Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Matemáticas
Escuela Profesional de Matemática

Convergencia del algoritmo de punto proximal en espacios de Hilbert: Implementación y simulaciones en Python

TESIS

Para optar el Título Profesional de Licenciado en Matemática

AUTOR

Miguel Francisco TÚLLUME LLUÉN

ASESOR

Dr. Alex Armando CRUZ HUALLPARA

Lima, Perú

2024



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

Túllume, M. (2024). *Convergencia del algoritmo de punto proximal en espacios de Hilbert: Implementación y simulaciones en Python*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Matemática]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Miguel Francisco Túllume Lluén
Tipo de documento de identidad	DNI
Número de documento de identidad	765828258
URL de ORCID	https://orcid.org/0009-0003-9318-424X
Datos de asesor	
Nombres y apellidos	Alex Armando Cruz Huallpara
Tipo de documento de identidad	DNI
Número de documento de identidad	40949706
URL de ORCID	https://orcid.org/0000-0002-5453-6885
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Víctor Hilario Tarazona Miranda
Tipo de documento	DNI
Número de documento de identidad	09264893
Miembro del jurado 1	
Nombres y apellidos	Luis Guillermo Huamanlazo Ricci
Tipo de documento	DNI
Número de documento de identidad	09197486
Datos de investigación	

Línea de investigación	Optimización
Grupo de investigación	Matemática Aplicada - Métodos Numéricos (METNUM)
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	Lima, Perú.
Año o rango de años en que se realizó la investigación	2024
URL de disciplinas OCDE	MATEMATICA PURA https://purl.org/pe-repo/ocde/ford#1.01.00

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
FACULTAD DE CIENCIAS MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMATICA

ACTA DE SUSTENTACIÓN DE TESIS PARA LA OBTENCIÓN DEL TÍTULO PROFESIONAL DE LICENCIADO EN MATEMATICA (PROGRAMA DE TITULACIÓN PROFESIONAL 2024)

En la UNMSM – Ciudad Universitaria – Facultad de Ciencias Matemáticas, siendo las 12:20 horas del sábado 28 de septiembre del 2024, se reunieron los docentes designados como Miembros del Jurado Evaluador (PROGRAMA DE TITULACIÓN PROFESIONAL 2024): Dr. Víctor Hilario Tarazona Miranda (PRESIDENTE), Mg. Luis Guillermo Huamanlazo Ricci (MIEMBRO) y el Dr. Alex Armando Cruz Huallpara (MIEMBRO ASESOR), para la sustentación de la Tesis titulada: **"CONVERGENCIA DEL ALGORITMO DE PUNTO PROXIMAL EN ESPACIOS DE HILBERT: IMPLEMENTACIÓN Y SIMULACIONES EN PYTHON"**, presentado por el señor **Bachiller MIGUEL FRANCISCO TÚLLUME LLUÉN**, para optar el Título Profesional de Licenciado en Matemática.

Luego de la exposición de la Tesis, el Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los Miembros del Jurado Evaluador, el expositor mereció la aprobación de sobresaliente, con un calificativo promedio de dieciocho (18).

A continuación, los Miembros del Jurado Evaluador dan manifiesto que el participante **Bachiller MIGUEL FRANCISCO TÚLLUME LLUÉN**, en vista de haber aprobado la sustentación de su Tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Matemática.

Siendo las 13:00 horas se levantó la sesión firmando para constancia la presente Acta.

Dr. Víctor Hilario Tarazona Miranda
PRESIDENTE

Mg. Luis Guillermo Huamanlazo Ricci
MIEMBRO

Dr. Alex Armando Cruz Huallpara
MIEMBRO ASESOR

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

Vicerrectorado de Investigación y Posgrado

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

CERTIFICADO DE SIMILITUD

Yo Dr. Alex Armando Cruz Huallpara en mi condición de asesor acreditado con la Resolución Decanal N°001330-2024-D-FCM/UNMSM de la tesis/monografía/informe de investigación/trabajo académico, cuyo título es **"CONVERGENCIA DEL ALGORITMO DE PUNTO PROXIMAL EN ESPACIOS DE HILBERT: IMPLEMENTACIÓN Y SIMULACIONES EN PYTHON",** presentado por el bachiller **Miguel Francisco Túllume Lluén**, para optar el título Profesional de Licenciado en Matemática CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante el software de similitud textual, el documentoevaluado cuenta con el porcentaje de 5% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado/ título/ especialidad correspondiente.

