\circ(x, v)| \leq K||v||.$$

ii. El valor $f^\circ(x, v)$ es semicontinua superior como función de (x, v) y Lipschitz con $K > 0$ como una función de v sobre $\mathbb{R}^n$.

iii. $(-f)^\circ(x, v) = f^\circ(x, -v)$.

Demostración. Véase en Makela and Neittaanmaki (1992), teorema 3.1.2. □

Definición 2.4.10. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una frnción localmente Lipschitz en $\mathbb{R}^n$. Entonces el subdiferencial de f en x es el conjunto*

$$\partial f^\circ(x) = \{\xi \in \mathbb{R}^n : f^\circ(x, v) \geq \xi^T.v \text{ para todo } v \in \mathbb{R}^n\}.$$

Teorema 2.4.11. *Dado $f : \mathbb{R}^n \to \mathbb{R}$ localmente Lipschitz en x con $K > 0$. Entonces*

i. El conjunto $\partial f^\circ(x)$ es distinto del vacío, convexo y compacto tal que $\partial f^\circ(x) \subset B(0, K)$.

ii. $f^\circ(x, v) = \text{máx}\{\xi^T.v | \xi \in \partial f^\circ(x)\}$.

iii. La $\partial f^\circ(.) : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ es semicontinua superior.

Demostración. Véase en Makela and Neittaanmaki (1992), teorema 3.1.4. □

Teorema 2.4.12. *Si se tiene que la función f es convexa, entonces*

a) $f'(x, v) = f^\circ(x, v)$ *para cualquier $v \in \mathbb{R}^n$.*

b) $\partial_c f(x) = \partial f^\circ(x)$.

Demostración. Véase en Makela and Neittaanmaki (1992), teorema 3.1.8. □

Ejemplo 2.4.13. *Dada la función* $g(x) = x^2.sen(\frac{1}{x})$. *Esta función es diferenciable en todo* $\mathbb{R}$ *excepto en el cero. Vamos a determinar el subdiferencial de Clarke en* $x = 0$.

Por la Definición 2.4.8 se tiene

$$g^\circ(0, v) = \lim_{\substack{y\to 0\\ t\to 0^+}} \sup \frac{(y+tv)^2 sen\frac{1}{(y+tv)} - (y)^2 sen\frac{1}{y}}{t}$$

y haciendo uso de la Definición 2.4.10

$$d \in \partial g^\circ(0) \Leftrightarrow v.d \leq g^\circ(x, v) = \lim_{\substack{x\to y\\ t\to 0^+}} \sup \frac{g(y+tv) - g(y)}{t} \text{ para todo } v \in \mathbb{R}$$

$$v.d \leq \lim_{\substack{y\to 0\\ t\to 0^+}} \sup \frac{(y+tv)^2 sen\frac{1}{(y+tv)} - (y)^2 sen\frac{1}{y}}{t}$$

como el límite

$$\lim_{y\to 0} \sup \frac{(y+tv)^2 sen\frac{1}{(y+tv)} - (y)^2 sen\frac{1}{y}}{t} = \frac{(tv)^2 sen(\frac{1}{tv})}{t}$$

existe y es único, entonces el límite superior de dicho límite tiene el mismo valor, es decir

$$v.d \leq \lim_{t\to 0^+} \sup \frac{(tv)^2 sen(\frac{1}{tv})}{t}$$

$$v.d \leq \lim_{t\to 0^+} \sup \frac{v.sen(\frac{1}{tv})}{\frac{1}{tv}}$$

$$v.d \leq v \text{ para todo } v \in \mathbb{R}^n$$

Ahora si

- $v \geq 0$ *entonces se tiene que* $d \leq 1$.
- $v < 0$ *entonces se tiene que* $d \geq -1$.

De ambas condiciones se concluye que $\partial f^\circ(0) = [-1, 1]$.

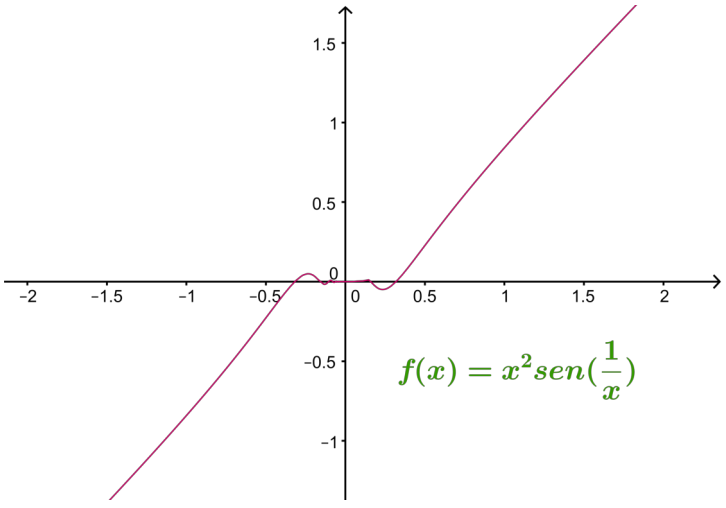


Figura 2.7: La gráfica de la función $f(x) = x^2.sen(\frac{1}{x})$

2.4.3. Subdiferenciabilidad de Fréchet

Definición 2.4.14. *Sea $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función de valores extendidos, f es propia semicontinua inferior y $x \in \mathbb{R}^n$*

(1) El conjunto $\widehat{\partial} f(x)$ de vectores $r \in \mathbb{R}^n$, tal que

$$f(x) + \langle r, y - x\rangle + \varphi(||y - x||) \leq f(y) \text{ donde } \lim_{\substack{x \neq y \\ y \to x}} \frac{\varphi(||y - x||)}{||y - x||} = 0$$

es llamado el subdiferencial de Fréchet de f en x.
o equivalentemente

$$\widehat{\partial} f(x) = \begin{cases} \left\{ s \in \mathbb{R}^n : \liminf\limits_{\substack{x \neq y \\ y \to x}} \frac{f(y) - f(x) - \langle s, y - x\rangle}{||y - x||} \geq 0 \right\} & ; \text{ si } x \in dom f \\ \emptyset & ; si\ x \notin dom f \end{cases}$$

(2) el subdiferencial limitante de f en $x \in \mathbb{R}^n$, $\partial f(x)$ es definido como sigue

$$\partial f(x) := \left\{ s \in \mathbb{R}^n : \exists\, x^k \to x, f(x^k) \to f(x), s^k \in \widehat{\partial} f(x^k) \to s \right\}.$$

Teorema 2.4.15. *(relación entre subgradiente) Dado una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ y un punto $\bar{x}$ donde la función $f(x)$ es finito, los conjuntos $\partial f(\bar{x})$ y $\widehat{\partial} f(x)$ son cerrados, el conjunto $\widehat{\partial} f(\bar{x})$ es convexo y $\widehat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$.*

Demostración. Véase en Rockafellar and Wets (2009), página 302. □

Proposición 2.4.16. *(Regla generalizada de Fermat) Si la función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ tiene un elemento mínimo local en $\bar{x} \in dom f$ en consecuencia $0 \in \widehat{\partial} f(\bar{x})$.*

Demostración. Como $\bar{x} \in dom f$ es un mínimo local, entonces existe $\epsilon > 0$ tal que

$$f(\bar{x}) \leq f(x),\ \forall x \in B_\epsilon(\bar{x}) \cap dom f$$

donde $B_\epsilon(\bar{x}) = \{x \in \mathbb{R}^n |\ ||x - \bar{x}|| < \epsilon\}$.

Consideremos una función constante $\varphi(||x - \bar{x}||) = 0$, entonces se tiene

$$f(x) \geq f(\bar{x}) + \langle 0, x - \bar{x}\rangle + \varphi(||x - \bar{x}||) \text{ para todo } x \in B_\epsilon(\bar{x}) \cap dom f$$

es decir,

$$\liminf_{\substack{x \neq y \\ y \to x}} \frac{f(x) - f(\bar{x}) - \langle 0, x - \bar{x}\rangle}{||x - \bar{x}||} \geq 0 \text{ donde } \lim_{\substack{x \neq y \\ y \to x}} \frac{\varphi(||x - \bar{x}||)}{||x - \bar{x}||} = 0$$

Por lo tanto $0 \in \widehat{\partial} f(\bar{x})$. □

Proposición 2.4.17. *(subgradiente versus gradiente) Son equivalentes las siguientes expresiones:*

(a) Si f es diferenciable en $\bar{x}$, entonces $\widehat{\partial} f(\bar{x}) = \{\bigtriangledown f(\bar{x})\}$, así $\bigtriangledown f(\bar{x}) \in \partial f(\bar{x})$.

(b) Si la función f es diferenciable sobre una vecindad de $\bar{x}$, entonces $\partial f(\bar{x}) = \bigtriangledown f(\bar{x})$.

(c) Si $h = g + f$ con g es finito en $\bar{x}$ y f es suave sobre una vecindad de $\bar{x}$, entonces $\widehat{\partial} h(\bar{x}) = \widehat{\partial} g(\bar{x}) + \bigtriangledown f(\bar{x})$, $\partial h(\bar{x}) = \partial g(\bar{x}) + \bigtriangledown f(\bar{x})$.

Demostración. Véase la prueba en Rockafellar and Wets (2009), Ejercicio 8.8 página 304. □

Proposición 2.4.18. *Si f_1 es una función localmente Lipschitz continua en $\bar{x}$, f_2 es semicontinua inferior y propia con $f_2(\bar{x})$ finito, entonces*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Demostración. Véase en Rockafellar and Wets (2009), ejercicio 10.10 página 431. □

Proposición 2.4.19. *Sea $f : I\!R^n \to I\!R \cup \{\pm\infty\}$una función localmente semicontinua inferior en $\bar{x}$ con $f(\bar{x})$. Entonces las siguientes condiciones son equivalentes*

(1) f es una función localmente Lipschitz continua en $\bar{x}$.

(2) Se tiene la aplicación $\widehat{\partial} : x \to \widehat{\partial} f(x)$ es localmente acotada en $\bar{x}$.

(3) La aplicación $\partial f(x) : x \to \partial f(x)$ es localmente acotada en $\bar{x}$.

Demostración. Véase en Rockafellar and Wets (2009), teorema 9.3 página 358. □

Ejemplo 2.4.20. *Dada la función*

$$f(x) = \begin{cases} x^2.sen(\frac{1}{x}) & ; \text{ cuando } x \neq 0 \\ 0 & ; \text{ cuando } x = 0 \end{cases}$$

vamos a determinar el subdiferencial de Fréchet y el subdiferencial limitante. Como sabemos la función f es continua en todo $I\!R$, y su derivada cuando $x \neq 0$

$$f'(x) = 2xsen\left(\frac{1}{x}\right) - cos\left(\frac{1}{x}\right)$$

Por otro lado,

$$\frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x} = xsen\left(\frac{1}{x}\right)$$

Luego

$$f'(0) = \lim_{x \to 0} xsen\left(\frac{1}{x}\right) = 0$$

Entonces por la Proposición 2.4.17 se tiene que el subdiferencial de Fréchet es

$$\begin{array}{cc} \widehat{\partial} f(x) = 2xsen\left(\frac{1}{x}\right) - cos\left(\frac{1}{x}\right) & \text{cuando } x \neq 0 \\ \widehat{\partial} f(0) = 0 & \text{cuando } x = 0 \end{array}$$

Por otro lado, la función f es discontinua en el punto 0. Entonces por la Proposición 2.4.17 se tiene que el subdiferencial limitante

$$\partial f(x) = 2xsen\left(\tfrac{1}{x}\right) - cos\left(\tfrac{1}{x}\right) \quad \textit{cuando } x \neq 0$$

Y para $\partial f(0) = [-1, 1]$ *cuando* $x = 0$.
En efecto, sea $\{x^k\}$ *una sucesión que converge a cero y* $f(x^k) \to f(0)$. *Entonces por la Definición 2.4.14 (2) se tiene*

$$s^k \in \widehat{\partial} f(x^k) = 2xsen\left(\frac{1}{x^k}\right) - cos\left(\frac{1}{x^k}\right)$$

tomando límite

$$\lim_{k\to\infty} s^k = \lim_{k\to\infty} 2x^k sen\left(\frac{1}{x^k}\right) - cos\left(\frac{1}{x^k}\right)$$

$$\lim_{k\to\infty} s^k = -\lim_{k\to\infty} cos\left(\frac{1}{x^k}\right)$$

entonces $s \in [-1, 1]$ *y esto implica que* $\partial f(0) = [-1, 1]$.

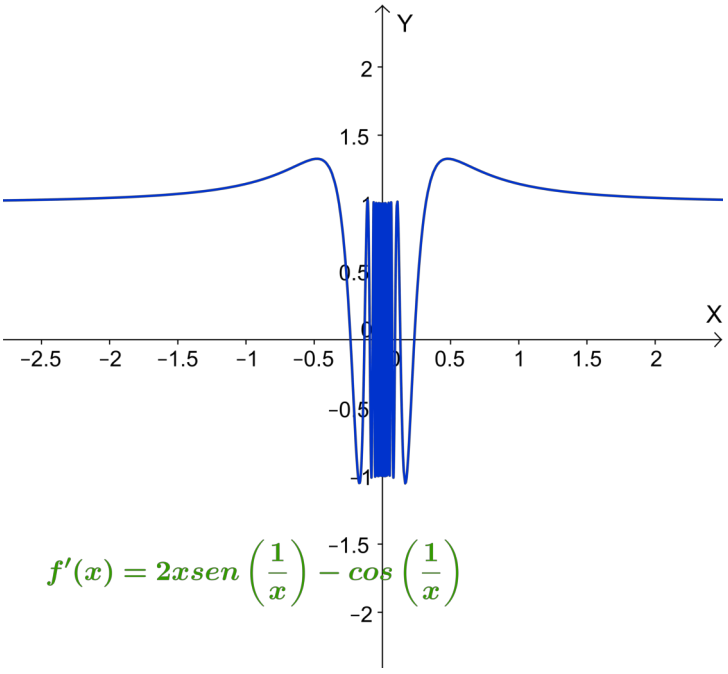


Figura 2.8: La derivada de la función $f(x) = x^2 sen(1/x)$ con $x \neq 0$

Capítulo 3

MÉTODO DEL PUNTO PROXIMAL

Este capítulo se desarrolla tomando como referencia a los trabajos realizados por Iusem (1995) e Illesca Cangalaya (2016).

3.1. Método del punto proximal clásico

Dada una función $f : \mathbb{R}^n \to \mathbb{R}$ convexa y acotada inferiormente. Vamos a considerar el siguiente problema

$$(P) \begin{cases} min f(x) \\ s.a \\ x \in \mathbb{R}^n \end{cases} \tag{3.1}$$

Observación 3.1.1. *Sin algunos supuestos al problema (3.1) no se podría garantizar la existencia de soluciones. Incluso si la función f es diferenciable no se garantiza la existencia de un mínimo en f.*

Ejemplo 3.1.2. *La función $f(x) = 2^x$ es convexa, diferenciable y es acotada inferiormente, pero aún así no existe un punto de mínimo.*

Sea f una función acotada inferiormente y convexa, $g : \mathbb{R}^n \to \mathbb{R}$ una función estrictamente convexa y coerciva. El problema (3.1) puede no tener solución, si embargo plantearemos un problema que si tiene solución (llamado problema regularizado).

$$\min_{x \in \mathbb{R}^n} f(x) + \lambda g(x) \quad \text{con } \lambda > 0. \tag{3.2}$$

Esto es verdad pues en la Proposición 2.2.22 se afirma que $f + \lambda g$ es una función estrictamente convexa (usando el hecho que f es convexa) y coerciva, bajo la suposición que f sea sci (semicontinua inferior) se reduce el problema a un conjunto compacto, combinando el Corolario 2.2.16 y el Teorema 2.2.24 se garantiza la existencia y unicidad de la solución (3.2) cuya solución única, que denotamos por $x(\lambda)$. Bajo algunas hipótesis se prueba que $\lim_{\lambda \to 0} x(\lambda)$ existe y resuelve (3.1). El problema de este enfoque se debe a que, aunque $f + \lambda g$ es una función estrictamente convexa y coerciva para cualquier λ positivo, para un λ pequeño, esta función será numéricamente casi tan mal portado como f.
Para vencer la dificultad mencionada, será deseable desarrollar un método de regularización que no requiera que el parámetro de regularización λ se aproxime a cero.

3.1.1. Algoritmo de punto proximal Clásico

El algoritmo de punto proximal va generar una sucesión de puntos $\{x^k\} \subset I\!R^n$ determinados de la siguiente forma:

Inicio: Tomar un punto arbitrario $x^0 \in I\!R^n$
Iteración: Para $k = 0, 1, \ldots,$ dado x^k, obtener un punto x^{k+1} satisfaciendo

$$x^{k+1} = argmin_{x \in I\!R^n}\{f(x) + \lambda_k \parallel x - x^k \parallel^2\} \tag{3.3}$$

cuyo parámetro λ_k es un número real que satisface

$$0 < \lambda_k < \tilde{\lambda} \tag{3.4}$$

Detiene: Si $x^k = x^{k+1}$ o $0 \in \partial f(x^{k+1})$, entonces se detiene el paso iterativo, en otro caso $k \leftarrow k + 1$ y retornar al paso de Iteración.

Bajo algunas hipótesis dadas al problema (P) demostraremos que la función $f_k(x) = f(x) + \lambda_k \parallel x - x^k \parallel^2$ está bien definida, es decir existe la sucesión $\{x^k\}$ para todo $k \in \mathbb{N}$.

Lema 3.1.3. *Dado una función $g : I\!R^n \to I\!R$ convexa y limitada inferiormente. Entonces, la función $g_k(x) = g(x) + \lambda_k \parallel x - x^k \parallel^2$ es coerciva, con $\lambda_k > 0$ para todo $k \in I\!N$.*

Demostración. En este caso consideremos una sucesión crítica $\{x^m\}$ en $I\!R^n$, por la Definición 2.2.15 vamos a tener dos posibilidades

$$\lim_{k\to\infty} ||x^m|| = +\infty \quad \text{o} \quad \lim_{k\to\infty} x^m = x \in \overline{I\!R^n} - I\!R^n$$

como se sabe que $\overline{\mathbb{R}^n} - \mathbb{R}^n = \emptyset$ entonces tenemos que asumir que

$$\lim_{k\to\infty} ||x^m|| = +\infty.$$

Por otra parte, como $k \in \mathbb{N}$ es un número fijo y $||x^m|| \to +\infty$ cuando $m \to +\infty$, entonces se tiene que existe $N_0 \in \mathbb{N}$ tal que se verifica

$$||x^m - x^k|| > 1, \text{ para } m > N_0.$$

Luego $||x^m - x^k|| < ||x^m - x^k||^2$.
Esto implica

$$||x^m - x^k|| \leq ||x^m - x^k||^2$$

$$| \, ||x^m|| - ||x^k|| \, | \leq ||x^m - x^k|| \leq ||x^m - x^k||^2$$

$$| \, ||x^m|| - ||x^k|| \, | \leq ||x^m - x^k||^2.$$

Tomando límite a ambos lados de la desigualdad

$$| \lim_{m\to\infty} ||x^m|| - ||x^k|| \, | \leq \lim_{m\to\infty} ||x^m - x^k||^2$$

$$| +\infty - ||x^k|| | \leq \lim_{m\to\infty} ||x^m - x^k||^2$$
$$+\infty \leq \lim_{m\to\infty} ||x^m - x^k||^2$$
$$\lim_{m\to\infty} ||x^m - x^k||^2 = +\infty$$

Como $\lambda_k > 0$, para todo $k \in \mathbb{N}$

$$\lim_{m\to\infty} \lambda_k ||x^m - x^k||^2 = +\infty. \tag{3.5}$$

Luego
Por ser g limitada inferiormente

$$\exists M \in I\!R / \quad M \leq g(x); \quad \forall x \in I\!R^n.$$

Para algún $k \in \mathbb{N}$, se tiene

$$M + \lambda_k ||x^m - x^k||^2 \leq g(x) + \lambda_k ||x^m - x^k||^2$$
$$\lim_{m\to\infty} \left(M + \lambda_k ||x^m - x^k||^2\right) \leq \lim_{m\to\infty} \left(g(x^m) + \lambda_k ||x^m - x^k||^2\right).$$

De (3.5)

$$M + (+\infty) \leq \lim_{m\to\infty} \left(g(x^m) + \lambda_k ||x^m - x^k||^2\right)$$
$$+\infty \leq \lim_{m\to\infty} \left(g(x^m) + \lambda_k ||x^m - x^k||^2\right)$$
$$\lim_{m\to\infty} \left(g(x^m) + \lambda_k ||x^m - x^k||^2\right) = +\infty$$
$$\lim_{m\to\infty} g_k(x^m) = +\infty.$$

Por lo tanto f_k es coerciva para todo $k \in \mathbb{N}$. □

Lema 3.1.4. *Dado una función $g : I\!R^n \to I\!R$ convexa y limitada inferiormente entonces la función $g_k(x) = g(x) + \lambda_k \parallel x - x^k \parallel^2$ es continua para todo $k \in \mathbb{N}$.*

Demostración. Por hipótesis g es convexa, $I\!R^n$ es un conjunto convexo y abierto, por el Teorema 2.3.3 se tiene que g es una función continua en $I\!R^n$. La función $h(x) = ||x - x^k||$ es continua en todo $I\!R^n$ lo cual implica que $\lambda_k ||x - x^k||^2$ es continua.
De lo afirmado en la parte superior se concluye que g_k es continua para todo $I\!R^n$. □

Lema 3.1.5. *Probar que la función $g(x) = ||x - x^k||^2$ es estrictamente convexa.*

Demostración. Sea dado $x \neq y$

$$||x - y||^2 > 0$$
$$||x - y||^2 > 0$$
$$2\langle x - y, x - y\rangle > 0$$
$$\langle 2x - 2y, x - y\rangle > 0$$
$$\langle 2(x - x^k) - 2(y - x^k), x - y\rangle > 0$$
$$\langle \bigtriangledown g(x) - \bigtriangledown g(y), x - y\rangle > 0$$

Por Corolario 2.2.27

se cumple que $g(x)$ es estrictamente convexo.

□

Teorema 3.1.6. *Dado una función $g : \mathbb{R}^n \to \mathbb{R}$ convexa y acotada inferiormente. Si la sucesión $\{x^k\}$ es generada por APP. se tiene entonces que la sucesión $\{x^k\}$ está bien definida.*

Demostración. Mostraremos que existe el mínimo de la siguiente función

$$g_k(x) = g(x) + \lambda_k \parallel x - x^k \parallel^2$$

Del Lema 3.1.3 y del Lema 3.1.4 se tiene g_k es continua y coerciva en todo $\mathbb{R}^n$. Del Corolario 2.2.16 se tiene la existencia de un punto que es un mínimo global de g_k en $\mathbb{R}^n$.
Ahora mostraremos que existe un único mínimo. Como g es convexa y del Lema 3.1.5

$\lambda_k||x - x^k||$ es una función estrictamente convexa.

Entonces por la Proposición 2.2.22

$$g(x) + \lambda_k||x - x^k||^2$$

es estrictamente convexa.
Luego por el Teorema 2.2.24

$$g(x) + \lambda_k||x - x^k||^2$$

tiene un único mínimo. □

3.2. Análisis de la convergencia del algoritmo de punto proximal

En esta sección se va exponer la convergencia de la sucesión generada por el APP para funciones de tipo convexa.

Definición 3.2.1. *Sea $\{y^k\} \subset \mathbb{R}^n$ sucesión. Se dice que es Féjer convergente a un conjunto $U \subset \mathbb{R}^n$ si*

$$||y^{k+1} - u|| \leq ||y^k - u||, \text{ para todo } k \in \mathbb{N} \text{ y para todo } u \in U.$$

Lema 3.2.2. *Dado una función $f : \mathbb{R}^n \to \mathbb{R}$ convexa y continuamente diferenciable. Si el conjunto U de minimizadores de la función f es no vacío. Se concluye que la sucesión $\{y^k\}$ es Féjer convergente al conjunto U, es decir*

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

Demostración. Sea un elemento $\bar{y} \in U$, entonces

$$\begin{aligned} ||y^k - y^{k+1} &+ y^{k+1} - \bar{y}||^2 = ||y^k - \bar{y}||^2 \\ ||y^k - y^{k+1}||^2 + ||y^{k+1} &- \bar{y}||^2 + 2\langle y^k - y^{k+1}, y^{k+1} - \bar{y}\rangle = \\ ||y^k - \bar{y}||^2 \dashv ||y^k - y^{k+1}||^2 &- ||y^{k+1} - \bar{y}||^2 = 2\langle y^k - y^{k+1}, y^{k+1} - \bar{y}\rangle. \end{aligned} \quad (3.6)$$

Por otro lado la función

$$f_k(y) = f(y) + \lambda_k ||y - y^k||^2$$

es convexa y continuamente diferenciable pues por hipótesis f es convexa y continuamente diferenciable

$$\bigtriangledown f_k(y) = \bigtriangledown f(y) + 2\lambda_k(y - y^k).$$

Como y^{k+1} es el mínimo de f_k, entonces por el Teorema 2.2.29

$$\begin{aligned} \bigtriangledown f_k(y^{k+1}) &= \bigtriangledown f(y^{k+1}) + 2\lambda_k(y^{k+1} - y^k) = 0 \\ \bigtriangledown f(y^{k+1}) &= 2\lambda_k(y^k - y^{k+1}) \\ y^k - y^{k+1} &= \frac{1}{2\lambda_k} \bigtriangledown f(y^{k+1}) \end{aligned} \quad (3.7)$$

Reemplazando (3.7) en (3.6)

$$\begin{aligned} ||y^k - \bar{y}||^2 - ||y^k - y^{k+1}||^2 - ||y^{k+1} - \bar{y}||^2 &= 2\langle \frac{1}{2\lambda_k} \bigtriangledown f(y^{k+1}), y^{k+1} - \bar{y}\rangle \\ ||y^k - \bar{y}||^2 - ||y^k - y^{k+1}||^2 - ||y^{k+1} - \bar{y}||^2 &= \frac{1}{\lambda_k}\langle \bigtriangledown f(y^{k+1}), y^{k+1} - \bar{y}\rangle. \end{aligned} \quad (3.8)$$

Como f es convexa y continuamente diferenciable entonces del Teorema 2.2.26

$$\langle \bigtriangledown f(x), x - z\rangle \geq f(x) - f(z); \quad \forall x, z \in I\!R^n.$$

En particular para $x = y^{k+1}$ y $z = \bar{y}$

$$\langle \bigtriangledown f(y^{k+1}), y^{k+1} - \bar{y}\rangle \geq f(y^{k+1}) - f(\bar{y}).$$

Como $\lambda_k > 0$

$$\frac{1}{\lambda_k}\langle \bigtriangledown f(y^{k+1}), y^{k+1} - \bar{y}\rangle \geq \frac{1}{\lambda_k}(f(y^{k+1}) - f(\bar{y})). \quad (3.9)$$

Reemplazando (3.9) en (3.8)

$$\frac{1}{\lambda_k}[f(y^{k+1}) - f(\bar{y})] \leq ||y^k - \bar{y}||^2 - ||y^k - y^{k+1}||^2 - ||y^{k+1} - \bar{y}||^2. \quad (3.10)$$

Como $\bar{y}$ es un minimizador de f se cumple: $f(y^{k+1}) \geq f(\bar{y}); \quad \forall k \geq 0$

$$f(y^{k+1}) - f(\bar{y}) \geq 0; \quad \forall k \geq 0$$

como $\lambda_k > 0$

$$\frac{1}{\lambda_k}[f(y^{k+1}) - f(\bar{y})] \geq 0; \quad \forall k \geq 0. \tag{3.11}$$

Reemplazando (3.11) en (3.10)

$$0 \leq ||y^k - \bar{y}||^2 - ||y^k - y^{k+1}||^2 - ||y^{k+1} - \bar{y}||^2$$
$$||y^k - \bar{y}||^2 - ||y^{k+1} - \bar{y}||^2 \geq ||y^{k+1} - y^k||^2. \tag{3.12}$$

También se cumple

$$0 \leq ||y^k - \bar{y}||^2 - ||y^{k+1} - \bar{y}||^2$$
$$||y^{k+1} - \bar{y}||^2 \leq ||y^k - \bar{y}||^2; \quad \forall k > 0$$
$$||y^{k+1} - \bar{y}|| \leq ||y^k - \bar{y}||; \quad \forall k > 0.$$

Por lo tanto

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

□

Lema 3.2.3. *Dado una función $f : \mathbb{R}^n \to \mathbb{R}$ convexa y continuamente diferenciable. Si el conjunto U de minimizadores de la función f es no vacío. Se concluye que la sucesión $\{x^k\}$ generada por APP satisface la siguiente igualdad*

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

Demostración. Del Lema 3.2.2 se tiene

$$||x^{k+1} - u|| \leq ||x^k - u||; \quad \forall k > 0, \quad \forall u \in U.$$

En particular para $\bar{x} \in U$

$$||x^{k+1} - \bar{x}|| \leq ||x^k - \bar{x}||; \quad \forall k > 0.$$

Lo cual nos afirma que la sucesión $\{||x^k - \bar{x}||\}$ es real y decreciente, entonces

$$\{||x^k - \bar{x}||^2\} \text{ es convergente.}$$

Por otro lado se cumple

$$\lim_{k\to\infty}(||x^k - \bar{x}||^2 - ||x^{k+1} - \bar{x}||^2) = 0. \tag{3.13}$$

De la ecuación (3.12)

$$||x^{k+1} - x^k||^2 \leq ||x^k - \bar{x}||^2 - ||x^{k+1} - \bar{x}||^2; \quad \forall k \geq 0$$
$$0 \leq ||x^{k+1} - x^k||^2 \leq ||x^k - \bar{x}||^2 - ||x^{k+1} - \bar{x}||^2.$$

Tomando límite

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

De la ecuación (3.13)

$$\lim_{k\to\infty} ||x^{k+1} - x^k||^2 = 0$$

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

Por la Proposición (2.1.6)

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

□

Lema 3.2.4. *Dado una función $f : \mathbb{R}^n \to \mathbb{R}$ convexa y continuamente diferenciable. Si el conjunto U de minimizadores de la función f es no vacío. Se concluye que la sucesión $\{x^k\}$ generada por el algoritmo de punto proximal posee puntos de acumulación y todos pertenecen a U.*

Demostración. De la afirmación, Lema 3.2.2 la sucesión $\{x^k\}$ es convergente

$$||x^{k+1} - u||^2 \leq ||x^k - u||^2; \quad \forall k \geq 0, \forall u \in U.$$

Para $u_0 \in U$

$$||x^{k+1} - u_0|| \leq ||x^k - u_0||;$$

$$||x^{k+1} - u_0|| \leq ||x^0 - u_0||; \quad \forall k \geq 0$$

$$||x^{k+1}|| - ||u_0|| \leq ||x^{k+1} - u_0|| \leq ||x^0 - u_0||; \quad \forall k \geq 0$$

$$||x^{k+1}|| \leq ||x^0 - u_0|| + ||u_0||; \quad \forall k \geq 0.$$

Entonces

$$\{x^k\} \text{ es una sucesión limitada.}$$

Si tenemos del Teorema 2.1.10 que existe $\{x^{k_j}\} \subset \{x^k\}$ convergente a $x^{k_j} \to a$, entonces $''a''$ es un punto de acumulación. Por lo tanto

$$\{x^k\} \text{ tiene puntos de acumulación.}$$

Consideremos en este caso $\bar{x}$ un punto de acumulación de $\{x^k\}$
es decir, existe $\{x^{k_j}\} \subset \{x^k\}$ que verifica:

$$\lim_{k\to\infty} x^{k_j} = \bar{x}; \quad x^{k_j} \neq \bar{x} \quad \forall k_j \geq 0 \tag{3.14}$$

si se considera $x^{k_j} = \bar{x}$ para algún $j \in \mathbb{N}$, así como está definido el algoritmo el punto proximal clásico, $\bar{x}$ sería un mínimo de $f_\lambda(x^{k_j})$ y por lo tanto pertenece al conjunto U.
Del Lema 3.2.3

$$\lim_{k\to\infty} [x^{k+1} - x^k] = 0$$

$$\lim_{j\to\infty} [x^{k_j+1} - x^{k_j}] = 0. \tag{3.15}$$

De (3.14) y (3.15) se tiene

$$\lim_{j\to\infty} x^{k_j+1} = \bar{x}.$$

Como $\{\lambda_{k_j}\}$ es acotada y de (3.15)

$$\lim_{j\to\infty} \lambda_{k_j}(x^{k_j+1} - x^{k_j}) = 0$$

$$\lim_{j\to\infty} 2\lambda_{k_j}(x^{k_j+1} - x^{k_j}) = 0. \tag{3.16}$$

Como x^{k+1} es el mínimo de f_k, entonces por el Teorema 2.2.29

$$\bigtriangledown f_k(x^{k+1}) = \bigtriangledown f(x^{k+1}) + 2\lambda_k(x^{k+1} - x^k) = 0$$

$$\bigtriangledown f(x^{k+1}) = 2\lambda_k(x^k - x^{k+1}).$$

Como $\{k_j\} \subset \{k\}$, entonces se tiene

$$\bigtriangledown f(x^{k_j+1}) = 2\lambda_{k_j}(x^{k_j} - x^{k_j+1}).$$

Tomando límite a la ecuación anterior

$$\lim_{j\to\infty} \bigtriangledown f(x^{k_j+1}) = \lim_{j\to\infty} 2\lambda_{k_j}(x^{k_j} - x^{k_j+1}).$$

De la ecuación (3.16), se tiene

$$\lim_{j\to\infty} \bigtriangledown f(x^{k_j+1}) = 0.$$

Como la función f es C^1, entonces

$$\bigtriangledown f(\lim_{j\to\infty} x^{k_j+1}) = 0$$

como $x^{k_j+1} \to \bar{x}$ se tiene

$$\bigtriangledown f(\bar{x}) = 0. \tag{3.17}$$

Por hipótesis f es convexa y continuamente diferenciable en $\bar{x}$, de la ecuación (3.17) y del Teorema 2.2.28, se tiene

$$\bar{x} \in U.$$

□

Teorema 3.2.5. *Dado una función $f : \mathbb{R}^n \to \mathbb{R}$ convexa y continuamente diferenciable. Si $U \neq \emptyset$ es el conjunto de minimizadores. En tal caso la sucesión $\{x^k\}$ generada por APP converge a un punto $\bar{x} \in U$.*

Demostración. Del Lema (3.2.4)
La sucesión $\{x^k\}$ tiene una subsucesión convergente a un punto $\bar{x}$ qu pertenece a U, es decir

$$\text{existe } \{x^{k_j}\} \subset \{x^k\} / \quad \lim_{j\to\infty} x^{k_j} = \bar{x}. \tag{3.18}$$

Por el Lema 3.2.2

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

En particular para $\bar{x} \in U$.

$$||y^{k+1} - \bar{x}|| \leq ||y^k - \bar{x}||; \forall k \geq 0.$$

Lo que afirma que $\{||x^k - \bar{x}||\}$ es una sucesión decreciente y acotada para todo $k \in \mathbb{N}$. Entonces

la sucesión $\{||x^k - \bar{x}||\}$ es convergente

De la ecuación 3.18

$$\lim_{j\to\infty} x^{k_j} - \bar{x} = 0$$

$$\lim_{j\to\infty} ||x^{k_j} - \bar{x}|| = 0.$$

Se cumple

$$\lim_{k\to\infty} ||x^k - \bar{x}|| = 0$$

$$\lim_{k\to\infty} x^k - \bar{x} = 0$$

$$\lim_{\to\infty} x^k = \bar{x}.$$

□

Ejemplo 3.2.6. *Sea $f(x) = (x-2)^2$ una función convexa, limitada inferiormente y continuamente diferenciable*

$$(P)\begin{cases} minf(x) & = (x-2)^2 \\ s.a & x \in \mathbb{R} \end{cases}$$

(1) Dado $x_0 = 0$.

(2) a) para $k = 0$, $x_0 = 0$ y $\lambda_0 = 1$ donde $\lambda_k = \frac{1}{1+k}$

$$(P_0)\begin{cases} minf_0(x) & = (x-2)^2 + 1.|x-0|^2 \\ s.a & x \in \mathbb{R} \end{cases}$$

Como f es una función diferenciable

$$\bigtriangledown f_0(x) = 2(x-2) + 2x = 0$$

$$4x - 4 = 0$$

$$\Rightarrow x = 1.$$

Como existe el mínimo global de f_0.

$$x = 1 \text{ es solución de } P_0$$

b) para $k = 1$, $x_1 = 1$ y $\lambda_0 = \frac{1}{2}$ donde $\lambda_k = \frac{1}{1+k}$

$$(P_1)\begin{cases} minf_1(x) & = (x-2)^2 + \frac{1}{2}.|x-1|^2 \\ s.a & x \in \mathbb{R} \end{cases}$$

Como f_1 es una función diferenciable

$$\bigtriangledown f_1(x) = 2(x-2) + x - 1 = 0$$

$$3x - 5 = 0$$

$$\Rightarrow x = \frac{5}{3}$$

Como existe el mínimo global de f_1.

$$x = \frac{5}{3} \text{ es solución de } P_1$$

c) para $k = 2$, $x_2 = \frac{5}{3}$ *y* $\lambda_2 = \frac{1}{3}$ *donde* $\lambda_k = \frac{1}{1+k}$

$$(P_2) \begin{cases} min f_2(x) & = (x-2)^2 + \frac{1}{3}.|x - \frac{5}{3}|^2 \\ s.a & x \in \mathbb{R} \end{cases}$$

Como f_2 *es una función diferenciable*

$$\bigtriangledown f_2(x) = 2(x-2) + \tfrac{2}{3}(x - \tfrac{5}{3}) = 0$$

$$\tfrac{8}{3}x - \tfrac{46}{9} = 0$$

$$\Rightarrow x = \tfrac{23}{12}$$

Como existe el mínimo global de f_2.

$$x = \frac{23}{12} \text{ es solución de } P_2$$

d) para $k = 3$, $x_3 = \frac{23}{12}$ *y* $\lambda_3 = \frac{1}{4}$ *donde* $\lambda_k = \frac{1}{1+k}$

$$(P_3) \begin{cases} min f_3(x) & = (x-2)^2 + \frac{1}{4}.|x - \frac{23}{12}|^2 \\ s.a & x \in \mathbb{R} \end{cases}$$

Como f_3 *es una función diferenciable*

$$\bigtriangledown f_3(x) = 2(x-2) + \tfrac{2}{4}(x - \tfrac{23}{12}) = 0$$

$$\tfrac{5}{2}x - \tfrac{119}{24} = 0$$

$$\Rightarrow x = \tfrac{119}{60}$$

Como existe el mínimo global de f_3.

$$x = \frac{119}{60} \text{ es solución de } P_3$$

Después de la cuarta iteración el algoritmo continua iterando.

Ejemplo 3.2.7. *Sea una* $f(x, y) = 1 + (x-2)^2 + (y+1)^2$ *función convexa, limitada inferiormente y continuamente diferenciable.*

(1) Dado $x_0 = (0, 0)$

(2) a*) para* $k = 0,\ x_0 = (0, 0)$ *y* $\lambda_0 = 1$ *donde* $\lambda_k = \frac{1}{1+k}$.

$$(P_0) \begin{cases} min f_0(x, y) & = 1 + (x-2)^2 + (y+1)^2 + 1||(x, y)||^2 \\ s.a & (x, y) \in \mathbb{R}^2 \end{cases}$$

Como f_0 *es una función diferenciable*

$$\bigtriangledown f_0(x, y) = (2(x-2) + 2x, 2(y+1) + 2y) = (0, 0)$$

$$4x - 4 = 0, \quad 4y + 2 = 0$$

$$\Rightarrow x = 1 \text{ y } y = \tfrac{-1}{2}$$

como existe el mínimo global de f_0.

$$(1, -\tfrac{1}{2}) \text{ es solución de } P_0.$$

b) para $k = 1,\ x_1 = (1, -\frac{1}{2})$ *y* $\lambda_1 = 1/2$ *donde* $\lambda_k = \frac{1}{1+k}$.

$$(P_1) \begin{cases} min f_1(x,y) & = 1 + (x-2)^2 + (y+1)^2 + \frac{1}{2}||(x,y) - (1, -\frac{1}{2})||^2 \\ s.a & (x,y) \in \mathbb{R}^2 \end{cases}$$

Como f_1 *es una función diferenciable*

$$\bigtriangledown f_1(x,y) = \big(2(x-2) + (x-1), 2(y+1) + (y+\tfrac{1}{2})\big) = (0,0)$$

$$3x - 5 = 0, \quad 3y + \tfrac{5}{2} = 0$$

$$\Rightarrow x = \tfrac{5}{3} \text{ y } y = -\tfrac{5}{6}$$

como existe el mínimo global de f_1.

$$(\tfrac{5}{3}, -\tfrac{5}{6}) \text{ es solución de } P_1.$$

c) para $k = 2,\ x_2 = (\frac{5}{3}, -\frac{5}{6})$ *y* $\lambda_2 = 1/3$ *donde* $\lambda_k = \frac{1}{1+k}$.

$$(P_2) \begin{cases} min f_2(x,y) & = 1 + (x-2)^2 + (y+1)^2 + \frac{1}{3}||(x,y) - (\frac{5}{3}, -\frac{5}{6})||^2 \\ s.a & (x,y) \in \mathbb{R}^2 \end{cases}$$

Como f_2 *es una función diferenciable*

$$\bigtriangledown f_2(x,y) = \big(2(x-2) + \tfrac{2}{3}(x - \tfrac{5}{3}), 2(y+1) + \tfrac{2}{3}(y + \tfrac{5}{6})\big) = (0,0)$$

$$\tfrac{8x}{3} - \tfrac{46}{9} = 0, \quad \tfrac{8y}{3} + \tfrac{46}{18} = 0$$

$$\Rightarrow x = \tfrac{23}{12} \text{ y } y = -\tfrac{23}{24}$$

como existe el mínimo global de f_2.

$$(\tfrac{23}{12}, -\tfrac{23}{24}) \text{ es solución de } P_2.$$

d) para $k = 3,\ x_3 = (\frac{23}{12}, -\frac{23}{24})$ *y* $\lambda_3 = 1/4$ *donde* $\lambda_k = \frac{1}{1+k}$.

$$(P_3) \begin{cases} min f_2(x,y) & = 1 + (x-2)^2 + (y+1)^2 + \frac{1}{4}||(x,y) - (\frac{23}{12}, -\frac{23}{24})||^2 \\ s.a & (x,y) \in \mathbb{R}^2 \end{cases}$$

Como f_3 *es una función diferenciable*

$$\bigtriangledown f_3(x,y) = \big(2(x-2) + \tfrac{1}{2}(x - \tfrac{23}{12}), 2(y+1) + \tfrac{1}{2}(y + \tfrac{23}{24})\big) = (0,0)$$

$$\tfrac{5x}{2} - \tfrac{119}{24} = 0, \quad \tfrac{5y}{2} + \tfrac{119}{48} = 0$$

$$\Rightarrow x = \tfrac{119}{60} \text{ y } y = -\tfrac{119}{120}$$

como existe el mínimo global de f_3.

$$(\tfrac{119}{60}, -\tfrac{119}{120}) \text{ es solución de } P_3.$$

Después de la cuarta iteración el algoritmo continua iterando.

Capítulo 4

CUASI-DISTANCIA

La idea de distancia entre elementos de un conjunto fue un tema ampliamante abordado por los matemáticos cuya idea abstracta fue generalizado en los espacios métricos. La importancia de los espacios métricos se debe en parte a que estos ayudan a dar una idea de proximidad entre dos puntos. La norma determina la proximidad de sus elementos sobre la magnitud de un vector, de modo que hay una estrecha relación, pues, puede definirse una métrica teniendo como base un espacio normado.
Las distancias no simétricas (llamadas también asimétricas) fueron mencionadas en su estudio sobre conjuntos por Hausdorff (2021).
En este capítulo se va hacer un estudio sobre la cuasi-distancia un tipo de distancia no simétrica así como también algunas de sus propiedades.

4.1. Espacios con distancias no simétricas

Los espacios con distancias no simétricas son conjuntos donde la función distancia entre elementos cualesquiera tienen orientación, es decir, la distancia de A a B, puede ser distinta de la distancia de B a A. De manera similar, en un espacio normado con una norma no simétrica la magnitud de un vector v puede no tener la misma magnitud del vector $-v$. Estas distancias entre dos puntos en un espacio no simétrico serán de nuestro interés en este capítulo.

Definición 4.1.1. *Dado $X \neq \emptyset$. La aplicación $q : X \times X \to \mathbb{R}_+$ es llamado una cuasi-distancia o cuasi-simétrica si para todo $x, y, z \in X$*

(1) $q(x, y) = q(y, x) = 0 \Leftrightarrow x = y$.

(2) $q(x, y) \leq q(x, z) + q(z, y)$

Ejemplo 4.1.2. *Definamos la siguiente función $\rho_1 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$*

$$\rho_1(x, y) = \begin{cases} c_1(y - x) & x \leq y \\ c_2(x - y) & x > y \end{cases} \qquad \textit{con } c_1, c_2 > 0$$

Afirmamos que la función ρ_1 es una cuasi-distancia.
De la Definición 4.1.1 en (1)*,*

$$\rho_1(x,y) = \text{máx}\{c_1(y-x), c_2(x-y)\} = \rho_1(y,x) = \text{máx}\{c_1(x-y), c_2(y-x)\} = 0$$

Lo que implica que $x = y$

Por otro lado se cumple $\rho_1(x,x) = 0$

Por lo tanto se cumple lo afirmado por la Definición 4.1.1 en (1).

De la Definición 4.1.1 en (2)*; verifiquemos la condición*

$$\rho_1(x,y) \leq \rho_1(x,z) + \rho_1(z,y), \quad \forall x,y,z \in I\!R. \tag{4.1}$$

Considerando la definición equivalente del máximo de dos números reales, tenemos

$$\text{máx}\{a,b\} = \frac{a+b+|a-b|}{2} \tag{4.2}$$

En nuestro caso $a = c_1(y-x)$ *y* $b = c_2(x-y)$*,*

$$\text{máx}\{c_1(y-x), c_2(x-y)\} = \frac{c_1(y-x) + c_2(x-y) + |c_1(y-x) - c_2(x-y)|}{2} \tag{4.3}$$

haciendo algunas operaciones con $z \in I\!R$

$$c_1(y-x) = c_1(y-z) + c_1(z-x) \quad y \quad c_2(x-y) = c_2(x-z) + c_2(z-y)$$

Reemplazando en (4.3)

$$\begin{aligned} \text{máx}\{c_1(y-x), c_2(x-y)\} &= \frac{c_1(y-z) + c_1(z-x) + c_2(x-z) + c_2(z-y)}{2} \\ &+ \frac{|c_1(y-z) + c_1(z-x) - [c_2(x-z) + c_2(z-y)]|}{2} \end{aligned}$$

agrupando convenientemente

$$\begin{aligned} \text{máx}\{c_1(y-x), c_2(x-y)\} &= \frac{[c_1(y-z) + c_2(z-y)] + [c_1(z-x) + c_2(z-z)]}{2} \\ &+ \frac{|[c_1(y-z) - c_2(z-y)] + [c_1(z-x) - c_2(z-z)]|}{2} \end{aligned}$$

y utilizando la desigualdad triangular

$$\begin{aligned} \text{máx}\{c_1(y-x), c_2(x-y)\} &\leq \frac{c_1(y-z) + c_2(z-y) + |c_1(y-z) - c_2(z-y)|}{2} \\ &+ \frac{c_1(z-x) + c_2(z-z) + |c_1(z-x) - c_2(z-z)|}{2} \end{aligned}$$

Por la condición (4.2) se tiene

$$\text{máx}\{c_1(y-x), c_2(x-y)\} \leq \text{máx}\{c_1(y-z), c_2(z-y)\} + \text{máx}\{c_1(z-x), c_2(x-z)\}$$

esto es equivalente

$$\rho_1(x,y) \leq \rho_1(x,z) + \rho_1(z,y)$$

Por lo tanto se cumple la Definición 4.1.1, (2)*, y* ρ_1 *es una cuasi-distancia.*

Ejemplo 4.1.3. *Sea $X = \mathbb{R}^n$ y consideremos $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, x_n) \in X$ y $c_i^+, c_i^- > 0$ definimos*

$$C_i(x_i, y_i) = \begin{cases} c_i^+(y_i - x_i), & \textit{si } y_i - x_i > 0 \\ c_i^-(x_i - y_i), & \textit{si } y_i - x_i < 0 \\ 0 & \textit{si } x_i - y_i = 0 \end{cases}$$

donde cada $C_i(x_i, y_i)$ es una cuasi-distancia para $i = 1, \ldots, n$, entonces $C(x, y) = \sum_{i=1}^n C_i(x_i, y_i)$ es también una cuasi-distancia.
Esta cuasi-distancia también se puede expresar como

$$C(x, y) = \sum_{i=1}^{n} \text{máx}\{c_i^+(y_i - x_i), c_i^-(x_i - y_i)\}$$

(i) En efecto. De la Definición 4.1.1 en (1)

$$C(x, y) = C(y, x) = 0 \Rightarrow x = y.$$

Haremos la demostración por contradicción, supongamos que $x \neq y$, entonces existen $r_1, \ldots, r_k \in \{1, 2, \ldots, n\}$ con $k < n$ tal que $x_{r_j} \neq y_{r_j}, j = 1, \ldots, k$.
Por hipótesis
$C(x, y) = 0$ entonces tenemos

$$C(x, y) = \sum_{j=1}^{k} \text{máx}\{c_{r_j}^+(y_{r_j} - x_{r_j}), c_{r_j}^-(x_{r_j} - y_{r_j})\} = 0$$

Por ser $\text{máx}\{c_{r_j}^+(y_{r_j} - x_{r_j}), c_{r_j}^-(x_{r_j} - y_{r_j})\} > 0$, pues $x_{r_j} \neq y_{r_j}$ se tiene una contradicción pues la suma de números positivos no da cero.

Por lo tanto $x = y$

El reciproco de la Definición 4.1.1 en (1), tenemos $x = y$ es evidente que $C(x, x) = 0$.

(ii) De la Definición 4.1.1 en (2) vamos a verificar

$$C(x, y) \leq C(x, z) + C(z, y)$$

Como cada $C_i(x_i, y_i) \leq C_i(x_i, z_i) + C_i(z_i, y_i)$, $i = 1, 2, \ldots, n$ cumple (vea Ejemplo 4.1.2), entonces se verifica la desigualdad de arriba.

Por lo tanto $C(x, y) = \sum_{i=1}^n C_i(x_i, y_i)$ es una cuasi-distancia.

Ejemplo 4.1.4. *Sea $g : X \to \mathbb{R}_+$ una función inyectiva. Entonces para $x, y \in X$ definimos $q_g(x, y) = \text{máx}\{g(x) - g(y), 0\}$, así q_g genera una cuasi-distancia sobre X.*
En efecto. De la Definición 4.1.1 en (1).

Para el caso cuando $q_g(x, y) = q_g(y, x) = 0 \Rightarrow x = y$

vamos a demostrar por contradicción, supongamos que $x \neq y$ por ser g inyectiva se verifica uno de los casos $g(x) - g(y) > 0$ o $g(y) - g(x) > 0$ entonces se tendría en uno de los casos que $q_g(x, y) > 0$ lo cual es una contradicción por la hipótesis.
Para el caso en que $x = y$ se cumple $q_g(x, x) = 0$.
De la Definición 4.1.1 en (2), para el caso de la desigualdad triangular

$$q_g(x, y) = \text{máx}\{g(x) - g(y), 0\} = \frac{g(x) - g(y) + |g(x) - g(y)|}{2}$$

estamos considerando la ecuación 4.2, se cumple

$$q_g(x, y) = \text{máx}\{g(x) - g(y), 0\} = \frac{g(x) - g(z) + g(z) - g(y) + |g(x) - g(z) + g(z) - g(y)|}{2}$$

por desigualdad triangular se obtiene

$$q_g(x, y) \leq q_g(x, z) + q_g(z, y)$$

Por lo tanto q_g es una cuasi-distancia.

Ejemplo 4.1.5. *Sea $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ definido por*

$$q(x, y) = \begin{cases} 0 & \text{si } x \leq y \\ 1 & \text{en otro caso} \end{cases}$$

genera una cuasi-distancia sobre $\mathbb{R}$. (vea Gerla et al. (2006))

Observación 4.1.6. *Dado $\bar{z} \in \mathbb{R}^n$ fijo, tenemos*

$$C(x, \bar{z}) = \sum_{i=1}^{n} C_i(x_i, \bar{z}_i) = \sum_{i=1}^{n} max\{c_i^+(\bar{z}_i - x_i), c_i^-(x_i - \bar{z}_i)\}$$

Afirmamos que la función $C(., \bar{z})$ es convexa en $\mathbb{R}^n$.
Sea $x, y \in \mathbb{R}^n$ y $\lambda \in [0, 1]$, entonces $w = (1 - \lambda)x + \lambda y \in \mathbb{R}^n$, luego

$$C(w, \bar{z}) = \sum_{i=1}^{n} C_i(w_i, \bar{z}_i) = \sum_{i=1}^{n} max\{c_i^+(\bar{z}_i - w_i), c_i^-(w_i - \bar{z}_i)\}$$

$$C(w, \bar{z}) = \sum_{i=1}^{n} max\{c_i^+(\bar{z}_i - (1 - \lambda)x_i - \lambda y_i), c_i^-((1 - \lambda)x_i + \lambda y_i - \bar{z}_i)\}$$

$$C(w, \bar{z}) = \sum_{i=1}^{n} max\{c_i^+((1-\lambda)(\bar{z}_i - x_i) + \lambda(\bar{z}_i - y_i), c_i^-((1-\lambda)(x_i - \bar{z}_i) + \lambda(y_i - \bar{z}_i)\}$$

$$C(w, \bar{z}) \leq \sum_{i=1}^{n} (1-\lambda) max\{c_i^+(x_i - \bar{z}_i), c_i^-(\bar{z}_i - x_i)\} + \lambda max\{c_i^+(y_i - \bar{z}_i), c_i^-(\bar{z}_i - y_i)\}$$

$$C(w, \bar{z}) \leq (1 - \lambda)C(x, \bar{z}) + \lambda C(y, \bar{z})$$

Observación 4.1.7. *En general no es verdad que toda cuasi-distancia es convexa y coerciva ni en el primer argumento ni en el segundo. Veamos el Ejemplo 4.1.4: Aquí supongamos que $\bar{z} \in \mathbb{R}$ y $q(\bar{z}, .)$ es una función real, como g es una función inyectiva que en particular puede ser creciente o decreciente, ésta no necesariamente es convexa*
En el caso que g sea una función convexa, las funciones $q_g(., \bar{z})$ y $q_g(\bar{z}, .)$ son respectivamente convexa y concava.

En todo este trabajo, consideremos la siguiente condición sobre la cuasi- distancia q: hay constantes positivas α y β tal que

$$\alpha||x-y|| \leq q(x,y) \leq \beta||x-y|| \tag{4.4}$$

Los Ejemplo 4.1.2 y Ejemplo 4.1.3 satisfacen la condición (4.4), sin embargo el ejemplo 3.4 de Moreno et al. (2012) no la satisface.

Proposición 4.1.8. *Sea $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ una cuasi-distancia que verifica (4.4). Entonces para cada $\bar{z} \in \mathbb{R}^n$ las funciones $q(\bar{z},.)$ y $q(.,\bar{z})$ son Lipschitz continuas.*

Demostración. Vamos a probar que la función $q(\bar{z},.)$ es Lipschitz continua.
Por la desigualdad triangular, sean $x, y \in \mathbb{R}^n$

$$q(\bar{z},x) \leq q(\bar{z},y) + q(y,x) \tag{4.5}$$
$$q(\bar{z},y) \leq q(\bar{z},x) + q(x,y) \tag{4.6}$$

De (4.5) y (4.6) se obtiene

$$-q(x,y) \leq q(\bar{z},x) - q(\bar{z},y) \leq q(y,x).$$

En valor absoluto se tiene

$$|q(\bar{z},x) - q(\bar{z},y)| \leq \text{máx}\{q(x,y), q(y,x)\} \leq q(x,y) + q(y,x)$$

Por la condición (4.4) se tiene que existe $\beta > 0$ tal que

$$|q(\bar{z},x) - q(\bar{z},y)| \leq 2\beta||x-y||$$

siendo $k = 2\beta$. Por lo tanto $q(\bar{z},.)$ es Lipschitz continua.
La prueba es análoga para la función $q(.,\bar{z})$ □

Proposición 4.1.9. *Sea $\bar{z} \in \mathbb{R}^n$. Si q verifica (4.4) entonces $q^2(\bar{z},.)$ y $q^2(.,\bar{z})$ son funciones localmente Lipschitz continua sobre $\mathbb{R}^n$*

Demostración. Sea $\bar{x} \in \mathbb{R}^n$ y $\epsilon > 0$. Afirmamos que si $w \in B_\epsilon(\bar{x})$ entonces existe $k_{\bar{x}} > 0$ tal que $|q(\bar{z},w)| \leq k_{\bar{x}}$. En efecto

$$|q(\bar{z},w)| = |q(\bar{z},w) - q(\bar{z},\bar{z})| \leq k||w - \bar{x} + \bar{x} - \bar{z}||$$

La desigualdad anterior se da por la Proposición 4.1.8 y utilizando propiedad de la norma

$$|q(\bar{z},w)| \leq k(||w-\bar{x}|| + ||\bar{x}-\bar{z}||)$$

donde consideramos $k_{\bar{x}} = k(\epsilon + ||\bar{x}-\bar{z}||)$

Sean $x, y \in B_\epsilon(\bar{x})$ entonces

$$|q^2(\bar{z},x) - q^2(\bar{z},y)| = |q(\bar{z},x) + q(\bar{z},y)||q(\bar{z},x) - q(\bar{z},y)|$$

utilizando nuevamente la Proposición 4.1.8 se tiene

$$|q^2(\bar{z},x) - q^2(\bar{z},y)| \leq k|q(\bar{z},x) + q(\bar{z},y)|||x-y|| = 2k_{\bar{x}}.k||x-y||,$$

como $\bar{x}$ es un punto arbitrario, $q^2(\bar{z},.)$ es localmente Lipschitz sobre $\mathbb{R}^n$
La prueba es análoga para el otro caso $q^2(.,\bar{z})$. □

Proposición 4.1.10. *Sea $\bar{z} \in I\!R^n$. Si q verifica (4.4) entonces $q^2(\bar{z}, .)$ y $q^2(., \bar{z})$ son funciones coercivas.*

Demostración. Sea $\{x^m\}$ una sucesión critica, es decir

$$\lim_{m\to\infty} ||x^m|| = +\infty \vee \lim_{m\to\infty} x^m = \bar{x} \in \overline{I\!R^n} - I\!R^n$$

como $\overline{I\!R^n} - I\!R^n = \emptyset$ se verifica la primera condición.
Por (4.4) se tiene que existe $\alpha > 0$ tal que $\alpha||\bar{z} - x^m|| \leq q(\bar{z}, x^m)$ par todo $m \in I\!N$

$$\alpha(||x^m|| - ||\bar{z}||) \leq \alpha||\bar{z} - x^m|| \leq q(\bar{z}, x^m)$$

$$\alpha||x^m|| - \alpha||\bar{z}|| \leq q(\bar{z}, x^m)$$

tomando límite a ambos lados de la desigualdad

$$\lim_{m\to\infty}(\alpha||x^m|| - \alpha||\bar{z}||) \leq \lim_{m\to\infty} q(\bar{z}, x^m)$$

$$+\infty - \alpha||\bar{z}|| \leq \lim_{m\to\infty} q(\bar{z}, x^m)$$

$$\lim_{m\to\infty} q(\bar{z}, x^m) = +\infty$$

Por lo tanto

$$\lim_{m\to\infty} q^2(\bar{z}, x^m) = +\infty$$

De forma análoga se prueba para $q^2(., \bar{z})$ □

Lema 4.1.11. *Sea $\bar{z} \in I\!R^n$ y $\lambda > 0$. Si la función $f : I\!R^n \to I\!R \cup \{+\infty\}$ es limitada inferiormente y $q^2(\bar{z}, .)$ es una función coerciva, entonces la función $f + \lambda q^2(\bar{z}, .) : I\!R^n \to I\!R \cup \{+\infty\}$ es coerciva.*

Demostración. Sea $\{x^k\}$ una sucesión crítica tal que

$$\lim_{k\to\infty} ||x^k|| = +\infty \vee \lim_{k\to\infty} x^k = \bar{x} \in \overline{I\!R^n} - I\!R^n = \emptyset$$

Por hipótesis la función f es acotada inferiormente, es decir, existe $M > 0$ tal que

$$M \leq f(x);\ \forall x \in I\!R^n$$

Sea $\{x^k\} \subseteq I\!R^n$ entonces

$$M + \lambda q^2(\bar{z}, x^k) \leq f(x^k) + \lambda q^2(\bar{z}, x^k)$$

Tomando límite a la desigualdad anterior

$$\lim_{k\to\infty} \left[M + \lambda q^2(\bar{z}, x^k)\right] \leq \lim_{k\to\infty} \left[f(x^k) + \lambda q^2(\bar{z}, x^k)\right]$$

como $q^2(\bar{z}, .)$ es coerciva por la Proposición 4.1.10

$$M + \infty \leq \lim_{k\to\infty} \left[f(x^k) + \lambda q^2(\bar{z}, x^k)\right]$$

$$\lim_{k\to\infty} \left[f(x^k) + \lambda q^2(\bar{z}, x^k)\right] = +\infty.$$

Por lo tanto $f(.) + q^2(\bar{z}, .)$ es coerciva. □

Capítulo 5

ALGORITMO PROXIMAL INEXACTO

5.1. Antecedentes del problema

El algoritmo del punto proximal clásico (APP) fue propuesto por primera vez por Martinet (1970) y posteriormente fue estudiado por Rockafellar (1976) quien la uso para hallar ceros en operadores monótonos. Este algoritmo se propuso con la intención de resolver un problema de la forma.

$$\begin{array}{c} \min f(x) \\ s.a. \quad x \in I\!R^n \end{array} \tag{5.1}$$

donde la función $f : I\!R^n \to I\!R$ es convexa. El APP genera una sucesión $\{x^k\}$ de la siguiente forma: dado un punto inicial $x^0 \in I\!R^n$ se desea encontrar

$$x^{k+1} = argmin\{f(x) + \frac{\lambda_k}{2}||x - x^k||^2; x \in I\!R^n\} \tag{5.2}$$

este algoritmo tiene como elemento a $\lambda_k > 0$ parámetro de regularización, a la norma Euclideana $||.||$, y x^{k+1} es el mínimo global de $f(.) + \frac{\lambda_k}{2}||. - x^k||^2$ el cual es indicado por $argmin$. Si el parámetro de regularización cumple $\sum_{k=1}^{+\infty}(1/\lambda_k) = +\infty$, se prueba la convergencia de la sucesión $\{f(x^k)\}$ a un ínfimo de f; en el caso que el conjunto de soluciones óptimas de f es distinto del vacío, se tendría la convergencia de la sucesión $\{x^k\}$ a una solución óptima de (5.1). vea Güler (1991).

Si en el problema (5.1) la función f es no convexa, solucionar el problema alterno (5.2) de encontrar un punto mínimo global se tendría quizas un problema de dificil solución al problema original. En este aspecto se han hecho muchas investigaciones para la generalización del (APP). Una de estas extensiones es generalizando el término regularizador en (5.2) con norma Euclideana $||x - .||^2$ a una distancia generalizada como la cuasi-distancia $q^2(x, .)$ dando forma a las siguiente iteración

$$x^{k+1} = argmin\{f(x) + \frac{\lambda_k}{2}q^2(x, x^k); x \in I\!R^n\} \tag{5.3}$$

La iteración (5.3) genera una sucesión $\{x^k\}$ cuya convergencia fue estudiada primeramente por Moreno et al. (2012) y los autores probaron que si f satisface la condición

de Kurdyka Lojasiewics sobre el conjunto de puntos críticos entonces toda sucesión $\{x^k\}$ acotada, converge a un punto crítico.

La versión inexacta del (APP) fue realizada por Attouch et al. (2013) donde garantiza la convergencia de $\{x^k\}$ dado por (5.2) para una sucesión acotada, con un criterio de error relativo

H_2 : Condición de error relativo. Para cada $k \in I\!N$ existe $w^{k+1} \in \partial f(x^{k+1})$

$$||w^{k+1}|| \leq b||x^{k+1} - x^k|| \tag{5.4}$$

con $b > 0$ y ∂ la subdiferencial limitante.

En 2015, Bento and Soubeyran (2015b), motivado por los paper de Attouch et al. (2013) y Moreno et al. (2012) presentaron una versión inexacta del algoritmo de punto proximal usando cuasi-distancia dado por la siguiente iteración:
Tómese $k = 0, 1, 2, 3, \ldots$, sea $x^k \in I\!R^n$ hallar un x^{k+1} y $\varphi^{k+1} \in I\!R^n$ por consiguiente

$$f(x^{k+1}) \leq f(x^k) - \lambda_{k+1}(1-\sigma)\Gamma(q(x^k, x^{k+1})) \tag{5.5}$$
$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{5.6}$$
$$\psi^{k+1} \in \partial q(x^k, .)(x^{k+1}) \tag{5.7}$$
$$||\varphi^{k+1}|| \leq b\Gamma'(q(x^k, x^{k+1}))||\psi^{k+1}|| \tag{5.8}$$

donde $\Gamma : [0, +\infty) \to [0, +\infty)$ satisface algunas condiciones adecuadas definidas en (2) y (3) de Bento and Soubeyran (2015b), $\sigma \in (0, 1)$, ∂ denota la subdiferencial limitante de f. Asumiendo que la sucesión $\{x^k\}$ es acotada, f satisface la condición de la desigualdad de Kurdyka Lojasiewicz. los autores prueban que la sucesión converge a un punto crítico de f, estudian la tasa de convergencia de la sucesión y introducen una aplicación para el cambio dinámico en conducta de la ciencia.

En este trabajo introducimos una iteración inexacta del método del APP usando cuasi-distancias que considera errores vectoriales; tómese $k = 0, 1, 2, \ldots$, sea $x^k \in I\!R^n$, hallar $x^{k+1} \in I\!R^n$ y $\varphi^{k+1} \in I\!R^n$ con la propiedad

$$f(x^{k+1}) \leq f(x^k) - \tfrac{\lambda_{k+1}}{2} q(x^k, x^{k+1}) \tag{5.9}$$
$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{5.10}$$
$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1}) \tag{5.11}$$

donde

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}, \tag{5.12}$$

$\sigma > 0$ y

$$\psi^{k+1} \in \partial q(x^k, .)(x^{k+1}) \tag{5.13}$$

Para diferenciar nuestro trabajo del artículo de Bento and Soubeyran (2015b), realizaremos una comparación de los dos algoritmos. Observamos inicialmente que las condiciones (5.11) y (5.8) son diferentes, sin embargo, una condición suficiente y bien difícil de probar en cada iteración para que exista una relación estrecha entre ambas es que se cumpla que $\lambda_k||\psi^{k+1}|| \geq 1$, $\sigma < 1$ and $\{\lambda_k\}$ sea acotado inferiormente. En efecto,

$$||e^{k+1}||^2 = ||\varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}||^2 \leq \sigma^2 q(x^k, x^{k+1})^2$$

$$\leq \sigma^2(||\varphi^{k+1}||^2 + q(x^k, x^{k+1})^2)$$
$$\leq \sigma^2(||\varphi^{k+1}||^2 + ||\lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}||^2)$$

Como $\sigma < 1$, entonces aplicando el Lemma 4.1 de Attouch et al. (2013)

$$||\varphi^{k+1}|| \leq M\lambda_{k+1} q(x^k, x^{k+1})||\psi^{k+1}||$$

Como $\{\lambda_k\}$ está acotado, obtenemos (5.8). Así, podemos concluir que los algoritmos inexactos propuestos por Bento and Soubeyran (2015b) y el algoritmo propuesto en este artículo son, en general, diferentes.

5.2. El algoritmo

En esta sección se estudia un problema de optimización:

$$min\{f(x);\, x \in I\!R^n\} \tag{5.14}$$

Para lo cual $f : I\!R^n \to I\!R \cup \{+\infty\}$ es una función de valores extendidos y satisface las siguientes condiciones

$(\mathbf{H_1})$ La función f es propia y sci (semicontinua inferior).

$(\mathbf{H_2})$ f es coerciva, que es, $\lim_{||x||\to+\infty} f(x) = +\infty$

$(\mathbf{H_3})$ $q : I\!R^n \times I\!R^n \to I\!R_+$ es una cuasi-distancia satisfaciendo: existe un $\alpha > 0$ y $\beta > 0$ tal que

$$\alpha||x - y|| \leq q(x, y) \leq \beta||x - y||. \tag{5.15}$$

Aquí vamos a estudiar el algoritmos de punto proximal inexacto.

5.3. Algoritmo Proximal Inexacto (API)

Sea $q(.,.)$ una cuasi-distancia definida sobre $I\!R^n$ satisfaciendo la hipótesis (H_3) y f una función objetivo del problema (5.14) satisfaciendo las hipótesis (H_1) y (H_2).

Inicialización: Elegimos una sucesión acotada de parámetros positivos $\{\lambda_k\}$, un valor real $\sigma > 0$ y un punto cualquiera inicial

$$x^0 \in I\!R^n, \tag{5.16}$$

Paso Principal: Tómese $k = 0, 1, 2, ...$, sea $x^k \in I\!R^n$, hallar $(x^{k+1}, \varphi^{k+1}) \in I\!R^n \times I\!R^n$ con la propiedad

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) \tag{5.17}$$

$$\varphi^{k+1} \in \partial f(x^{k+1}) \tag{5.18}$$

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1}) \tag{5.19}$$

donde

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1} \tag{5.20}$$

y $\psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$

Criterio de Parada: Si $x^k = x^{k+1}$ verifica la condición o $0 \in \partial f(x^{k+1})$, entonces finalizar. Por otro lado, si no se verifica reemplazar $k \leftarrow k+1$ y volver al Paso Principal.

5.3.1. Existencia de los iterados

Teorema 5.3.1. *Bajo la suposición* $(\mathbf{H_1})$-$(\mathbf{H_3})$, *para cada* k, *existe* (x^{k+1}, φ^{k+1}) *satisfaciendo (5.17)- (5.20), que es el algoritmo proximal inexacto (API), entonces está bien definido.*

Demostración. Dado $x^k \in I\!R^n$ y considere el problema

$$\text{mín}\{f(x) + \frac{\lambda_{k+1}}{2} q^2(x^k, x) : x \in I\!R^n\}$$

De la semicontinuidad inferior de f y $q^2(x^k, .)$ y la coercitividad de f tenemos que existe x^{k+1} tal que

$$f(x^{k+1}) + \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1}) \leq f(x) + \frac{\lambda_{k+1}}{2} q^2(x^k, x),$$

para todo $x \in I\!R^n$. Tomando $x = x^k$ tenemos que (5.17).
Por otro lado, de la condición de optimalidad, véase la Proposición 2.4.16, tenemos que

$$0 \in \hat{\partial} \left(f(.) + \frac{\lambda_{k+1}}{2} q^2(x^k, .) \right) (x^{k+1}).$$

Por la Proposición (2.4.18) se obtiene (5.18) - (5.20) con $e^{k+1} = 0$ □

5.3.2. Propiedades de convergencia para el caso general

Proposición 5.3.2. *Bajo la suposición* $(\mathbf{H_1})$, $(\mathbf{H_2})$, $(\mathbf{H_3})$ *y* $0 < \bar{\lambda} < \lambda_{k+1} < \tilde{\lambda}$ *tenemos que la sucesión* $\{x^k\}$ *generada por API satisface:*

i $\{f(x^k)\}$ *es una sucesión convergente.*

ii $q(x^k, x^{k+1})$ *converge a cero.*

iii $\{x^k\}$ *es una sucesión acotada.*

iv Si x^{k_j} *converge a* $\bar{x}$, *entonces* $\{x^{k_j+1}\}$ *converge a* $\bar{x}$.

Demostración. .

(i) De (5.17) tenemos

$$f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1})$$

Dividiendo por λ_{k+1} y asociando tenemos

$$(1/2)\, q^2(x^k, x^{k+1}) \leq \frac{1}{\lambda_{k+1}} f(x^k) - \frac{1}{\lambda_{k+1}} f(x^{k+1}) \tag{5.21}$$

Por lo tanto, de (5.21) tenemos:

$$f(x^{k+1}) < f(x^k) \tag{5.22}$$

Del Corolario 2.2.16 obtenemos que f está acotado inferiormente, con lo cual la sucesión $\{f(x^k)\}$ decreciente es convergente.
Del Lema 2.1.15 obtenemos que $\{f(x^k)\}$ converge.

(ii) De la acotación de $\{\lambda_k\}$, la convergencia de $\{f(x^k)\}$ y de (5.21) obtenemos

$$\lim_{k\to\infty} q(x^k, x^{k+1}) = 0$$

(iii) De (5.22) obtenemos:

$$f(x^k) < f(x^0)$$

para todo $k = 1, ...$ Así $x^k \in L_f(a_0)$ donde $a_0 = f(x^0)$. Como $L_f(a_0)$ es acotada pues f es una función coerciva, entonces se concluye que la sucesión $\{x^k\}$ es limitada.

(iv) Utilizando la propiedad de la desigualdad triangular

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

por la condición (5.15) existe $\beta > 0$

$$||x^{k_j+1} - \bar{x}|| \leq \beta q(x^{k_j}, x^{k_j+1}) + ||x^{k_j} - \bar{x}||$$

Como $q(x^{k_j}, x^{k_j+1})$ converge a cero por item(ii), se tiene $\{x^{k_j+1}\}$ converge a $\bar{x}$.

□

Ahora, se define el conjunto

$$U := \{x \in I\!R^n; f(x) \leq \lim_{k\to\infty} f(x^k)\}.$$

Teniendo en cuenta la suposición $(\mathbf{H_1})$ - $(\mathbf{H_3})$ el conjunto U es no vacío, cerrado y acotado.

Teorema 5.3.3. *Bajo la suposición* $(\mathbf{H_1})$ - $(\mathbf{H_3})$ *y* $0 < \bar{\lambda} < \lambda_{k+1} < \tilde{\lambda}$*, cada subsucesión convergente de la sucesión* $\{x^k\}$ *generada por el (API) converge a un punto que pertenece a U.*

Demostración. Dado $\bar{x} \in I\!R^n$ y la sucesión $\{x^k\}$ tal que existe una subsucesión $x^{k_j} \to \bar{x}$. Por la Proposición 5.3.2 (i) la sucesión $\{f(x^k)\}$ converge y por el Teorema 2.1.9 se tiene que $f(x^{k_j})$ converge. Entonces

$$f(\bar{x}) \leq \lim_{j\to\infty} f(x^{k_j}) = \lim_{k\to\infty} f(x^k)$$

que es, $\bar{x} \in U$. □

Considere la siguiente suposición:
$(\mathbf{H_1})'$ La función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es semicontinua inferior, propia y continua sobre su dominio efectivo dom (f).

Teorema 5.3.4. *Bajo la suposición* $(\mathbf{H_1})'$, $(\mathbf{H_2})$, $(\mathbf{H_3})$, *y que* $0 < \bar{\lambda} < \lambda_{k+1} < \tilde{\lambda}$, *cada subsucesión convergente de la sucesión* $\{x^k\}$ *generada por (API), será punto crítico limitante de* f.

Demostración. Dado $\bar{x} \in \mathbb{R}^n$ y la sucesión $\{x^k\}$. Se tiene que existe $\{x^{k_j}\}$ tal que $x^{k_j} \to \bar{x}$ y por la Proposición 5.22 (iv), $x^{k_j+1} \to \bar{x}$. De (5.20) y la desigualdad triangular tenemos

$$||\varphi^{k_j+1}|| \leq ||e^{k_j+1}|| + \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})||\psi^{k+1}||$$

por la condición (5.19) se obtiene

$$||\varphi^{k_j+1}|| \leq \left(\sigma + \lambda_{k_j+1}||\psi^{k_j+1}||\right) q(x^{k_j}, x^{k_j+1}) \tag{5.23}$$

Como $q(., x^k)$ es localmente Lipschitz en $\mathbb{R}^n$ entonces $\partial q(., x^k)$ es localmente acotada (vea la Proposición 2.4.19) y del Lema 5.1 de Moreno et al. (2012) existe un $M > 0$ tal que $||\psi^{k_j+1}|| \leq M$ para todo $j \in \mathbb{N}$.
Tomando límite a ambos lados de la desigualdad (5.23) para $j \to \infty$ tenemos

$$\lim_{j\to\infty} ||\varphi^{k_j+1}|| \leq \lim_{j\to\infty} \left(\sigma + \lambda_{k_j+1} M\right) q(x^{k_j}, x^{k_j+1})$$

por lo tanto

$$\lim_{j\to\infty} \varphi^{k_j+1} = 0$$

Así, existe $x^{k_j+1} \to \bar{x}$ con $f(x^{k_j+1}) \to f(\bar{x})$ pues f es continua en $dom(f)$ y $\varphi^{k_j+1} \to 0$ con $\varphi^{k_j+1} \in \partial f(x^{k_j+1})$ entonces de la Definición 2.4.14, $0 \in \partial f(\bar{x})$. □

5.3.3. Propiedad de convergencia para funciones localmente Lipschitz

Asumir que la función f es localmente Lipschitz sobre $\mathbb{R}^n$, entonces el $dom(f)$ es abierto, entonces podemos considerar la subdiferencial de Clarke dado en la Definición 2.4.10

$$\partial f^\circ(x) = \{\xi \in \mathbb{R}^n \mid f^\circ(x, v) \geq \xi^T.v \text{ para todo } v \in \mathbb{R}^n\}$$

donde

$$f^\circ(x, v) = \lim_{\substack{y\to x \\ t\to 0^+}} \sup \frac{f(y + tv) - f(y)}{t}$$

Entonces, podemos sustituir en el (API) la subdiferencial limitante ∂ por ∂°.
Asumamos la siguiente suposición:
$(\mathbf{H_1})''$ La función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es semicontinua inferior, propia y localmente Lipschitz sobre $\mathbb{R}^n$.

Teorema 5.3.5. *Bajo la suposición* $(\mathbf{H_1})''$, $(\mathbf{H_2})$, $(\mathbf{H_3})$, *y que* $0 < \bar{\lambda} < \lambda_{k+1} < \tilde{\lambda}$ *cada punto de acumulación de la sucesión* $\{x^k\}$ *generado por el (API) (cambiando la subdiferencial* ∂ *por* ∂°*) es un punto crítico (en el sentido de la subdiferencial de Clarke) de el problema (5.14).*

Demostración. Dado $\bar{x} \in I\!R^n$ y la sucesión $\{x^k\}$, se tiene que existe x^{k_j} tal que $x^{k_j} \to \bar{x}$ y de la Proposición 5.22 (iv), $x^{k_j+1} \to \bar{x}$.
De (5.20) tenemos

$$\varphi^{k+1} = e^{k+1} - \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1} \in \partial^\circ f(x^{k+1})$$

así

$$f^\circ(x^{k_j+1}, v) \geq \langle e^{k_j+1} - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})\psi^{k_j+1}, v\rangle \text{ para todo } v \in I\!R^n$$

$$f^\circ(x^{k_j+1}, v) \geq \langle e^{k_j+1}, v\rangle - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})\langle \psi^{k_j+1}, v\rangle \text{ para todo } v \in I\!R^n$$

utilizando la propiedad de Cauchy se tiene

$$f^\circ(x^{k_j+1}, v) \geq -||e^{k_j+1}||\,||v|| - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})||\psi^{k_j+1}||\,||v||$$

por la condición (5.19)

$$f^\circ(x^{k_j+1}, v) \geq -q(x^{k_j}, x^{k_j+1})||v||(\sigma + \bar{\lambda}||\psi^{k_j+1}||)$$

Como $q(., x^k)$ es localmente Lipschitz en $I\!R^n$ entonces $\partial q(., x^k)$ es localmente acotada (vea la Proposición 2.4.19) y del Lema 5.1 de Moreno et al. (2012) existe un $M > 0$ tal que $||\psi^{k_j+1}|| \leq M$ para todo $j \in I\!N$. Entonces

$$f^\circ(x^{k_j+1}, v) \geq -q(x^{k_j}, x^{k_j+1})||v||(\sigma + \bar{\lambda}M) \tag{5.24}$$

Consideramos $v^{k_j+1} = v$ para $j \in I\!N$. Tomando $\limsup$ a ambos de la desigualdad (5.24)

$$\limsup_{j\to\infty} f^\circ(x^{k_j+1}, v^{k_j+1}) \geq \limsup_{j\to\infty} \left(-q(x^{k_j}, x^{k_j+1})||v^{k_j+1}||(\sigma + \bar{\lambda}M)\right)$$

Por el Teorema 2.4.9, (ii) en función de (x, v) la función $f^\circ(x, v)$ es semicontinua superior y como $q(x^{k_j}, x^{k_j+1}) \to 0$; tenemos

$$f^\circ(\bar{x}, v) \geq \limsup_{j\to\infty} f^\circ(x^{k_j+1}, v^{k_j+1}) \geq 0, \;\; \forall v \in I\!R^n$$

se sigue que $0 \in \partial^\circ f(\bar{x})$. □

Proposición 5.3.6. *Sea* $\{x^k\}$ *la sucesión generada por el (API) y denotamos por* $\omega(x_0)$ *el conjunto de todos sus puntos de acumulación. Si* f *satisface* $(\mathbf{H_1})'$ *entonces* f *es finito y constante sobre* $\omega(x_0)$.

Demostración. Dado $\bar{x} \in I\!R^n$ y la sucesión $\{x^k\}_{k\in I\!N}$, tal que existe $x^{k_j} \to \bar{x}$, como $\{f(x^k)\}$ es una sucesión decreciente y acotada inferiormente se tiene $\lim_{k\to\infty} f(x^k) = \beta$, por la continuidad de f se cumple que $\lim_{j\to\infty} f(x^{k_j}) = f(\bar{x}) = \beta$. Como $\bar{x}$ es un punto arbitrario se cumple lo afirmado. □

Observación 5.3.7. *1) Se tiene que $\omega(x_0)$ es un subconjunto de U, por lo cual se puede afirmar que es acotado y como f es constante en $\omega(x_0)$ se tiene que ella es cerrada.*

2) De la condición $(\mathbf{H_1})''$ toda función localmente Lipschitz es continua, entonces ella verifica la condición de la Proposición 5.3.6.

5.3.4. Propiedad de convergencia para funciones convexas

Del Teorema 5.3.4 concluimos que si la función f es convexa y continua sobre $dom(f)$, entonces cada punto crítico de f es un mínimo del problema (5.14). Similarmente del Teorema 5.3.5 si el $dom(f)$ es abierto, entonces cada punto crítico de f es un mínimo del problema (5.14).

Desafortunadamente existen muchas funciones convexas que no son continuas sobre su dominio ni su $dom(f)$ es abierto. En esta sección completamos la convergencia del método para funciones convexas arbitrarias. Para esto asumiremos la siguiente condición:
$(\mathbf{H_1})'''$ f es una función semicontinua inferior, propia y convexa.

Teorema 5.3.8. *Bajo la suposición $(\mathbf{H_1})'''$, $(\mathbf{H_2})$, $(\mathbf{H_3})$, y que λ_k es acotada, cada punto de acumulación de $\{x^k\}$ es una solución optima del problema (5.14).*

Demostración. Dado $x \in I\!R^n$. Por hipótesis existe $\{x^{k_j}\}$ tal que $x^{k_j} \to \bar{x}$. Como f es semicontinua inferior y $f(x^{k_j})$ converge entonces

$$f(\bar{x}) \leq \lim_{j\to\infty} f(x^{k_j}) \tag{5.25}$$

Por otro lado como f es convexa, entonces para cada $x \in I\!R^n$ tenemos

$$f(x) \geq f(x^{k_j+1}) + \langle s^{k_j+1}, x - x^{k_j+1}\rangle \tag{5.26}$$

donde $s^{k_j+1} \in \partial f(x^{k_j+1})$
De (5.20) y $q^2(x^k, .)$ es localmente Lipschitz

$$e^{k_j+1} - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})\psi^{k_j+1} \in \partial f(x^{k_j+1}) \tag{5.27}$$

donde $\psi^{k_j+1} \in \partial q(x^{k_j}, .)(x^{k_j+1})$.
De (5.26) y (5.27) tenemos

$$f(x) \geq f(x^{k_j+1}) + \langle e^{k_j+1} - \lambda_{k_j+1} q(x^{k_j}, x^{k_j+1})\psi^{k_j+1}, x - x^{k_j+1}\rangle$$

Por la propiedad de Cauchy en $I\!R^n$, la condición (5.19) y $\lambda_{k_j+1} < \tilde{\lambda}$

$$f(x) \geq f(x^{k_j+1}) - \sigma q(x^{k_j}, x^{k_j+1})||x - x^{k_j}|| - \tilde{\lambda} q(x^{k_j}, x^{k_j+1})||\psi^{k_j+1}||\,||x - x^{k_j+1}||$$

Como $q(., x^k)$ es localmente Lipschitz en $I\!R^n$ entonces $\partial q(., x^k)$ es localmente acotada (vea la Proposición 2.4.19) y del lema 5.1 de Moreno et al. (2012) existe un $M > 0$ tal que $||\psi^{k_j+1}|| \leq M$ para todo $j \in I\!N$. Entonces

$$f(x) \geq f(x^{k_j+1}) - q(x^{k_j}, x^{k_j+1})||x - x^{k_j}|| \left(\sigma - \tilde{\lambda} M\right) \tag{5.28}$$

Tomando límite a ambos lados de la desigualdad (5.28) cuando $j \to \infty$

$$\lim_{j\to\infty} f(x) \geq \lim_{j\to\infty} f(x^{k_j+1}) - \lim_{j\to\infty} \left(q(x^{k_j}, x^{k_j+1}) ||x - x^{k_j}|| \left(\sigma - \tilde{\lambda} M \right) \right)$$

De la Proposición 5.22 (iv), $x^{k_j+1} \to \bar{x}$ y como $q(x^{k_j}, x^{k_j+1}) \to 0$ tenemos

$$\lim_{j\to\infty} f(x) \geq \lim_{j\to\infty} f(x^{k_j+1}) \quad \forall x \in I\!R^n$$

□

Como $f(x^{k_j+1})$ es una subsucesión de $f(x^k)$ ambos convergen a un mismo límite y de la condición (5.25) tenemos

$$f(x) \geq \lim_{j\to\infty} f(x^{k_j+1}) = \lim_{j\to\infty} f(x^{k_j}) \geq f(\bar{x}) \quad \forall x \in I\!R^n$$

Por lo tanto $f(\bar{x})$ es un mínimo global.

5.3.5. Convergencia para funciones que satisface la condición de Kurdyka Lojasiewicz

En esta subsección probaremos la convergencia de las sucesiones para funciones que satisfacen la propiedad de Kurdyka Lojasiewicz. Nuestros resultados son motivados del paper de Moreno et al. (2012).

Recordaremos una caracterización, sobre el contexto de la subdiferencial limitante, de la desigualdad de Kurdyka Lojasiewicz. Dado $\bar{r} \in (0, +\infty]$, tenemos

$$K(0, \bar{r}) := \{\phi \in C[0, \bar{r}) \cap C^1(0, \bar{r}) : \phi(0) = 0 \text{ y } \phi'(r) > 0, \forall r \in (0, \bar{r})\},$$

donde $C[0, \bar{r})$ (respectivamente $C^1(0, \bar{r})$) denota el conjunto de funciones continuas sobre $[0, \bar{r})$ (respectivamente, funciones de clase C^1 sobre $(0, \bar{r})$). Sea $f : I\!R^n \to I\!R \cup \{+\infty\}$ una función semicontinua inferior. Para $\eta_1 < \eta_2 \leq +\infty$ se tiene el siguiente conjunto

$$[\eta_1 < f < \eta_2] = \{x \in I\!R^n; \eta_1 < f(x) < \eta_2\}.$$

Definición 5.3.9. *La función f se dice que tiene la propiedad de Kurdyka Lojasiewicz en $\bar{x} \in dom(\partial f)$ si existe un $\eta \in (0, +\infty]$, una vecindad U de $\bar{x}$ y una función concava $\varphi \in K(0, \eta)$, tal que para todo $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$ la desigualdad de Kurdyka Lojasiewicz cumple*

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1 \tag{5.29}$$

Dado $x \in I\!R^n$ y $\epsilon > 0$ fijo, denotamos por $B_q(x, \epsilon)$ la bola abierta con respecto a la cuasi-distancia q de centro x y radio $\epsilon > 0$, es definido como el conjunto de puntos $B_q(x, \epsilon) = \{y \in I\!R^n; q(x, y) < \epsilon\}$. En caso especial si q es la distancia Euclideana, $B_q(x, \epsilon)$ es denotada como $B(x, \epsilon)$.

Teorema 5.3.10. *Sea $\{x^k\}$ la sucesión generada por el (API), f es una función que satisface $(\mathbf{H_1})'$ y $(\mathbf{H_2})$, q es una cuasi-distancia que satisface $(\mathbf{H_3})$, $\bar{x}$ es un punto de acumulación de la sucesión $\{x^k\}$ donde f satisface la desigualdad de Kurdyka-Lojasiewicz. Sea $U \subset \mathbb{R}^n$ una vecindad de $\bar{x}$, $\eta > 0$ y $\phi : [0, \eta) \to \mathbb{R}^+$ una función cóncava que satisfacen las condiciones de la definición de la condición de Kurdyka-Lojasiewicz. Dados $\delta > 0$ y $r \in (0, 1)$ tal que $B(\bar{x}, \frac{\delta}{\alpha}) \subset U$ y $\overline{M} = 2\left(\frac{\sigma + M\lambda_+}{r(1-r)\lambda_-}\right)$ entonces existe un $N_0 \in \mathbb{N}$ satisfaciendo*

$$f(\bar{x}) < f(x^k) < f(\bar{x}) + \eta, \quad \textit{para todo} \quad k \geq N_0 \tag{5.30}$$

$$q(\bar{x}, x^{N_0}) + \left(\frac{r}{1-r}\right)\sqrt{\frac{2}{\lambda_-}\left[f(x^{N_0-1}) - f(\bar{x})\right]} + \overline{M}\left[\phi(f(x^{N_0})) - \phi(f(\bar{x}))\right] + \sqrt{\frac{2}{\lambda_-}\left[f(x^{N_0}) - f(\bar{x})\right]} < \delta \tag{5.31}$$

y

$$x^{N_0+j} \in B_q(\bar{x}, \delta) \textit{ con } j = 1, 2, \ldots \tag{5.32}$$

Además, la sucesión $\{x^k\}$ converge al punto $\bar{x}$ y en particular

$$\sum_{j=1}^{+\infty} q(x^{N_0+j}, x^{N_0+j+1}) < +\infty \tag{5.33}$$

Demostración. Sin perdida de generalidad suponemos que $f(x) := f(x) - \text{ínf}_{k\geq 0} f(x^k)$. Así, de la Proposición 5.3.2, (i), obtenemos que $\lim_{k\to\infty} f(x^k) = 0$. De la propiedad de Kurdyka lojasiewicz existe $\eta \in (0, +\infty]$, una función estrictamente concava $\phi \in K(0, \eta)$. Entonces usando la caracterización de la diferenciabilidad de la función concava $\phi(s)$ para $s, t \in (0, \eta)$ tenemos

$$\phi(t) - \phi(s) \geq \phi'(t)(t - s) \tag{5.34}$$

De la convergencia de $\{f(x^k)\}$ a cero, existe un $k_0 \in \mathbb{N}$ tal que $0 < f(x^k) < \eta$, para todo $k \geq k_0$. definimos la sucesión $\{b^k\}$ dado por

$$b^k = q(\bar{x}, x^k) + \left(\frac{1}{1-r}\right)\sqrt{\frac{2}{\lambda_-} f(x^k)} + \sqrt{\frac{2}{\lambda_-} f(x^{k-1})} + \overline{M}[\phi(f(x^{k+1}))]$$

La sucesión $\{b^k\}$ definida así tiene un punto de acumulación en cero, esto se debe a la continuidad de la función f, $q(\bar{x}, .)$ y ϕ en $\bar{x}$ (por condición $\lim_{k\to\infty} \phi(f(x^k)) = \phi(0) = 0$), entonces existe un $j_0 \in \mathbb{N}$ tal que para todo $j > j_0$ se tiene $b^{k_j} < \delta$. Definiendo $N_0 = \max\{k_0, k_{j_0}\}$, se cumplen las condiciones (5.31) y (5.30).
Observe que (5.31) implica que $x^{N_0} \in B_q(\bar{x}, \delta)$ y por $(\mathbf{H_3})$ esto significa que $x^{N_0} \in B(\bar{x}, \delta/\alpha) \subset U$. Por lo tanto de (5.30) tenemos que

$$x^{N_0} \in U \cap [0 < f < \eta]$$

Tomando $s = f(x^{N_0+1})$ y $t = f(x^{N_0})$ en (5.34) tenemos

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \phi'(f(x^{N_0}))[f(x^{N_0}) - f(x^{N_0+1})] \tag{5.35}$$

De la condición (5.17), reemplazados en (5.35) obtenemos para todo $k \in I\!N$

$$\phi(f(x^k)) - \phi(f(x^{k+1})) \geq \phi'(f(x^k))[\frac{1}{2}\lambda_{k+1} q^2(x^k, x^{k+1})]. \tag{5.36}$$

Por otro lado, tomando una vecindad alrededor de $\bar{x}$ se tiene por la condición (5.29): $\phi'(f(x))dist(0, \partial f(x)) \geq 1$ para todo $x \in B(\bar{x}, \epsilon) \cap [0 < f < \eta]$ con $\epsilon > 0$, de esto se sigue que $0 \notin \partial f(x^{N_0})$. En particular para $x^{N_0} \in U \cap [0 < f < \eta]$ se tiene

$$\phi'(f(x^{N_0}))\, ||\varphi^{N_0}|| \geq 1, \text{ con } \varphi^{N_0} \in \partial f(x^{N_0})$$

donde

$$\phi'(f(x^{N_0})) \geq \frac{1}{||\varphi^{N_0}||} \tag{5.37}$$

Ahora, de (5.18) y (5.20) tenemos que $\varphi^k \in \partial f(x^k)$ y

$$||\varphi^k|| \leq ||e^k|| + \lambda_k q(x^{k-1}, x^k)||\psi^k|| \tag{5.38}$$

donde $\psi^k \in \partial q(x^{k-1}, .)(x^k)$.
De (5.19) en (5.38) obtenemos

$$||\varphi^k|| \leq \sigma q(x^{k-1}, x^k) + \lambda_k q(x^{k-1}, x^k)||\psi^k|| \quad \text{para } k = 1, 2, \ldots \tag{5.39}$$

Como $\lambda_- < \lambda_k < \lambda_+$ y del Lema 5.1 de Moreno que afirma $||\psi^k|| \leq M, k = 1, 2, \ldots$, se tiene en (5.39)

$$||\varphi^k|| \leq (\sigma + \lambda_+ M)\, q(x^{k-1}, x^k) \quad \text{para } k = 1, 2, \ldots \tag{5.40}$$

De (5.40) y (5.37)

$$\phi'(f(x^{N_0})) \geq \frac{1}{\left(\sigma + \tilde{\lambda} M\right) q(x^{N_0-1}, x^{N_0})}. \tag{5.41}$$

De (5.41) y (5.36) obtenemos para $k = N_0$

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \frac{\frac{1}{2}\lambda_{N_0+1} q^2(x^{N_0}, x^{N_0+1})}{(\sigma + \lambda_+ M)\, q(x^{N_0-1}, x^{N_0})}.$$

factorizando

$$\phi(f(x^{N_0})) - \phi(f(x^{N_0+1})) \geq \frac{q(x^{N_0}, x^{N_0+1})}{q(x^{N_0-1}, x^{N_0})} \left(\frac{\frac{1}{2}\lambda_{N_0+1} q(x^{N_0}, x^{N_0+1})}{\sigma + \lambda_+ M} \right). \tag{5.42}$$

fijando un $r \in (0, 1)$ consideramos dos casos:
caso 1: Si $\frac{q(x^{N_0}, x^{N_0+1})}{q(x^{N_0-1}, x^{N_0})} \geq 1$ entonces tenemos

$$q(x^{N_0}, x^{N_0+1}) \leq 2 \left(\frac{\sigma + \lambda_+ M}{r\lambda_{N_0+1}} \right) [\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))]$$

Caso 2: Ahora, si $\frac{q(x^{N_0},x^{N_0+1})}{q(x^{N_0-1},x^{N_0})} \leq 1$, entonces

$$q(x^{N_0}, x^{N_0+1}) \leq rq(x^{N_0-1}, x^{N_0}).$$

De ambos casos obtenemos

$$q(x^{N_0}, x^{N_0+1}) \leq rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_{N_0+1}}\right)[\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))]$$

Finalmente como $\lambda_- < \lambda_k < \lambda_+$ se obtiene

$$q(x^{N_0}, x^{N_0+1}) \leq rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_-}\right)[\phi(f(x^{N_0})) - \phi(f(x^{N_0+1}))] \quad (5.43)$$

Probaremos por inducción que $x^{N_0+j} \in B_q(\bar{x}, \delta)$ para $j \in I\!N$
Para $j = 1$. Por la desigualdad triangular se tiene

$$q(\bar{x}, x^{N_0+1}) \leq q(\bar{x}, x^{N_0}) + q(x^{N_0}, x^{N_0+1})$$

por (5.17)

$$q(\bar{x}, x^{N_0+1}) \leq q(\bar{x}, x^{N_0}) + \sqrt{\frac{2}{\lambda_-} f(x^{N_0}) - f(x^{N_0+1})} < \delta$$

donde la última desigualdad es debido a (5.31), lo cual establece que $x^{N_0+1} \in B_q(\bar{x}, \delta)$.
Supongamos que (5.32) cumple para todo $k = N_0 + 1, \ldots, N_0 + j - 1$. Probaremos que se cumple para $N_0 + j$.
Observemos que de forma análoga a lo realizado para x^{N_0}, podemos verificar que (5.43) cumple para $N_0 + 1, \ldots, N_0 + j - 1$ y por lo tanto

$$\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \leq r \sum_{k=N_0}^{N_0+j-1} q(x^{k-1}, x^k) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_-}\right) \sum_{k=N_0}^{N_0+j-1} [\phi(f(x^k)) - \phi(f(x^{k+1}))]$$

Restando a ambas desigualdades el término $-r\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1})$ y usando la no negatividad de la cuasi-distancia obtenemos

$$(1-r)\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \leq rq(x^{N_0-1}, x^{N_0}) + 2\left(\frac{\sigma + \lambda_+ M}{r\lambda_-}\right) \sum_{k=N_0}^{N_0+j-1} [\phi(f(x^k)) - \phi(f(x^{k+1}))].$$

Dividiendo por $(1 - r)$ y denotando $\overline{M} = 2\left(\frac{\sigma+\lambda_+ M}{r(1-r)\lambda_-}\right)$

$$\sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \leq (\frac{r}{1-r}) q(x^{N_0-1}, x^{N_0}) + \overline{M}[\phi(f(x^{N_0})) - \phi(f(x^{N_0+j}))] \quad (5.44)$$

Estamos ahora en condiciones de probar que la propiedad es verdad para $N_0 + j$.
De la desigualdad triangular, tenemos

$$\begin{aligned} q(\bar{x}, x^{N_0+j}) &\leq q(\bar{x}, x^{N_0}) + q(x^{N_0}, x^{N_0+1}) + \ldots + q(x^{N_0+j-1}, x^{N_0+j}) \\ &\leq q(\bar{x}, x^{N_0}) + \sum_{k=N_0}^{N_0+j-1} q(x^k, x^{k+1}) \\ &\leq q(\bar{x}, x^{N_0}) + \left(\frac{r}{1-r}\right) q(x^{N_0}, x^{N_0+1}) + \overline{M}[\phi(f(x^{N_0})) - \phi(f(x^{N_0+j}))] \end{aligned}$$

La última desigualdad se cumple por (5.44).
Usando el hecho que $-\phi(f(x^{N_0+j})) < 0$, obtenemos

$$q(\bar{x}, x^{N_0+j}) \leq q(\bar{x}, x^{N_0}) + \frac{r}{1-r}\sqrt{\frac{2}{\lambda_-}f(x^{N_0-1})} + \overline{M}[\phi(f(x^{N_0}))]$$

la cual, de (5.31), nos permite concluir la prueba por inducción.
Observe que de (5.44) cuando $j \to \infty$ se obtiene (5.33). Por otro lado, por $(\mathbf{H_3})$ se tiene $\alpha||x^k - x^{k+1}|| \leq q(x^k, x^{k+1})$, combinado con la condición (5.33) se tiene

$$\sum_{k=N_0}^{\infty} ||x^k - x^{k+1}|| < +\infty$$

concluimos que $\{x^k\}$ es una sucesión de Cauchy y por lo tanto $\{x^k\}$ es convergente a $\bar{x}$ que es un punto crítico según el Teorema 5.3.4. □

5.3.6. Algunos ejemplos

Presentamos algunos ejemplos donde se hace uso del algoritmo de punto proximal inexacto.

Ejemplo 5.3.11. *Dada la siguiente función*

$$f(x) = (x-2)^2$$

Planteamos a continuación un problema de optimización

$$\begin{array}{c} \text{mín } (x-2)^2 \\ s.a. \;\; x \in I\!R \end{array}$$

que verifica las condiciones $(\mathbf{H_1})$, $(\mathbf{H_2})$, $(\mathbf{H_3})$ *dadas por el algoritmo.*
Consideremos un caso particular del Ejemplo 4.1.3, donde la siguiente función es una cuasi-distancia.

$$q(x,y) = \begin{cases} 3(x-y) & ; x > y \\ y - x & ; x \leq y \end{cases}$$

Vamos a resolver el siguiente subproblema dado por el algoritmo

$$(P_k) \begin{cases} \text{mín } f_k(x) = (x-2)^2 + \frac{\lambda_{k+1}}{2}q^2(x^k, x) \\ s.a. \quad x \in I\!R \end{cases}$$

(1) Dado $x_0 = 0$ *y* $\lambda_{k+1} = \frac{2+k}{1+k}$.

(2) a) Para $k = 0$ *tenemos* $x_0 = 0$ *y* $\lambda_1 = 2$.

$$(P_0) \begin{cases} \text{mín } f_0(x) = (x-2)^2 + \frac{2}{2}q^2(0, x) \\ s.a. \quad x \in I\!R \end{cases}$$

Consideramos dos casos para q.

i) Si $x > 0$ *entonces* $f_0(x) = (x-2)^2 + (x-0)^2$.
Ya que f_0 *es diferenciable*

$$\nabla f_0(x) = 2(x-2) + 2x = 0 \Rightarrow x = 1$$

ii) Si $x < 0$ *entonces* $f_0(x) = (x-2)^2 + (3(0-x))^2$.
Ya que f_0 *es diferenciable*

$$\nabla f_0(x) = 2(x-2) + 18x = 0 \Rightarrow x = \frac{1}{5}$$

La afirmación se verifica con el item i), entonces se obtiene el siguiente punto consecutivo $x^1 = 1$.
Debemos de verificar (5.17)
Dados

- $x_0 = 0$, $x_1 = 1$, $\lambda_1 = 2$.

$$f(x^1) \leq f(x^0) - \frac{\lambda_1}{2} q^2(x^0, x^1) \Rightarrow 1 \leq 4 - \frac{2}{2}(1)^2$$

Lo cual verifica la desigualdad.

Ahora debemos de calcular el error vectorial e^k *de*

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$,
Por otro lado f *es diferenciable,* $\nabla f(x) = 2(x-2) \Rightarrow \nabla f(x^1) = \nabla f(1) = -2$

$$\varphi^1 = -2 \in \partial f(x^1) = \nabla f(x^1)$$

Como $q(0, .)(x)$ *es una función diferenciable en* $x \neq 0$ *entonces* $\nabla q(0, x) = 1$ *en* $x^1 = 1 > 0$ *y se tiene*

$$\psi^1 = 1 \in \partial q(0, x^1) = \nabla q(0, x^1)$$

$$e^1 = \varphi^1 + \lambda_1 q(x^0, x^1) \psi^1 \Rightarrow e^1 = -2 + 2 \Rightarrow e^1 = 0$$

El error e^{k+1} *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.

b) para $k = 1$ *tenemos que* $x^1 = 1$ *y* $\lambda_2 = \frac{3}{2}$ *(* $\lambda_{k+1} = \frac{2+k}{1+k}$*).*

$$(P_1) \begin{cases} \text{mín } f_1(x) = (x-2)^2 + \frac{3}{4} q^2(1, x) \\ s.a. \quad x \in I\!R \end{cases}$$

Consideramos dos casos para q.

i) Si $x > 1$ *entonces* $f_1(x) = (x-2)^2 + \frac{3}{4}(x-1)^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x) = 2(x-2) + \frac{3}{4} 2(x-1) = 0 \Rightarrow x = \frac{11}{7}.$$

ii) Si $x < 1$ *entonces* $f_1(x) = (x-2)^2 + \frac{3}{4}(3(1-x))^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x) = 2(x-2) + \frac{3}{4}18(x-1) = 0 \Rightarrow x = \frac{35}{31}.$$

La afirmación se verifica con el item i), entonces se obtiene el siguiente punto consecutivo $x^2 = \frac{11}{7}$.
Debemos verificar (5.17)

- $x^1 = 1$, $x^2 = 11/7$, $\lambda_2 = 3/2$.

$$f(x^2) \leq f(x^1) - \frac{\lambda_2}{2} q^2(x^1, x^2) \Rightarrow \frac{9}{49} \leq 1 - \frac{3}{4}\left(\frac{4}{7}\right)^2.$$

Por otro lado f *es diferenciable,* $\nabla f(x) = 2(x-2) \Rightarrow \nabla f(x^2) = \nabla f(\frac{11}{7}) = -\frac{6}{7}$

$$\varphi^2 = -\frac{6}{7} \in \partial f(x^2) = \nabla f(x^2)$$

Debemos de verificar

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$.
Como $q(1, .)(x)$ *es una función diferenciable en* $x \neq 1$ *entonces* $\nabla q(1, x) = 1$ *en* $x^2 = 11/7 > 1$ *y se tiene*

$$\psi^2 = 1 \in \partial q(1, x^2) = \nabla q(1, x^2)$$

$$e^2 = \varphi^2 + \lambda_2 q(x^1, x^2)\psi^2 \Rightarrow e^2 = -\frac{6}{7} + \frac{6}{7} \Rightarrow e^2 = 0$$

El error e^2 *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
c) Para $k = 2$ *tenemos que* $x^2 = \frac{11}{7}$ *y* $\lambda_3 = \frac{4}{3}$ *donde* $\lambda_{k+1} = \frac{2+k}{1+k}$.

$$(P_2) \begin{cases} \text{mín } f_2(x) = (x-2)^2 + \frac{4}{6}q^2(\frac{11}{7}, x) \\ s.a. \quad x \in I\!R \end{cases}$$

Consideramos dos casos para q.

i) Si $x > \frac{11}{7}$ *entonces* $f_2(x) = (x-2)^2 + \frac{4}{6}(x - \frac{11}{7})^2$.
Ya que f_2 *es diferenciable*

$$\nabla f_2(x) = 2(x-2) + \frac{4}{6}2(x - \frac{11}{7}) = 0 \Rightarrow x = \frac{128}{70}.$$

ii) Si $x < \frac{11}{7}$ *entonces* $f_2(x) = (x-2)^2 + \frac{4}{3}(3(\frac{11}{7} - x))^2$
Ya que f_2 *es diferenciable*

$$\nabla f_2(x) = 2(x-2) + 24(x - \frac{11}{7}) = 0 \Rightarrow x = \frac{292}{182}.$$

La cual verifica el item i), entonces $x^3 = \frac{128}{70}$, *además* $\lambda_3 = 4/3$.
Debemos verificar

$$\boxed{f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1})}$$

- $x^2 = 11/7$, $x^3 = 128/70$, $\lambda_3 = 4/3$.

$$f(x^3) \leq f(x^2) - \frac{\lambda_3}{2} q^2(x^2, x^3) \Rightarrow \frac{144}{4900} \leq \frac{9}{49} - \frac{4}{6}\left(\frac{128}{70} - \frac{11}{7}\right)^2.$$

Por otro lado f *es diferenciable,* $\nabla f(x) = 2(x-2) \Rightarrow \nabla f(x^3) = \nabla f(\frac{128}{70}) = -\frac{24}{70}$

$$\varphi^3 = -\frac{24}{70} \in \partial f(x^3) = \nabla f(x^3).$$

Debemos de verificar

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$.
Como $q(\frac{11}{7}, .)(x)$ *es una función diferenciable en* $x \neq \frac{11}{7}$ *entonces* $\nabla q(\frac{11}{7}, x) = 1$ *en* $x^3 = 128/70 > \frac{11}{7}$ *y se tiene*

$$\psi^3 = 1 \in \partial q(\frac{11}{7}, x) = \nabla q(\frac{11}{7}, x)$$

$$e^3 = \varphi^3 + \lambda_3 q(x^2, x^3)\psi^3 \Rightarrow e^3 = -\frac{24}{70} + \frac{24}{70} \Rightarrow e^3 = 0.$$

El error e^3 *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
d) Para $k = 3$ *tenemos que* $x^3 = \frac{128}{70}$ *y* $\lambda_4 = \frac{5}{4}$ *donde* $\lambda_{k+1} = \frac{2+k}{1+k}$.

$$(P_3) \begin{cases} \min f_3(x) = (x-2)^2 + \frac{5}{8} q^2(\frac{128}{70}, x) \\ s.a. \quad x \in I\!R \end{cases}$$

Consideramos dos casos para q.

i) Si $x > \frac{128}{70}$ *entonces* $f_3(x) = (x-2)^2 + \frac{5}{8}(x - \frac{128}{70})^2$
Ya que f_2 *es diferenciable*

$$\nabla f_3(x) = 2(x-2) + \frac{5}{8} 2(x - \frac{128}{70}) = 0 \Rightarrow x = \frac{176}{91}$$

ii) Si $x < \frac{128}{70}$ *entonces* $f_3(x) = (x-2)^2 + \frac{5}{8}(3(\frac{128}{70} - x))^2$
Ya que f_3 *es diferenciable*

$$\nabla f_3(x) = 2(x-2) + \frac{45}{4}(x - \frac{128}{70}) = 0 \Rightarrow x = \frac{3440}{1855}.$$

La cual verifica el item i), entonces $x^4 = \frac{176}{91}$, *además* $\lambda_4 = 5/4$.
Debemos verificar

$$\boxed{f(x^{k+1}) \leq f(x^k) - \frac{\lambda_{k+1}}{2} q^2(x^k, x^{k+1})}$$

- $x^3 = 128/70$, $x^4 = 176/91$, $\lambda_4 = 5/4$.

$$f(x^4) \leq f(x^3) - \frac{\lambda_4}{2} q^2(x^3, x^4) \Rightarrow \frac{36}{8281} \leq \frac{144}{4900} - \frac{5}{8}\left(\frac{176}{91} - \frac{128}{70}\right)^2$$

Por otro lado f *es diferenciable,* $\nabla f(x) = 2(x-2) \Rightarrow \nabla f(x^4) = \nabla f(\frac{176}{91}) = -\frac{12}{91}$

$$\varphi^4 = -\frac{12}{91} \in \partial f(x^4) = \nabla f(x^4)$$

Debemos de verificar

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$.
Como $q(\frac{128}{70}, .)(x)$ *es una función diferenciable en* $x \neq \frac{128}{70}$ *entonces* $\nabla q(\frac{128}{70}, x) = 1$ *en* $x = \frac{176}{91} > \frac{128}{70}$ *y se tiene*

$$\psi^4 = 1 \in \partial q(\frac{128}{70}, x^4) = \nabla q(\frac{128}{70}, x^4)$$

$$e^4 = \varphi^4 + \lambda_4 q(x^3, x^4)\psi^4 \Rightarrow e^4 = -\frac{12}{91} + \frac{12}{91} \Rightarrow e^4 = 0$$

El error e^4 *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
Así sucesivamente se sigue la iteración.

Ejemplo 5.3.12. *Dado la siguiente función* $f(x, y) = x^2 - 1 + (y-2)^2$
Planteamos a continuación un problema de optimización

$$\begin{array}{c} \text{mín} : x^2 - 1 + (y-2)^2 \\ s.a. \ \ (x, y) \in I\!R^2 \end{array}$$

que verifica las condiciones $(\mathbf{H_1})$, $(\mathbf{H_2})$, $(\mathbf{H_3})$ *dadas por el algoritmo.*
Consideremos un caso particular del Ejemplo 4.1.3, donde la siguiente función es una cuasi-distancia.

$$q_1(\bar{x}, x) = \begin{cases} \bar{x} - x & ; \bar{x} > x \\ 2(x - \bar{x}) & ; \bar{x} < x \end{cases}$$

$$q_2(\bar{y}, y) = \begin{cases} \frac{1}{2}(\bar{y} - y) & ; \bar{y} > y \\ y - \bar{y} & ; \bar{y} < y \end{cases}$$

donde

$$q\left((\bar{x}, \bar{y}), (x, y)\right) = \sum_{i=1}^{2} q_i(\bar{x}_i, x_i)$$

Vamos a resolver el siguiente subproblema dado por el algoritmo

$$(P_k) \begin{cases} \text{mín } f_k(x) = x^2 - 1 + (y-2)^2 + \frac{\lambda_{k+1}}{2} q^2(\bar{x}^k, (x, y)) \\ s.a. \quad (x, y) \in \mathbb{R}^2 \end{cases}$$

(1) Dado $x^0 = (2,3)$ *y* $\lambda_{k+1} = \frac{k+1}{2k+1}$.

(2) a) Para $k = 0$ *tenemos* $x^0 = (2,3)$ *y* $\lambda_1 = 1$.

$$(P_0)\begin{cases} \text{mín } f_0(x) = x^2 - 1 + (y-2)^2 + \frac{1}{2}\left[q_1(2,x) + q_2(3,y)\right]^2 \\ s.a. \quad (x.y) \in \mathbb{R}^2 \end{cases}$$

i) Si $x > 2, y > 3$ *entonces* $f_0(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{2}\left[2(x-2) + y - 3\right]^2$. *Ya que* f_0 *es diferenciable*

$$\nabla f_0(x,y) = (6x + 2y - 14; 2x + 3y - 11) = (0,0) \Rightarrow (x,y) = (10/7, 19/7)$$

ii) Si $x < 2, y < 3$ *entonces* $f_0(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{2}\left[2 - x + \frac{1}{2}(3-y)\right]^2$. *Ya que* f_0 *es diferenciable*

$$\nabla f_0(x,y) = (3x + \frac{y}{2} - \frac{7}{2}; \frac{x}{2} + \frac{9y}{4} - \frac{23}{4}) \Rightarrow (x,y) = (\frac{10}{3}; \frac{31}{13})$$

.

iii) Si $x < 2, y > 3$ *entonces* $f_0(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{2}\left[2 - x + y - 3\right]^2$. *Ya que* f_0 *es diferenciable*

$$\nabla f_0(x,y) = (3x - y + 1; 3y - x - 5) = (0,0) \Rightarrow (x,y) = (\frac{1}{4}; \frac{7}{4})$$

iv) Si $x > 2, y < 3$ *entonces* $f_0(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{2}\left[2(x-2) + \frac{1}{2}(3-y)\right]^2$. *Ya que* f_0 *es diferenciable*

$$\nabla f_0(x,y) = (6x - y - 5; \frac{9y}{4} - x - \frac{11}{4}) = (0,0) \Rightarrow (x,y) = (\frac{28}{25}; \frac{43}{25})$$

La afirmación se verifica con el item ii), entonces se obtiene el siguiente punto consecutivo $x^1 = \left(\frac{10}{13}; \frac{31}{13}\right)$.
Debemos de verificar (5.17). Dado

- $x^0 = (2,3)$, $x^1 = (\frac{10}{13}; \frac{31}{13})$, $\lambda_1 = 1$.

$$f(x^1) \leq f(x^0) - \frac{\lambda_1}{2} q^2(x^0, x^1) \Rightarrow \frac{-44}{169} \leq 4 - \frac{1}{2}(20/13)^2$$

Lo cual verifica la desigualdad.

Ahora debemos de calcular el error vectorial e^k *de*

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$,
Por otro lado f *es diferenciable,* $\nabla f(x.y) = (2x, 2(y-2)) \Rightarrow \nabla f(x^1) = \nabla f(\frac{10}{13}, \frac{31}{13}) = (\frac{20}{13}, \frac{10}{13})$

$$\varphi^1 = (\frac{20}{13}, \frac{10}{13}) \in \partial f(x^1) = \nabla f(x^1)$$

Como $q(x^0,.)(x) = 2 - x + \frac{1}{2}(3-y)$ *es una función diferenciable en* $x \neq 0$ *entonces* $\nabla q(x^0,x) = (-1,-\frac{1}{2})$ *se tiene*

$$\psi^1 = (-1,-\frac{1}{2}) \in \partial q(x^0,x^1) = \nabla q(x^0,x^1)$$

$$e^1 = \varphi^1 + \lambda_1 q(x^0,x^1)\psi^1 \Rightarrow e^1 = (\frac{20}{13},\frac{10}{13}) + \frac{40}{26}(-1,-\frac{1}{2}) \Rightarrow e^1 = (0,0)$$

El error e^{k+1} *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k,x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
b) Para $k = 1$ *tenemos* $x^1 = (\frac{10}{13},\frac{31}{13})$ *y* $\lambda_2 = \frac{2}{3}$.

$$(P_1)\begin{cases} \text{mín } f_1(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{3}\left[q_1(\frac{10}{13},x) + q_2(\frac{31}{13},y)\right]^2 \\ s.a. \quad (x.y) \in I\!R^2 \end{cases}$$

i) Si $x > \frac{10}{13}, y > \frac{31}{13}$ *entonces* $f_1(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{3}\left[2(x-\frac{10}{13}) + y - \frac{31}{13}\right]^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x,y) = (\frac{14x}{3} + \frac{4y}{3} - \frac{68}{13}; \frac{4x}{3} + \frac{8y}{3} - \frac{86}{13}) = (0,0) \Rightarrow (x,y) = (25/52, 233/104)$$

ii) Si $x < \frac{10}{13}, y < \frac{31}{13}$ *entonces* $f_1(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{3}\left[\frac{10}{13} - x + \frac{1}{2}(\frac{31}{13} - y)\right]^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x,y) = (\frac{8x}{3} + \frac{y}{3} - \frac{17}{13}; \frac{x}{3} + \frac{13y}{6} - \frac{121}{26}) \Rightarrow (x,y) = (\frac{50}{221}; \frac{467}{221})$$

.

iii) Si $x < \frac{10}{13}, y > \frac{31}{13}$ *entonces* $f_1(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{3}\left[\frac{10}{13} - x + y - \frac{31}{13}\right]^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x,y) = (\frac{8x}{3} - \frac{2y}{3} + \frac{14}{13}; \frac{8y}{3} - \frac{2x}{3} - \frac{66}{13}) = (0,0) \Rightarrow (x,y) = (\frac{1}{13}; \frac{25}{13})$$

iv) Si $x > \frac{10}{13}, y < \frac{31}{13}$ *entonces* $f_1(x,y) = x^2 - 1 + (y-2)^2 + \frac{1}{3}\left[2(x-\frac{10}{13}) + \frac{1}{2}(\frac{31}{13} - y)\right]^2$.
Ya que f_1 *es diferenciable*

$$\nabla f_1(x,y) = (\frac{14x}{3} - \frac{2y}{3} - \frac{6}{13}; \frac{13y}{6} - \frac{2x}{3} - \frac{101}{26}) = (0,0) \Rightarrow (x,y) = (\frac{140}{377}; \frac{719}{377})$$

La afirmación se verifica con el item ii), entonces se obtiene el siguiente punto consecutivo $x^2 = \left(\frac{50}{221}; \frac{467}{221}\right)$.
Debemos de verificar (5.17). Dado

- $x^1 = (\frac{10}{13},\frac{31}{13})$, $x^2 = (\frac{50}{221}; \frac{467}{221})$, $\lambda_2 = \frac{2}{3}$.

$$f(x^2) \leq f(x^1) - \frac{\lambda_1}{2}q^2(x^1,x^2) \Rightarrow -\frac{45716}{48841} \leq -\frac{44}{169} - \frac{1}{2}(20/13)^2$$

Lo cual verifica la desigualdad.

Ahora debemos de calcular el error vectorial e^k *de*

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$,
Por otro lado f *es diferenciable,* $\nabla f(x.y) = (2x, 2(y-2)) \Rightarrow \nabla f(x^1) = \nabla f(\frac{10}{13}, \frac{31}{13}) = (\frac{20}{13}, \frac{10}{13})$

$$\varphi^1 = (\frac{20}{13}, \frac{10}{13}) \in \partial f(x^1) = \nabla f(x^1)$$

Como $q(x^0, .)(x) = \frac{10}{13} - x + \frac{1}{2}(\frac{31}{13} - y)$ *es una función diferenciable en* $x \neq 0$ *entonces* $\nabla q(x^0, x) = (-1, -\frac{1}{2})$ *se tiene*

$$\psi^1 = (-1, -\frac{1}{2}) \in \partial q(x^0, x^1) = \nabla q(x^0, x^1)$$

$$e^1 = \varphi^1 + \lambda_1 q(x^0, x^1)\psi^1 \Rightarrow e^1 = (\frac{20}{13}, \frac{10}{13}) + \frac{40}{26}(-1, -\frac{1}{2}) \Rightarrow e^1 = (0, 0)$$

El error e^{k+1} *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
c) Para $k = 2$ *tenemos* $x^2 = (\frac{50}{221}, \frac{467}{221})$ *y* $\lambda_3 = \frac{3}{5}$. *(*$\lambda_{k+1} = \frac{k+1}{2k+1}$*)*

$$(P_2) \begin{cases} \min f_2(x, y) = x^2 - 1 + (y-2)^2 + \frac{3}{10}\left[q_1(\frac{50}{221}, x) + q_2(\frac{467}{221}, y)\right]^2 \\ s.a. \quad (x.y) \in \mathbb{R}^2 \end{cases}$$

i) Si $x > \frac{50}{221}, y > \frac{467}{221}$ *entonces* $f_2(x, y) = x^2 - 1 + (y-2)^2 + \frac{3}{10}\left[2(x - \frac{50}{221}) + y - \frac{467}{221}\right]^2$.
Ya que f_2 *es diferenciable*

$$\nabla f_2(x, y) = (\frac{22x}{5} + \frac{6y}{5} - \frac{3402}{1105}; \frac{6x}{5} + \frac{13y}{5} - \frac{6121}{1105}) = (0, 0) \Rightarrow (x, y) = (30/221, 457/221)$$

ii) Si $x < \frac{50}{221}, y < \frac{467}{221}$ *entonces* $f_2(x, y) = x^2 - 1 + (y-2)^2 + \frac{3}{10}\left[\frac{50}{221} - x + \frac{1}{2}(\frac{467}{221} - y)\right]^2$.
Ya que f_2 *es diferenciable*

$$\nabla f_2(x, y) = (\frac{13x}{5} + \frac{3y}{10} - \frac{1701}{2210}; \frac{3x}{10} + \frac{43y}{20} - \frac{19381}{4420}) \Rightarrow (x, y) = (\frac{150}{2431}; \frac{4937}{2431})$$

.

iii) Si $x < \frac{50}{221}, y > \frac{467}{221}$ *entonces* $f_2(x, y) = x^2 - 1 + (y-2)^2 + \frac{3}{10}\left[\frac{50}{221} - x + y - \frac{467}{221}\right]^2$.
Ya que f_2 *es diferenciable*

$$\nabla f_2(x, y) = (\frac{13x}{5} - \frac{3y}{5} + \frac{1251}{1105}; \frac{13y}{5} - \frac{3x}{5} - \frac{5671}{1105}) = (0, 0) \Rightarrow (x, y) = (\frac{75}{3536}; \frac{6997}{3536})$$

iv) Si $x > \frac{50}{221}, y < \frac{467}{221}$ *entonces* $f_2(x, y) = x^2 - 1 + (y-2)^2 + \frac{3}{10}\left[2(x - \frac{50}{221}) + \frac{1}{2}(\frac{467}{221} - y)\right]^2$.
Ya que f_2 *es diferenciable*

$$\nabla f_2(x, y) = (\frac{22x}{5} - \frac{3y}{5} - \frac{801}{1105}; \frac{43y}{20} - \frac{3x}{5} - \frac{18481}{4420}) = (0, 0) \Rightarrow (x, y) = (\frac{300}{2873}; \frac{5671}{2873})$$

La afirmación se verifica con el item ii), entonces se obtiene el siguiente punto consecutivo $x^3 = \left(\frac{150}{2431}; \frac{4937}{2431}\right)$.
Debemos de verificar (5.17). Dado

- $x^2 = (\frac{50}{221}; \frac{467}{221})$, $x^3 = (\frac{150}{2431}; \frac{4937}{2431})$, $\lambda_3 = \frac{3}{5}$.

$$f(x^3) \leq f(x^2) - \frac{\lambda_3}{2} q^2(x^2, x^3) \Rightarrow -\frac{5881636}{5909761} \leq -\frac{45716}{48841} - \frac{3}{10}(500/2431)^2$$

Lo cual verifica la desigualdad.

Ahora debemos de calcular el error vectorial e^k *de*

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1})\psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$,
Por otro lado f *es diferenciable,* $\nabla f(x.y) = (2x, 2(y-2)) \Rightarrow \nabla f(x^3) = \nabla f(\frac{150}{2431}, \frac{4937}{2431}) = (\frac{300}{2431}, \frac{150}{2431})$

$$\varphi^3 = (\frac{300}{2431}, \frac{150}{2431}) \in \partial f(x^3) = \nabla f(x^3)$$

Como $q(x^2, .)(x) = \frac{50}{221} - x + \frac{1}{2}(\frac{467}{221} - y)$ *es una función diferenciable entonces* $\nabla q(x^2, x) = (-1, -\frac{1}{2})$ *se tiene*

$$\psi^3 = (-1, -\frac{1}{2}) \in \partial q(x^2, x^3) = \nabla q(x^2, x^3)$$

$$e^3 = \varphi^3 + \lambda_3 q(x^2, x^3)\psi^3 \Rightarrow e^3 = (\frac{300}{2431}, \frac{150}{2431}) + \frac{300}{2431}(-1, -\frac{1}{2}) \Rightarrow e^3 = (0, 0)$$

El error e^{k+1} *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
d) Para $k = 3$ *tenemos* $x^3 = (\frac{150}{2431}, \frac{4937}{2431})$ *y* $\lambda_4 = \frac{4}{7}$. *(*$\lambda_{k+1} = \frac{k+1}{2k+1}$*)*

$$(P_3) \begin{cases} \text{mín } f_3(x, y) = x^2 - 1 + (y-2)^2 + \frac{2}{7}\left[q_1(\frac{150}{2431}, x) + q_2(\frac{4937}{2431}, y)\right]^2 \\ s.a. \quad (x.y) \in \mathbb{R}^2 \end{cases}$$

i) Si $x > \frac{150}{2431}, y > \frac{4937}{2431}$ *entonces* $f_3(x, y) = x^2 - 1 + (y-2)^2 + \frac{2}{7}\left[2(x - \frac{150}{2431}) + y - \frac{4937}{2431}\right]^2$.
Ya que f_3 *es diferenciable*

$$\nabla f_3(x, y) = (\frac{30x}{7} + \frac{8y}{7} - \frac{41896}{17017}; \frac{8x}{7} + \frac{18y}{7} - \frac{89016}{17017}) = (0, 0) \Rightarrow (x, y) = (\frac{1500}{41327}, \frac{83404}{41327})$$

ii) Si $x < \frac{150}{2431}, y < \frac{4937}{2431}$ *entonces* $f_3(x, y) = x^2 - 1 + (y-2)^2 + \frac{2}{7}\left[\frac{150}{2431} - x + \frac{1}{2}(\frac{4937}{2431} - y)\right]^2$.
Ya que f_3 *es diferenciable*

$$\nabla f_3(x, y) = (\frac{18x}{7} + \frac{2y}{7} - \frac{10474}{17017}; \frac{2x}{7} + \frac{15y}{7} - \frac{73305}{17017}) \Rightarrow (x, y) = (\frac{750}{46189}; \frac{92753}{46189})$$

.

iii) Si $x < \frac{150}{2431}, y > \frac{4937}{2431}$ *entonces* $f_3(x,y) = x^2 - 1 + (y-2)^2 + \frac{2}{7}\left[\frac{150}{2431} - x + y - \frac{4937}{2431}\right]^2$.
Ya que f_3 *es diferenciable*

$$\nabla f_3(x,y) = (\frac{18x}{7} - \frac{4y}{7} + \frac{19148}{17017}; \frac{18y}{7} - \frac{4x}{7} - \frac{87216}{17017}) = (0,0) \Rightarrow (x,y) = (\frac{150}{26741}; \frac{53332}{26741})$$

iv) Si $x > \frac{150}{2431}, y < \frac{4937}{2431}$ *entonces* $f_3(x,y) = x^2 - 1 + (y-2)^2 + \frac{2}{7}\left[2(x - \frac{150}{2431}) + \frac{1}{2}(\frac{4937}{2431} - y)\right]^2$.
Ya que f_3 *es diferenciable*

$$\nabla f_3(x,y) = (\frac{30x}{7} - \frac{4y}{7} + \frac{17348}{17017}; \frac{15y}{7} - \frac{4x}{7} - \frac{72405}{17017}) = (0,0) \Rightarrow (x,y) = (\frac{2100}{75361}; \frac{150197}{75361})$$

La afirmación se verifica con el item ii), entonces se obtiene el siguiente punto consecutivo $x^4 = \left(\frac{750}{46189}; \frac{92753}{46189}\right)$.
Debemos de verificar (5.17). Dado

- $x^3 = (\frac{150}{2431}; \frac{4937}{2431})$, $x^4 = (\frac{750}{46189}; \frac{92753}{46189})$, $\lambda_4 = \frac{4}{7}$.

$$f(x^4) \leq f(x^3) - \frac{\lambda_4}{2} q^2(x^3, x^4) \Rightarrow -\frac{2132720596}{2133423721} \leq -\frac{5881636}{5909761} - \frac{2}{7}(2625/46189)^2$$

Lo cual verifica la desigualdad.

Ahora debemos de calcular el error vectorial e^k *de*

$$\boxed{e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}}$$

donde $\varphi^{k+1} \in \partial f(x^{k+1}), \psi^{k+1} \in \partial q(x^k, .)(x^{k+1})$,
Por otro lado f *es diferenciable,* $\nabla f(x.y) = (2x, 2(y-2)) \Rightarrow \nabla f(x^4) = \nabla f(\frac{750}{46189}, \frac{92753}{46189}) = (\frac{1500}{46189}, \frac{750}{46189})$

$$\varphi^4 = (\frac{1500}{46189}, \frac{750}{46189}) \in \partial f(x^4) = \nabla f(x^4)$$

Como $q(x^3, .)(x) = \frac{150}{2431} - x + \frac{1}{2}(\frac{4937}{2431} - y)$ *es una función diferenciable entonces* $\nabla q(x^3, x) = (-1, -\frac{1}{2})$ *se tiene*

$$\psi^4 = (-1, -\frac{1}{2}) \in \partial q(x^3, x^4) = \nabla q(x^3, x^4)$$

$$e^4 = \varphi^4 + \lambda_4 q(x^3, x^4)\psi^4 \Rightarrow e^4 = (\frac{1500}{46189}, \frac{750}{46189}) + \frac{1500}{46189}(-1, -\frac{1}{2}) \Rightarrow e^4 = (0,0)$$

El error e^{k+1} *satisface:*

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

con $\sigma = \frac{1}{8} > 0$.
Para la cuarta iteración se sigue el mismo procedimiento y así sucesivamente.

Capítulo 6

VELOCIDAD DE CONVERGENCIA PARA FUNCIONES QUE SATISFACEN LA CONDICIÓN DE K.L.

En este capítulo se va a dar la tasa de convergencia de la sucesión generada por API que satisface la propiedad de Kurdyka Lojasiewicz.
De (5.44), para un número natural k suficientemente grande con $k \geq N_0$ se tiene

$$\sum_{p=k}^{N} q(x^p, x^{p+1}) \leq (\frac{r}{1-r}) q(x^{k-1}, x^k) + \overline{M}[\phi(f(x^k)) - \phi(f(x^{N+1}))] \tag{6.1}$$

cuando N tiende al infinito y tomando en cuenta que $f(x^k)$ decrece a cero, y $\phi(0) = 0$ en la desigualdad se tiene

$$\sum_{p=k}^{+\infty} q(x^p, x^{p+1}) \leq (\frac{r}{1-r}) q(x^{k-1}, x^k) + \overline{M}[\phi(f(x^k))] \quad \text{con } k \geq N_0. \tag{6.2}$$

Supongamos que la función concava dada en la propiedad de KL. este dado como $\phi(s) = s^{1-\theta}, \theta \in [0, 1)$ (aquí vamos considerar $s = f(x^k)$) y denotamos por $\Delta_k := \sum_{p=k}^{+\infty} q(x^p, x^{p+1})$. Así entonces se considera $\Delta_{k-1} - \Delta_k = q(x^{k-1}, x^k)$, y reescribiendo la desigualdad (6.2) se obtiene

$$\Delta_k \leq (\frac{r}{1-r})(\Delta_{k-1} - \Delta_k) + \overline{M}[(f(x^k))^{1-\theta}] \quad \text{con } k \geq N_0. \tag{6.3}$$

Aplicando la desigualdad (5.29) y la ecuación $\phi'(f(x^k)) = (1-\theta)[f(x^k)]^{-\theta}$ se obtiene

$$||\varphi^k||(1-\theta) \geq [f(x^k)]^{\theta} \quad \text{con } \varphi^k \in \partial f(x^k) \tag{6.4}$$

del Lema 5.1 de Moreno se tiene que existe $M > 0$ con la propiedad $||\psi^k|| \leq M$ ($\psi^k \in \partial q(x^{k-1}, .)(x^k)$) para $k \in \mathbb{N}$, y de la desigualdad (5.40) que reemplazado en (6.4) se tiene la siguiente desigualdad

$$\left[(1-\theta)(\sigma + \lambda_+ M) q(x^{k-1}, x^k)\right]^{\frac{1-\theta}{\theta}} \geq [f(x^k)]^{1-\theta}. \tag{6.5}$$

Reemplazando (6.5) en (6.3) y sabiendo $\Delta_{k-1} - \Delta_k = q(x^{k-1}, x^k)$

$$\Delta_k \leq \frac{r}{1-r}(\Delta_{k-1} - \Delta_k) + \overline{M}\left[(1-\theta)(\sigma + \lambda_+ M)(\Delta_{k-1} - \Delta_k)\right]^{\frac{1-\theta}{\theta}}, k > N_0 \quad (6.6)$$

Vamos a considerar en este trabajo $h(\theta) = \frac{1-\theta}{\theta} > 0, \theta \in [0, 1)$. La expresión (6.6) se transforma

$$\Delta_k \leq (\frac{r}{1-r})(\Delta_{k-1} - \Delta_k) + \overline{M}(\widetilde{C})^{h(\theta)}(\Delta_{k-1} - \Delta_k)^{h(\theta)}, \quad k > N_0 \quad (6.7)$$

donde $\widetilde{C} = (1-\theta)(\sigma + \lambda_+ M)$.

Teorema 6.0.1. *Asumiremos aquí que se satisface la función concava, dada por* $\phi(s) = s^{1-\theta},\ \theta \in [0, 1)$,*con* $h(\theta) = \frac{1-\theta}{\theta}$. *Asumiremos también todas las suposiciones de los anteriores teoremas ya vistos. Sea* $\bar{x}$ *el punto límite de la sucesión* $\{x^k\}$. *Entonces*

(i) Si $\theta = 0$, *la sucesión* $\{x^k\}$ *converge en un número finito de pasos;*

(ii) Si $h(\theta) \geq 1$, *entonces existe* $c_1 > 0,\ Q \in (0, 1)$ *y* $k_1 \in \mathbb{N}$ *tal que*

$$q(\bar{x}, x^k) \leq c_1 Q^k, \quad k \geq k_1;$$

(iii) Si $h(\theta) \in (0, 1)$, *entonces, existe* $c_2 > 0$ *tal que*

$$q(\bar{x}, x^k) \leq c_2 k^{\frac{h(\theta)}{h(\theta)-1}}, \quad k \geq k_1$$

Demostración. Primeramente, supongamos que $h(\theta) \in (0, 1)$, ya que $\Delta_{k-1} - \Delta_k$ es una sucesión que tiende a cero cuando k tiende al infinito, entonces existe $k_1 \geq N_0$ tal que

$$\Delta_{k-1} - \Delta_k \leq (\Delta_{k-1} - \Delta_k)^{h(\theta)}, \quad k \geq k_1$$

reemplazando la desigualdad anterior en la expresión (6.7)

$$\Delta_k \leq \left[\frac{r}{1-r} + \overline{M}(\widetilde{C})^{h(\theta)}\right](\Delta_{k-1} - \Delta_k)^{h(\theta)}, \quad k \geq k_1. \quad (6.8)$$

La expresión $\frac{r}{1-r} + \overline{M}(\widetilde{C})^{h(\theta)}$ es un número positivo y $\mathscr{C} = \left[\frac{r}{1-r} + \overline{M}(\widetilde{C})^{h(\theta)}\right]^{\frac{1}{h(\theta)}} > 0$, luego de (6.8) se tiene

$$(\Delta_k)^{\frac{1}{h(\theta)}} \leq \mathscr{C}(\Delta_{k-1} - \Delta_k), \quad k \geq k_1. \quad (6.9)$$

Aquí vamos a considerar la función $\Gamma : [0, +\infty) \to [0, +\infty)$ definida como $\Gamma(t) = t^{-\frac{1}{h(\theta)}}$, reescribiendo la expresión (6.9) obtenemos

$$1 \leq \mathscr{C}(\Delta_{k-1} - \Delta_k)\Gamma(\Delta_k). \quad (6.10)$$

Para $k \geq k_1$, sea $R \in (1, +\infty)$, se tiene los siguientes dos casos:

a) $\Gamma(\Delta_k) \leq R\Gamma(\Delta_{k-1})$

b) $\Gamma(\Delta_k) > R\Gamma(\Delta_{k-1})$

se establece que existe $\mu > 0$ y $\nu < 0$ tal que

$$\Delta_k^{\nu} - \Delta_{k-1}^{\nu} \geq \mu. \tag{6.11}$$

Supongamos que se cumpla a), entonces de (6.10), se sigue que

$$1 \leq R\mathscr{C}(\Delta_{k-1} - \Delta_k)\Gamma(\Delta_{k-1}). \tag{6.12}$$

Como la función Γ es diferenciable se tiene que es estrictamente decreciente y ya que $\triangle_k < \triangle_{k-1}$ se cumple $\Gamma(\Delta_{k-1}) = \min\limits_{s\in[\triangle_k,\triangle_{k-1}]} \Gamma(s)$, por lo cual se cumple

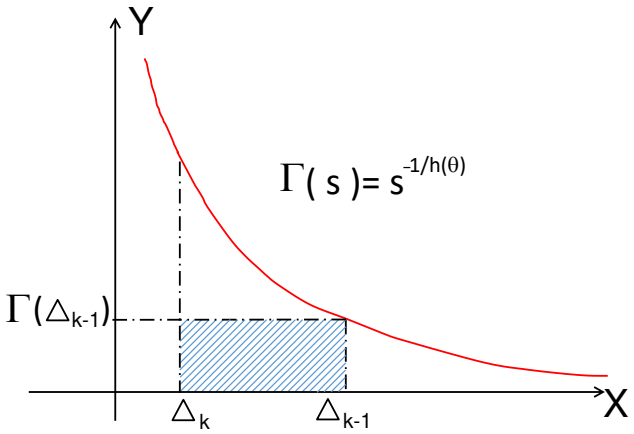


Figura 6.1: La región resaltada tiene por base $\triangle_{k-1} - \triangle_k$ y altura $\Phi(\triangle_{k-1})$.

de (6.12) se tiene

$$1 \leq R\mathscr{C} \int_{\Delta_k}^{\Delta_{k-1}} \Gamma(t)dt$$

integrando, se obtiene

$$1 \leq R\mathscr{C}\frac{h(\theta)}{h(\theta)-1}\left[\Delta_{k-1}^{\frac{h(\theta)-1}{h(\theta)}} - \Delta_k^{\frac{h(\theta)-1}{h(\theta)}}\right].$$

Observe que $h(\theta) \in (0,1)$ y esto nos lleva que $\frac{h(\theta)-1}{h(\theta)} < 0$ por lo tanto $\frac{1-h(\theta)}{h(\theta)} > 0$, luego se tiene la desigualdad (6.11) con

$$\mu = \frac{1-h(\theta)}{R\mathscr{C}h(\theta)} \quad \text{y} \quad \nu = \frac{h(\theta)-1}{h(\theta)}. \tag{6.13}$$

Caso b). Entonces de la definición de la función Γ y $h(\theta)$ se tiene $\Delta_k^{-\frac{1}{h(\theta)}} > R\Delta_{k-1}^{-\frac{1}{h(\theta)}}$, tomando la inversa tenemos $\Delta_{k-1}^{\frac{1}{h(\theta)}} > R\Delta_k^{\frac{1}{h(\theta)}}$, luego

$$\Delta_k < \left(\frac{1}{R}\right)^{h(\theta)} \Delta_{k-1}.$$

Considerando $p = \left(\frac{1}{R}\right)^{h(\theta)}$, entonces $\Delta_k \leq p\Delta_{k-1}$ y tomando una cantidad $\nu < 0$ obtenemos

$$\Delta_k^\nu \geq p^\nu \Delta_{k-1}^\nu.$$

Por lo tanto, restando a ambos lados de la desigualdad Δ_{k-1}^ν, se obtiene

$$\Delta_k^\nu - \Delta_{k-1}^\nu \geq (p^\nu - 1)\Delta_{k-1}^\nu.$$

Ya que $R \in (1, +\infty)$ entonces $p = \left(\frac{1}{R}\right)^{h(\theta)} \in (0, 1)$ y por lo tanto $p^\nu > 1$ (pues $\nu < 0$). Además como Δ_q tiende a cero cuando q va al infinito (de 6.3 se tiene que Δ_q es una serie convergente y se cumple dicha desigualdad), es decir, existe $q > k_1$ tal que $\Delta_q^\nu > \mu k_1$. De ambas afirmaciones se tiene que existe $\bar{\mu} > 0$ tal que $(p^\nu - 1)\Delta_{k-1}^\nu > \bar{\mu}$. Esto nos da el resultado deseado con $\mu = \bar{\mu}$ y $\nu < 0$ definido en (6.13). Por lo tanto se cumple la afirmación.
Para un $k \in \mathbb{N}$ mucho mayor que k_1. Sumando la desigualdad (6.11) desde k_1 hasta k se tiene

$$\sum_{i=k_1}^{k} (\Delta_i^\nu - \Delta_{i-1}^\nu) > \sum_{i=k_1}^{k} \mu. \quad \text{cuyo resultado es} \quad \Delta_k^\nu - \Delta_{k_1-1}^\nu > (k - k_1)\mu$$

sumando a ambos lados de la desigualdad $\Delta_{k_1-1}^\nu$ y elevando a $\frac{1}{\nu} < 0$ se obtiene

$$\Delta_k < \left[(k - k_1)\mu + \Delta_{k_1-1}^\nu\right]^{\frac{1}{\nu}}$$

Teniendo a $\nu < 0$, se cumple $\left[(k - k_1)\mu + \Delta_{k_1-1}^\nu\right]^{\frac{1}{\nu}} \leq ((k - k_1)\mu)^{\frac{1}{\nu}}$, se sigue que existe una constante ω dado por $\omega = \mu^{\frac{1}{\nu}}$ y por (6.13) tenemos $\frac{1}{\nu} = \frac{h(\theta)}{h(\theta)-1}$ tal que

$$\Delta_k \leq \omega(k - k_1)^{\frac{h(\theta)}{h(\theta)-1}}. \tag{6.14}$$

Por otro lado, combinando la condición ($\mathbf{H_3}$) y la definición de Δ_k, tenemos

$$\Delta_k \geq \alpha \lim_{N\to+\infty} \sum_{p=k}^{N} ||x^{p+1} - x^p|| \geq \alpha \lim_{N\to+\infty} \sum_{p=k}^{N} ||x^p - \bar{x}|| - ||x^{p+1} - \bar{x}|| = \alpha||x^k - \bar{x}|| \tag{6.15}$$

donde en la segunda desigualdad se utilizo la desigualdad triangular y en la igualdad la definición de $\bar{x}$. Así el item (iii) se sigue combinando las desigualdades (6.14) y (6.15) y la suposición en($\mathbf{H_3}$) con $c_2 = \frac{\beta\omega}{\alpha}$.
Supongamos que $h(\theta) \geq 1$, de (6.7) obtenemos

$$\Delta_k \leq \mathscr{C}_1(\Delta_{k-1} - \Delta_k), \quad k \geq k_1 \geq N_0 \tag{6.16}$$

donde $\mathscr{C}_1 = \mathscr{C}^{h(\theta)}$. Así la desigualdad anterior implica $\Delta_k \leq \frac{\mathscr{C}_1}{1+\mathscr{C}_1}\Delta_{k-1}$, para $k \geq k_1 \geq N_0$. Dado las siguientes expresiones

$$\begin{aligned}
\Delta_{k_1} &\leq \tfrac{\mathscr{C}_1}{1+\mathscr{C}_1}\Delta_{k_0} \\
\Delta_{k_2} &\leq \tfrac{\mathscr{C}_1}{1+\mathscr{C}_1}\Delta_{k_1} \\
&\vdots \\
\Delta_k &\leq \tfrac{\mathscr{C}_1}{1+\mathscr{C}_1}\Delta_{k-1}
\end{aligned}$$

multiplicando estas desigualdades y simplificando se obtiene $\Delta_k \leq \left(\frac{\mathscr{C}_1}{1+\mathscr{C}_1}\right)^{k-(k_1-1)} \Delta_{k_0}$, observe que $k_0 = k_1 - 1$. Resumiendo

$$\Delta_k \leq \omega_2 Q^k \tag{6.17}$$

donde $Q = \frac{\mathscr{C}_1}{1+\mathscr{C}_1} \in (0,1)$ y $\omega_2 = \frac{\Delta_{k_0}}{Q^{k_0}}$. Usando nuevamente (6.15), se sigue el item(ii) con $c_1 = \frac{\beta\,\omega_2}{\alpha}$.
Ahora, asumiendo que $\theta = 0$. Por contradicción, supongamos que la sucesión $\{x^k\}$ no converge a un número finito de pasos. Consideramos un $k_0 \in \mathbb{N}$ grande tal que $x^k \in B_q(\bar{x}, \delta) \cup [0 < f < \eta]$ para $k \geq k_0$ (estamos asumiendo que $f(\bar{x}) = s$). Ya que $\phi(s) = s$ de la desigualdad de KL, se tiene

$$1 \leq dist(0, \partial f(x^k)), \quad k \geq k_0 \tag{6.18}$$

De la sucesión $\{x^k\}$. para cada $k \in \mathbb{N}$ existe $\varphi^k \in \partial f(x^k)$ tal que la expresión (6.18) nos da $||\varphi^k|| \geq 1$. Por otro lado de (5.40) combinado con la desigualdad anterior se obtiene

$$1 \leq ||\varphi^k|| \leq (\sigma + \lambda_+ M) q(x^{k-1}, x^k) \tag{6.19}$$

Por la proposición (5.3.2) item (ii), se tiene que $q(x^{k-1}, x^k)$ tiende a cero cuando k va al infinito, esto implica que $1 \leq 0$ lo cual es una contradicción. □

Aquí vamos a dar una demostración análoga al artículo de Hedy Attouch referente al Lema 2.6, donde nos afirma que una función semicontinua inferior que satisface la propiedad de Kurdyka Lojasiewicz en un punto (no necesariamente un punto de acumulación), presenta una sucesión que converge a una vecindad del punto que satisface la propiedad de KL.

Lema 6.0.2. *Dado una función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ propia, semicontinua inferior que satisface la propiedad de KL en algún $x^* \in \mathbb{R}^n$. Denotamos por U, η y $\phi : [0, \eta\rangle \to \mathbb{R}_+$ los objetos que aparecen en la Definición 5.3.9 de la propiedad de KL en x^*. Sea δ, ρ con $\rho < \delta$ tal que $B_q(x^*, \delta) \subset U$.*
Considere la sucesión $\{x^k\}_{k\in\mathbb{N}}$ la cual satisface la condición (5.17)- (5.20), asuma además que

$$f(x^*) \leq f(x^0) < f(x^*) + \eta \tag{6.20}$$

$$q(x^*, x^0) + 2\sqrt{\frac{2[f(x^0) - f(x^*)]}{\lambda_-}} + \overline{M}\left[\phi(f(x^0) - f(x^*))\right] < \rho \tag{6.21}$$

con $\overline{M} = \frac{2(\sigma+\lambda_+ M)}{\lambda_-}$

$$\forall k \in \mathbb{N}, x^k \in B_q(x^*, \rho) \Rightarrow x^{k+1} \in B_q(x^*, \delta) \text{ con } f(x^{k+1}) \geq f(x^*) \tag{6.22}$$

Entonces la sucesión $\{x^k\}$ satisface

$$x^k \in B_q(x^*, \rho), \quad \forall k \in \mathbb{N}$$
$$f(x^k) \to f(x^*) \quad \text{cuando } k \to +\infty$$

y converge a un punto $\bar{x} \in B_q(x^*, \rho)$. tal que $f(\bar{x}) \leq f(x^*)$

Demostración. **i)** Probaremos inicialmente que si

$$f(x^k) < f(x^*) + \eta \quad \text{y} \quad x^k \in B_q(x^*, \rho)$$

implica

$$2q(x^k, x^{k+1}) \leq q(x^{k-1}, x^k) + \overline{M}[\phi(f(x^k) - f(x^*)) - \phi(f(x^{k+1}) - f(x^*))]$$

donde $\overline{M} = \frac{2(\sigma+\lambda_+ M)}{\lambda_-}$.
La desigualdad esta bien definida por la hipótesis (6.22) ($f(x^{k+1}) \geq f(x^*)$) y por (5.17) ($f(x^{k+1}) \leq f(x^k)$).
Si $x^{k+1} = x^k$ entonces se cumple trivialmente.
En efecto,

$$\begin{aligned} &q(x^k, x^{k+1}) = 0 \\ &q(x^{k-1}, x^k) \geq 0 \\ &\phi(f(x^k) - f(x^*)) - \phi(f(x^{k+1}) - f(x^*)) = 0 \end{aligned}$$

vamos a suponer que $x^k \neq x^{k+1}$ entonces de (5.17)

$$q(x^k, x^{k+1}) \leq \sqrt{\frac{2[f(x^k) - f(x^{k+1})]}{\lambda_{k+1}}} \quad \text{para } k = 0, 1, 2,$$

que nos afirma que $f(x^k) > f(x^{k+1})$ y como $f(x^{k+1}) \geq f(x^*)$ entonces

$$f(x^k) - f(x^*) > 0$$

Como $x^k \in B_q(x^*, \rho)$ y $\rho < \delta$ entonces $x^k \in U$. Además por hipótesis tenemos $f(x^*) < f(x^k) < f(x^*) + \eta$.
Así $x^k \in U \cap [f(x^*) < f < f(x^*) + \eta]$ lo cual implica por la desigualdad de KL

$$\phi'(f(x^k) - f(x^*)) dist(0, \partial f(x^k)) \geq 1$$

Por (5.18) se tiene que $\varphi^k \in \partial f(x^k)$ entonces

$$\phi'(f(x^k) - f(x^*)) ||\varphi^k|| \geq 1$$

esto es,

$$\phi'(f(x^k) - f(x^*)) \geq \frac{1}{||\varphi^k||}, \quad \text{pues } \varphi^k \neq 0 \tag{6.23}$$

pero por (5.20)

$$\varphi^k = e^k - \lambda_k q(x^{k-1}, x^k)\psi^k \quad \text{donde } \psi^k \in \partial q(x^{k-1}, .)(x^k)$$

Aplicando la norma y la desigualdad triangular

$$||\varphi^k|| \leq ||e^k|| + \lambda_k q(x^{k-1}, x^k)||\psi^k||$$

de (5.19), $\lambda_- \leq \lambda_k \leq \lambda_+$ y por Lema de Moreno existe $M > 0$ tal que $||\psi^k|| < M$.

$$\begin{aligned} ||\varphi^k|| \leq \quad & \sigma q(x^{k-1}, x^k) + \lambda_+ M q(x^{k-1}, x^k) \\ = \quad & (\sigma + \lambda_+ M) q(x^{k-1}, x^k). \end{aligned} \tag{6.24}$$

Aplicando en la desigualdad (6.23) tenemos

$$\phi'(f(x^k) - f(x^*)) \geq \frac{1}{(\sigma + \lambda_+ M) q(x^{k-1}, x^k)} \tag{6.25}$$

Ahora, por definición de KL existe una función concava $\phi : [0, \eta\rangle \to \mathbb{R}$ con $\eta > 0$, la expresión $f(x^k) - f(x^{k+1})$ es positiva por (5.17), para $s, t \in [0, \eta\rangle$ en ϕ se tiene

$$\phi(t) - \phi(s) \geq \phi'(t)(t - s). \tag{6.26}$$

Aplicando para $t = f(x^k) - f(x^*), s = f(x^{k+1}) - f(x^*)$

$$\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1} - f(x^*))\right) \geq \phi'\left(f(x^k) - f(x^*)\right)[f(x^k) - f(x^{k+1})].$$

Usando (6.25) y (5.17) se obtiene

$$\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1}) - f(x^*)\right) \geq \frac{\lambda_{k+1}}{2(\sigma + \lambda_+ M)} \frac{q^2(x^k, x^{k+1})}{q(x^{k-1}, x^k)} \tag{6.27}$$

usando nuevamente $\lambda_- \leq \lambda_{k+1}$ y recordando $\overline{M} = \frac{2(\sigma + \lambda_+ M)}{\lambda_-}$ se obtiene

$$\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1}) - f(x^*)\right) \geq \frac{1}{\overline{M}} \frac{q^2(x^k, x^{k+1})}{q(x^{k-1}, x^k)}$$

pasando a multiplicar

$$\overline{M} q(x^{k-1}, x^k)[\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1}) - f(x^*)\right)] \geq q^2(x^k, x^{k+1}).$$

Asociando y sacando raiz cuadrada a ambas expresiones positivas

$$\sqrt{\overline{M}[\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1}) - f(x^*)\right)]} \sqrt{q(x^{k-1}, x^k)} \geq q(x^k, x^{k+1}).$$

Como $\frac{a+b}{2} \geq \sqrt{a}\sqrt{b}$.
Se tiene

$$\overline{M}[\phi\left(f(x^k) - f(x^*)\right) - \phi\left(f(x^{k+1}) - f(x^*)\right)] + q(x^{k-1}, x^k) \geq 2q(x^k, x^{k+1}). \tag{6.28}$$

Esto completa la prueba de lo afirmado en el item i). **ii)** Ahora probaremos

$$x^j \in B_q(x^*, \rho), \quad \forall j \in \mathbb{N} \tag{6.29}$$

Probaremos la afirmación (6.29) por inducción sobre j. De (6.22) con $k = 0$ obtenemos que $x^1 \in B_q(x^*, \delta)$ y $f(x^1) \geq f(x^*)$, usando (5.17) y $\lambda_- < \lambda_k < \lambda_+$ obtenemos

$$q(x^0, x^1) \leq \sqrt{\frac{2[f(x^0) - f(x^1)]}{\lambda_-}} \leq \sqrt{\frac{2[f(x^0) - f(x^*)]}{\lambda_-}}. \tag{6.30}$$

Utilizando la expresión (6.30) y la desigualdad triangular

$$q(x^*, x^1) \leq q(x^*, x^0) + q(x^0, x^1) \leq q(x^*, x^0) + \sqrt{\frac{2[f(x^0) - f(x^*)]}{\lambda_-}} < \rho$$

que nos expresa que $x^1 \in B_q(x^*, \rho)$.
Ahora supongamos que (6.29) es verdadero para valores $1, 2, 3, \ldots, j$ y la expresión (6.28) es verdadero para $k = 1, 2, \ldots, j$, entonces se tiene

$$2\sum_{k=1}^{j} q(x^k, x^{k+1}) \leq \sum_{k=1}^{j} q(x^{k-1}, x^k) + \overline{M}\sum_{k=1}^{j}[\phi(f(x^k) - f(x^*)) - \phi(f(x^{k+1}) - f(x^*))]$$

$$\sum_{k=1}^{j} q(x^k, x^{k+1}) + q(x^j, x^{j+1}) \leq q(x^0, x^1) + \overline{M}[\phi(f(x^1) - f(x^*)) - \phi(f(x^{j+1}) - f(x^*))] \tag{6.31}$$

Utilizamos nuevamente la desigualdad triangular

$$q(x^*, x^{j+1}) \leq q(x^*, x^0) + q(x^0, x^1) + q(x^1, x^2) + \ldots + q(x^j, x^{j+1})$$

$$q(x^*, x^{j+1}) \leq q(x^*, x^0) + q(x^0, x^1) + q(x^0, x^1) + \overline{M}[\phi(f(x^1) - f(x^*)) - \phi(f(x^{j+1}) - f(x^*))]$$

De (6.30) se tiene

$$q(x^*, x^{j+1}) \leq q(x^*, x^0) + 2\sqrt{\frac{2[f(x^0) - f(x^*)]}{\lambda_-}} + \overline{M}[\phi(f(x^1) - f(x^*)) - \phi(f(x^{j+1}) - f(x^*))]$$

sabemos que $\phi(f(x^1) - f(x^*)) - \phi(f(x^{j+1}) - f(x^*)) \leq \phi(f(x^1) - f(x^*)) \leq \phi(f(x^0) - f(x^*))$ luego se tiene

$$q(x^*, x^{j+1}) \leq q(x^*, x^0) + 2\sqrt{\frac{2[f(x^0) - f(x^*)]}{\lambda_-}} + \overline{M}[\phi(f(x^0) - f(x^*))] < \rho$$

lo que implica que $x^{j+1} \in B_q(x^*, \rho)$.
Probaremos ahora que la sucesión $\{x_k\}$ es convergente.
De (6.31) cuando j tiende al infinito se obtiene

$$\sum_{k=1}^{+\infty} q(x^k, x^{k+1}) \leq q(x^0, x^1) + \overline{M}[\phi(f(x^1) - f(x^*))]$$

que nos afirma que la serie $\sum_{k=1}^{+\infty} q(x^k, x^{k+1}) < +\infty$ es convergente y por H_3 se tiene $\sum_{k=1}^{+\infty} ||x^k - x^{k+1}|| < +\infty$. Esto implica que la suceción es de Cauchy, por lo tanto $x^k \to \bar{x} \in B(x^*, \rho/\alpha)$. Por (6.24) inferimos que $\varphi^k \to 0$ y por (6.22) (ϕ' esta decreciente) $f(x^k) \to b \geq f(x^*)$. Si en el caso que $b > f(x^*)$ por la Definición 5.3.9 se tendría

$$\phi'(b - f(x^*))||\varphi^k|| \geq 1, \quad k = 1, 2, \ldots$$

lo cual es absurdo por que $\varphi^k \to 0$, por lo tanto $f(x^*) = b$ y por la semicontinuidad de f se cumple $f(\bar{x}) \leq f(x^*)$.
Si tenemos el caso que f es continua se cumple $f(\bar{x}) = f(x^*)$.

□

Bajo las hipótesis (5.19), (5.20) y la afirmaciones dada por Moreno, nos conduce a una condición de error relativo ($||\varphi^k|| \leq aq(x^{k+1}, x^k)$ con a una constante).
La hipótesis (5.17) considerando $\lambda_- \leq \lambda_{k+1}$ nos da como resultado

$$f(k^{k+1}) \leq f(x^k) - \frac{\lambda_-}{2} q^2(x^{k+1}, x^k) \tag{6.32}$$

que viene a ser una condición suficiente de decrecimiento. Aquí vamos considerar una versión análoga a la de Attouch para la demostración de convergencia a un mínimo local.
Para cualquier $\delta > 0$ existe $0 < \rho < \delta$ y ν tal que

$$H_4: \quad \left.\begin{array}{c} x \in B_q(x^*, \rho), f(x) < f(x^*) + \nu \\ y \notin B_q(x^*, \delta) \end{array}\right\} \Rightarrow f(x) < f(y) + \frac{\lambda_-}{2} q^2(x, y)$$

Lema 6.0.3. *(Convergencia a un mínimo local) Dado la función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ propia, semicontinua inferior que tiene la propiedad de KL en algún mínimo local x^*. Asuma que se cumple la propiedad H_4.*
Entonces para cualquier $r > 0$, existe $u \in (0, r)$ y $\mu > 0$ tal que la desigualdad

$$q(x^*, x^0) < u, \quad f(x^*) < f(x^0) < f(x^+) + \mu.$$

Implica que cualquier sucesión $\{x^k\}$ que empezando en x^0, que satisface (5.17)-(5.20) tiene la propiedad de permanecer en $B_q(x^*, r)$ y converge a algún $\bar{x}$ un punto critico de f con $f(\bar{x}) = f(x^*)$.

Demostración. Dado $r > 0$. Como f satisface la propiedad de Kurdyka Lojasiewicz en x^* entonces existe un $\eta > 0$, $\delta < r$ y una función concava ϕ tal que

- $\forall x \in B_q(x^*, \delta) \cap [f(x^*) < f < f(x^*) + \eta]$ tenemos

$$\phi'(f(x) - f(x^*)) dist(0, \partial f(x)) \geq 1$$

- $\forall x B_q(x^*, \delta)$ se cumple

$$f(x^*) \leq f(x)$$

Por H_4 se tiene que para δ existe un $\rho < \delta$ y $\nu > 0$ tal que

$$\left.\begin{array}{c} x \in B_q(x^*, \rho), f(x) < f(x^*) + \nu \\ y \notin B_q(x^*, \delta) \end{array}\right\} \Rightarrow f(x) < f(y) + \frac{\lambda_-}{2} q^2(x, y)$$

Sea $\eta = \min\{\eta_0, \nu\}$ y sea $k \in \mathbb{N}$. Si x^k es tal que $q(x^*, x^k) < \rho$, $f(x^k) < f(x^*) + \eta$ y de (6.32), entonces por H_4, implica que $x^{k+1} \in B_q(x^*, \delta)$ y $f(x^*) < f(x^{k+1})$ (x es un mínimo local sobre $B_q(x^*, \delta)$). Esto es precisamente la propiedad (6.22) del Lema 6.0.2.
Elegimos $u, \mu > 0$ tal que

$$u < \rho/3, \quad \mu < \eta, \quad 2\sqrt{2\frac{\mu}{\lambda_-}} + \overline{M}[\phi(\mu)] < \rho/3$$

Si x^0 es un punto que satisface $q(x^*, x^0) < u$ y $f(x^*) < f(x^0) < f(x^*) + \mu$ por lo tanto tenemos

$$q(x^*, x^0) + 2\sqrt{2[\frac{f(x^0) - f(x^*)}{\lambda_-}]} + \overline{M}\phi(f(x^0) - f(x^+)) < \rho.$$

Que es precisamente la propiedad (6.21) del Lema 6.0.2. Usando el Lema 6.0.2 concluimos que la sucesión $\{x^k\}$ es convergente y permanece en $B_q(x^*, \rho)$. □

Teorema 6.0.4. *(Convergencia para un mínimo global). Dado la función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ propia, semicontinua inferior que tiene la propiedad de KL en algún x^*, un punto mínimo global de f.*
Para cada $r > 0$, existe $u \in (0, r), \mu > 0$ tal que la desigualdad

$$q(x^*, x^0) < u, \quad f(x^*) < f(x^0) < f(x^*) + \mu$$

implica que cualquier sucesión $\{x^k\}_{k\in\mathbb{N}}$ que satisface (5.17) - (5.20) y que empieza en x^0 satisface

(i) Todos los elementos de la sucesión $\{x^k\}$ pertenecen a $B_q(x^*, r)$

(ii) La sucesión $\{x^k\}$ converge a $\bar{x}$ y $\sum_{k=1}^{+\infty} q(x^k, x^{k+1}) < +\infty$.

(iii) Se cumple $f(\bar{x}) = f(x^*)$.

Demostración. Sea $r > 0$. Ya que x^* es un mínimo global (punto critico) y f satisface la propiedad de Kurdyka Lojasiewicz, existe $\eta > 0$, $\delta \in (0, r)$ y la función concava $\phi : [0, \eta) \to \mathbb{R}$ tal que

- para todo $x \in B_q(x^*, \delta) \subset U \cap [f(x^*) < f < f(x^*) + \eta]$ se cumple la desigualdad

$$\phi'\left(f(x) - f(x^*)\right) dist(\partial f(x)) \geq 1$$

- para todo $x \in B_q(x^*, \delta)$ se cumple

$$f(x) \geq f(x^*)$$

Sea $\eta < \frac{\lambda_-}{2}(\delta - \rho)^2$ y $f(x^*) \leq f(x^k)$ para todo $k \in \mathbb{N}$.(Por la suposición que x^* es un punto mínimo global).
Dado $x^k \in B_q(x^*, \rho)$, por (5.17) se tiene

$$q(x^{k+1}, x^k) \leq \sqrt{\frac{2[f(x^k) - f(x^{k+1})]}{\lambda_{k+1}}} \leq \sqrt{\frac{2[f(x^k) - f(x^{k+1})]}{\lambda_-}} \leq \sqrt{\frac{2\eta}{\lambda_-}} < \delta - \rho$$

la tercera desigualdad de la derecha se debe a $\lambda_- \leq \lambda_{k+1}$.
Por otro lado utilizando la desigualdad triangular

$$q(x^{k+1}, x^*) \leq q(x^{k+1}, x^k) + q(x^k, x^*) < \delta$$

Este es precisamente la propiedad (6.22) del Lema 6.0.2.
Elegimos $u, \mu > 0$ tal que

$$u < \rho/3, \quad \mu < \eta, \quad 2\sqrt{\frac{2\mu}{\lambda_-}} + \overline{M}\varphi(\mu) < \frac{2\rho}{3}$$

Si x^0 satisface el conjunto de desigualdades $q(x^*, x^0) < u$ y $f(x^*) < f(x^0) < f(x^*) + \mu$ tenemos

$$q(x^*, x^0) + 2\sqrt{\frac{\lambda_-(f(x^0) - f(x^*))}{2}} + \overline{M}\phi(f(x^0) - f(x^*)) < \rho$$

que es precisamente la propiedad (6.21) del Lema 6.0.2. Usando el Lema 6.0.2 concluimos que la sucesión $\{x^k\}_{k\in\mathbb{N}}$ permanece en $B_q(x^*, \rho)$, converge a algún $\bar{x} \in B_q(x^*, \delta)$. Ya que $f(x^*)$ es el mínimo de f en $B_q(x^*, \delta)$ se cumple $f(\bar{x}) = f(x^*)$. □

6.0.1. Aplicaciones conductuales

Utilizando el ”enfoque de Racionalidad Variacional "(VR) de dinámica de permanencia y cambios que valen la pena propuesta por [Soubeyran (2009) y Soubeyran (2010)], un algoritmo proximal puede modelar bastante bien un proceso de habituación como se describe en Psicología para un agente o un proceso de rutinización, en Ciencias de la Gestión para una organización.

La Teoría de permanencia y cambio considera sucesiones de permanencias y cambios. Permanencias, se refiere a los hábitos, rutinas, equilibrio, trampas, reglas y convenciones, etc. Cambios, representa creación, destrucción, proceso de aprendizaje, innovación, actitudes así como la formación y revisión de creencias, problemas de autoregulación.

Estas son características en el contexto de Ciencias del Comportamiento, el enfoque de Racionalidad Variacional, muestra como modelizar el caso de las actividades humanas como una sucesión de estancias y cambios temporales que se equilibran, en cada paso, la motivación al cambio (la utilidad de las ventajas al cambio) y la resistencia al cambio (la desutilidad de los inconvenientes al cambio).

En nuestro trabajo se tiene que la resistencia al cambio es "suficientemente débil", donde el término de perturbación modela la dificultad (resistencia relativa) para poder cambiar, que es representada por una cuasi-distancia entre dos iteraciones sucesivas. Una cuasi-distancia modeliza los costos de poder cambiar como un índice de disimilitud entre acciones donde no es lo mismo el costo de poder cambiar de una acción a otra que el costo de poder cambiar de otra manera.

El reciente enfoque de ”Racionalidad Variacional"(VR), unifica muchas de las dinámicas de permanencia y cambios que valen la pena en las Ciencias del comportamiento (Bento and Soubeyran, 2015b) han demostrado que los algoritmos proximales inexactos generalizado usan en cada etapa una condición suficiente de descenso y una regla de parada marginal, representa un prototipo específico pero principal para el enfoque VR que permite modelar entre otras muchas aplicaciones potenciales la formación y ruptura de hábitos y rutinas.

Siguiendo esa misma linea dada por Bento and Soubeyran (2015b), pero en este caso considerando un caso particular y otro criterio de error. El caso del empresario innovador que puede optar por contratar y despedir diferentes tipos y números de trabajadores calificados y especializados: $J = [1, 2, 3, \ldots, l]$ para producir una cantidad elegida de un bien final. El problema es saber cuando cuando este emprendedor dejara de innovar. El perfil elegido de trabajadores calificados es dado por $x = (x^1, x^2, x^3, \ldots, x^l) \geq 0$ donde $x^j \geq 0$ es un número de trabajadores de tipo j. Para simplificar, en cada periodo, cada trabajador calificado utiliza una unidad de un medio no duradero especifico para producir, utilizando su conocimiento para producir una unidad de un componente especifico de tipo j. Luego, el empresario, combina estos diferentes componentes para producir $\Pi(x)$ unidades de un bien final de calidad endogena $s(x)$.

Esta función de producción es original por que mezcla tanto cantidad variable como calidad. El ingreso del empresario es $\psi(\Pi(x), s(x))$. Sus costos operativos $\rho(x)$ son las sumas de sus costos para comprar los medios no duraderos utilizado por cada trabajador y los salarios pagados a cada trabajador empleado. Entonces, en un periodo dado la ganancia del empresario que emplea el perfil $x \in X = \mathbb{R}^l$ de trabajadores calificados es $g(x) = \psi(\Pi(x), S(x)) - \rho(x) \in \mathbb{R}$. Vamos a considerar que $\bar{g} = sup\{g(x); x \in X\} < +\infty$ es la ganancia máxima que el empresario puede esperar alcanzar y $f(x) = \bar{g} - g(x) \geq 0$ su insatisfacción actual. Sean $x = x^k$ y $y = x^{k+1}$ los perfiles del último periodo y del periodo actual de los trabajadores calificados elegidos por el empresario. Las ventajas al cambio son $A(x, y) = g(y) - g(x) = f(x) - f(y) \geq 0$ y los inconvenientes de cambiar representa la diferencia $I(x, y) = C(x, y) - C(x, x) \geq 0$ donde $C(x, y)$ es el costo de cambiar del perfil x al perfil y, $C(x, x)$ es el costo de poderse quedar en el mismo perfil x. Vea Bento and Soubeyran (2015a) para la modelización precisa de los costos $C(x, y)$ para contratar y despedir trabajadores como una cuasi distancia $C(x, y) = q(x, y)$ entre perfiles de trabajadores. Las funciones de motivación y resistencia al cambio son $M(x, y) = U(A(x, y)) = [g(y) - g(x)]^\mu$ y $R(x, y) = D(I(x, y)) = q(x, y)^\nu$ con $\mu, \nu > 0$ la resistencia relativa al cambio es $q(x, y)^{\nu/\mu} = U^{-1}(D(I(x, y)))$, si $\alpha = \nu/\mu > 0$. En este contexto, vale la pena cambiar del perfil x al perfil y y si $M(x, y) \geq \xi R(x, y)$ es decir $[g(y) - g(x)]^\mu \geq \xi q(x, y)^\nu$ donde $\xi > 0$. Entonces un cambio que vale la pena en x es tal que

$$y \in W_\xi(x) \Leftrightarrow g(y) - g(x) \geq \lambda q(x, y), \quad \lambda = (\xi)^{1/\mu}$$

que define el conjunto que vale la pena cambiar en x. La sucesión de estancia y cambios temporales que valen la pena satisfacen

$$g(x^{k+1}) - g(x^k) = f(x^k) - f(x^{k+1}) \geq \lambda_{k+1} q^\alpha(x^k, x^{k+1}), \quad k \in \mathbb{N}$$

Este ejemplo nos muestra como se forma la rutina progresivamente, paso por paso, en una manera recurrente.

Capítulo 7

METODOLOGÍA

7.1. Tipo y Diseño de Investigación

El presente estudio es de tipo descriptivo y correlacional, pues en primer lugar se describe las propiedades de convergencia de los algoritmos del método de punto proximal inexacto con cuasi distancia para la minimización de funciones convexas y no convexas; y en segunda lugar se pretende relacionar dichos algoritmos con sus versiones inexactas.

7.2. Unidad de análisis

Para la realización de este trabajo, se estudio el artículo ”A proximal algorithm with quasi distance. Application to habit’s formation"propuesto por Moreno et al. (2012), el artículo ” Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss-Seidel methods” propuesto por Attouch et al. (2013) y el artículo ”A Generalized Inexact Proximal Point Method for Nonsmooth Functions thar Satisfies Kurdyka Lojasiewicz Inequality”. presentada por Bento and Soubeyran (2015b).

7.3. Población de estudio

En este trabajo no se tiene población de estudio pues es netamente abstracto, sin embargo, el estudio a desarrollar esta intimamente relacionado con el espacio Euclideano $\mathbb{R}^n$, los métodos de punto proximal y los problemas de optimización de funciones convexas y no convexas.

7.4. Técnica para la recolección de datos

La debida recolección de información a fin de avanzar en el presente tesis se hizo mediante la búsqueda en página web, así como también la revisión bibliográfica (artículos, libros, tesis, etc.) y el apoyo de mi asesor para guiarme en la correcta bibliografía.

7.5. Análisis e interpretación de la información

En el desarrollo de la tesis se empezo revisando y estudiando algunos textos relacionados con el tema de análisis variacional y análisis convexo, que nos ayudo para ver los distintos tipos de subdiferenciales tales como, la subdiferencial de Frechet, la subdiferencial en el limite y así como también la cuasi-distancia. De igual modo se busco en los libros y artículos métodos de optimización y algoritmos proximal con cuasi-distancia. Los resultados fueron presentadas periódicamente al asesor quien nos oriento y guió en la estructura de la tesis.

Capítulo 8

RESULTADO Y DISCUSIÓN

Presentación de resultados

El trabajo realizado aquí es de tipo teórico pero incentivado a la aplicación. Los avances obtenidos sobre la convergencia de la sucesión que es generada por el algoritmo del MPPI con cuasi-distancia fueron presentados en el capítulo 5 donde se tiene un criterio de error vectorial distinta a la presentada por Attouch la cual mencionamos en los siguientes items.

1.- Bajo las hipótesis $(\mathbf{H_1})$, $(\mathbf{H_2})$, $(\mathbf{H_3})$ se prueba para funciones f no convexas que todo punto de acumulación dada por la sucesión generada por API es para el problema (5.14) un punto crítico.

2.- Teniendo las mismas hipótesis del item 1, se prueba para funciones convexas, que cada punto de acumulación generada por el API es una solución optima del problema (5.14), es decir, cada punto de acumulación es un mínimo global de la función f.

3.- Siendo $\bar{x}$ un mínimo global de f y además es un punto que satisface la propiedad de Kurdyka Lojasiewicz para dicha función, se prueba que la sucesión que inicia cerca a una vecindad de $\bar{x}$ converge a dicho punto.

4.- El error e^{k+1} presentada por API satisface

$$||e^{k+1}|| \leq \sigma q(x^k, x^{k+1})$$

que es un error vectorial derivado de la subdiferencial de la función $f_k(x) = f(x) + \frac{\lambda_{k+1}}{2} q^2(x^k, x)$ dado por

$$e^{k+1} = \varphi^{k+1} + \lambda_{k+1} q(x^k, x^{k+1}) \psi^{k+1}. \qquad \varphi^{k+1} \in \partial f(x^{k+1}),\ \psi^{k+1} \in \partial q(x^k, .)(x^{k+1}).$$

es una condición que ayuda a la convergencia del algoritmo.

5.- En este trabajo se presenta la rapidez de convergencia para el algoritmo API.

Análisis, interpretación y discusión de resultados

1.- En esta tesis se presenta la extensión del método del punto proximal para la resolución de problemas de optimización no convexas, las cuales según estudio recientes, modelan de una manera más aproximada ciertos problemas reales.

2.- Algunas situaciones son modelados en funciones de tipo no convexas que en general son difíciles de resolver usando técnicas de optimización conocidas, es por ello que el aporte dado en esta tesis puede ser considerada como una contribución para la obtención de algoritmos de aproximación más eficientes.

Capítulo 9

CONCLUSIONES Y FUTURAS INVESTIGACIONES

9.1. Conclusiones

1.- Se presenta el API con cuasi-distancia con el fin de resolver un problema de minimización, que en cada proceso resuelve un subproblema cuya solución encontrada es aproximada, en este caso se presenta su variante inexacta con distinto criterio de error.

2.- Aquí exponemos el API con cuasi distancia para minimización de funciones objetivos reales convexas y no convexas definidas en espacio Euclideanos, se sigue el mismo enfoque seguida por Bento and Soubeyran (2015b).

3.- Bajo hipótesis razonables se prueba que $\{x^k\}$ es una sucesión convergente a un mínimo local del problema de optimización 5.14.

4.- Se tiene que los métodos de aproximación inexactos, se ajustan mejor a un modelo matemático siendo más eficaz que los métodos de aproximación exacta, debido a que generan menos trabajo computacional y mejor tasa de convergencia.

5.- Esperamos que este trabajo de investigación ayude a futuras tesis, pues el campo de investigación es muy amplio, tanto en la parte teorica como en la computacional, esto debido a su gran utilidad en distintas áreas de la matemática.

6.- El algoritmo de punto proximal inexacto en cada etapa del proceso resuelve un problema perturbado simple, es decir, se va generando un conjunto de puntos, que convergen a la solución de un problema.

9.2. Futuras investigaciones

A la realización de esta tesis hubo algunos interrogantes que no se han resuelto por completo, esperamos que los futuros trabajos de investigación resuelvan dichos dudas:

1.- Un trabajo de investigación a futuro se puede dar en la implementación computacional del método estudiado para problemas adaptados a la realidad, y su respectivas comparaciones con otros métodos.

2.- Un futuro trabajo puede ser para un algoritmo proximal inexacto con cuasi distancia con criterio de error tanto vectorial como escalar.

Bibliografía

Attouch, H. and Bolte, J. (2009). On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. *Mathematical Programming*, 116(1–2):5–16.

Attouch, H., Bolte, J., and Svaiter, B. F. (2013). Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods. *Mathematical Programming*, 137(1-2):91–129.

Barbolla, R., Tena, E. C., and Sanz, P. (2001). *Optimización: cuestiones, ejercicios y aplicaciones a la economía*. Pearson Educación. España.

Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. (2013). *Nonlinear programming: theory and algorithms*. John Wiley & Sons, New Jersey.

Benazic, R. (2000). *Topología en espacios Euclideanos, V. 2*. Sociedad Matemática Peruana. Lima.

Bento, G. and Soubeyran, A. (2015a). Generalized inexact proximal algorithms: routineś formation with resistance to change, following worthwhile changes. *Journal of Optimization Theory and Applications*, 166(1):172–187.

Bento, G. and Soubeyran, A. (2015b). A generalized inexact proximal point method for nonsmooth functions that satisfies kurdyka lojasiewicz inequality. *Set-Valued and Variational Analysis*, 23(3):501–517.

Bento, G. C. and Neto, JX Cruz y Soubeyran, A. (2014). A proximal point type method for multicriterio optimization. *Set-Valued and Variational Analysis*, 22:557–573.

Brattka, V. (2003). Recursive quasi-metric spaces. *Theoretical computer science*, 305(1-3):17–42.

Chen, J. S. and Pan, S. (2008). A proximal-like algorithm for a class of nonconvex programming. *Pacific Journal of Optimization*, 4(2):319–333.

dos Santos Gromicho, J. A. (2013). *Quasiconvex optimization and location theory, V. 9*. Springer Science & Business Media. Netherlands.

Gerla, G. et al. (2006). Quasi-metric spaces and point-free geometry. *Mathematical Structures in Computer Science*, 16(1):115–137.

Güler, O. (1991). On the convergence of the proximal point algorithm for convex minimization. *SIAM Journal on Control and Optimization*, 29(2):403–419.

Hasser, N. B., LaSalle Joseph, P., and Sullivan Joseph, A. (1970). *Análisis Matemático 2; Curso Intermedio*. Trillas. México.

Hausdorff, F. (2021). *Set theory, V. 119*. American Mathematical Soc. New York.

Illesca Cangalaya, O. F. (2016). *Solución de un problema de operadores monótonos maximales usando el algoritmo de punto proximal. Tesis de Licenciatura.* Universidad Nacional del Callao.

Iusem, A. N. (1995). Métodos de ponto proximal em otimizacao. In *20 colóquio Brasileiro de Matemática*. Instituto de matemática pura e aplicada.

Kaplan, A. and Tichatschke, R. (1998). Proximal point method and nonconvex optimization. *Journal of global Optimization*, 13(4):389–406.

Kiwiel, K. C. (1997). Proximal minimization methods with generalized bregman functions. *SIAM journal on control and optimization*, 35(4):1142–1168.

Künzi, H.-P. A. (2001). Nonsymmetric distances and their associated topologies: about the origins of basic ideas in the area of asymmetric topology. In *Handbook of the history of general topology*, pages 853–968. Springer.

Künzi, H.-P. A., Pajoohesh, H., and Schellekens, M. P. (2006). Partial quasi-metrics. *Theoretical Computer Science*, 365(3):237–246.

Lima, E. L. (2000). *curso de análise vol. 2, Projeto Euclides, V. 2*. Livros técnicos e científicos editora.

Makela, M. M. and Neittaanmaki, P. (1992). *Nonsmooth optimization: analysis and algorithms with applications to optimal control*. World Scientific. Singapore.

Martinet, B. (1970). Breve communication. régularisation d'inéquations variationnelles par approximations successives. *Revue francaise*, 4:154–158.

Mas Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic theory*, volume 1. Oxford university press New York.

Moreno, F. G., Oliveira, P. R., and Soubeyran, A. (2012). A proximal algorithm with quasi distance application to habit's formation. *Optimization*, 61(12):1383–1403.

Pan, S. and Chen, J.-S. (2007). Entropy like proximal algorithms based on a second-order homogeneous distance function for cuasi-convex programming. *Journal of Global Optimization*, 39(4):555–575.

Papa Quiroz, E. and Oliveira, P. R. (2012). Proximal point method for minimizing quasiconvex locally lipschitz functions on hadamard manifolds. *Nonlinear Analysis: Theory, Methods Applications*, 75(15):5924–5932.

Papa Quiroz, E., Ramirez, L. M., and Oliveira, P. R. (2015). A inexact proximal method for quasiconvex minimization. *European Journal of operational Research*, 246(3):721–729.

Papa Quiroz, E. A. (2019a). *optimización matemática y computacional, V.1*. Dhyblas-soft.

Papa Quiroz, Erik y Acuña, S. C. (2019b). Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa en espacios euclidianos. *Pesquimat*, 22(1):31–50.

Polyak, B. (1987). Introduction to optimization, ser. *Series of Translations in Mathematics and Engineering. New York: Optimization Software Publications Division Inc*.

Rocha, Rog 'e rio Azevedo y Oliveira, P. R. y. G. ó. r. R. M. y. S. M. (2016). Método de escalarización de punto proximal de cuasi distancia logaritmica para programación multiobjetivo. *Matemáticas aplicadas y computación*, 273:856–867.

Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. *SIAM journal on control and optimization*, 14(5):877–898.

Rockafellar, R. T. and Wets, R. J.-B. (2009). *Variational analysis*, volume 317. Springer Science & Business Media.

Romaguera, S. and Sanchis, M. (2003). Aplicaciones de funciones de utilidad definidas en espacios cuasi-métricos. *Diario de análisis y aplicaciones matemáticos*, 283(1):219–235.

Soubeyran, A. (2009). Variational rationality, a theory of individual stability and change: worthwhile and ambidextry behaviors. greqam. aix-marseillle university. *Preprint*.

Soubeyran, A. (2010). Variational rationality and the unsatisfied man: routines and the course pursuit between aspirations, capabilities and beliefs. *Preprint. GREQAM, Aix Marseillle University*.

Souza, S. d. S., Oliveira, P. R., da Cruz Neto, J. X., and Soubeyran, A. (2010). A proximal method with separable bregman distances for quasiconvex minimization over the nonnegative orthant. *European Journal of Operational Research*, 201(2):365–376.

Stojmirovic, A. (2003). Quasi-metric spaces with measure. *arXiv preprint math/0311070*.

Takayama, A. and Akira, T. (1985). *Mathematical economics*. Cambridge University Press.