Firma del Asesor

DNI:40949706

FICHA CATALOGRÁFICA

Miguel Francisco Túllume Lluén

Convergencia del Algoritmo de Punto Proximal en Espacios de Hilbert: Implementación y Simulaciones en Python, (Lima) 2024

VIII.,68p.,29.7cm (UNMSM, Titulo, Matemática, 2020) Tesis, Universidad Nacional Mayor de San Marcos. Facultad de Ciencias Matemáticas.

1. Matemática, UNMSM/FCM II. Título

(Series).

DEDICATORIA

Este trabajo está dedicado con todo mi amor y gratitud a:

Mi madre, Jesús Eusebia Lluén Túllume,
quien con su amor, sacrificio y apoyo incondicional han sido mi mayor inspiración y fuerza para alcanzar mis metas.

Mi padre, Manuel Túllume Custodio,
cuya sabiduría, paciencia y guía han sido fundamentales en cada paso de mi camino.

Mis hermanos, Maira y Víctor Manuel,
quienes siempre han estado a mi lado, brindándome su apoyo y cariño incondicional.

Y a **Pelusa**, por la compañía silenciosa durante las largas noches de estudio. Gracias por todo el cariño que me diste.

AGRADECIMIENTOS

"En primer lugar, agradezco a **Dios**, por ser mi guía y fortaleza en cada paso de este camino académico".

"A mis padres, **Jesús Eusebia Lluén Túllume** y **Manuel Túllume Custodio**, por su amor incondicional, su apoyo constante y por creer siempre en mí. Su ejemplo y sacrificio me han inspirado a perseverar y alcanzar mis metas".

A mis hermanos, **Maira Cruz** y **Víctor Manuel**, por su cariño, compañía y apoyo incondicional. Gracias por estar siempre a mi lado, alentándome en cada momento.

A mi asesor, el **Dr. Alex Cruz Huallpara**, por su guía experta, paciencia y dedicación durante todo este proceso. Sus enseñanzas y consejos han sido invaluables.

A mi enamorada, **Mary Campos**, por su amor, comprensión y apoyo constante en los momentos más desafiantes. Gracias por ser mi compañera y mi soporte emocional en todo momento.

A las secretarias de la escuela, **Juanita** y **Mirtha**, por su amabilidad, eficiencia y disposición para ayudar en todo lo que necesitaba. Su apoyo ha sido fundamental.

A mis amigos de la facultad, por su amistad, compañerismo y por compartir juntos esta etapa de nuestras vidas. Sus palabras de ánimo y apoyo han sido una fuente constante de motivación.

RESUMEN

Convergencia del Algoritmo de Punto Proximal en Espacios de Hilbert: Implementación y Simulaciones en Python

Miguel Francisco Túllume Lluén

2024

Asesor : Dr. Alex Armando, Cruz Huallpara

Titulo Obtenido : Licenciado en Matemática

En este trabajo de tesis, se estudia la convergencia del algoritmo de punto proximal en espacios de Hilbert utilizando las definiciones de subdiferenciales y distancias proyectadas sobre un subconjunto no vacío, convexo y cerrado para funciones convexas.

Para analizar la convergencia en funciones convexas, se utiliza el subdiferencial de Frchet, "definido como

$$\widehat{\partial}F(x) = \left\{ v \in \mathbb{R}^n \;\middle|\; \liminf_{\substack{y \to x \\ y \neq x}} \frac{F(y) - F(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

En el caso de funciones diferenciables, el subdiferencial de Fréchet coincide con la derivada usual (o el gradiente en el caso de varias variables)".

Además, se aplica la técnica de regularización de Moreau-Yosida, esta regularización transforma una función original en una nueva función suavizada (más manejable) manteniendo muchas de las propiedades de la función original. Gracias a esta regularización podemos "definir la Envolvente de Moreau-Yosida como

$$F_\lambda(x) = \min_{z \in \mathcal{H}} \{F(\lambda, x)(z)\} = \min_{z \in \mathcal{H}} \left\{ F(z) + \frac{1}{2\lambda}\|z - x\|^2 \right\} = F(x^*) + \frac{1}{2\lambda}\|x^* - x\|^2$$

Es especialmente útil en problemas donde la función objetivo tiene discontinuidades o es difícil de minimizar directamente debido a su no diferenciabilidad".

En el estudio de la convergencia de métodos de optimización, un resultado clave es que, bajo ciertas condiciones, las secuencias generadas por el algoritmo convergen hacia un minimizador de la función objetivo. Este tipo de resultados es particularmente relevante cuando se trabaja con espacios donde el conjunto de minimizadores X^* no es vacío.

Específicamente, si el conjunto de minimizadores X^* no está vacío, la sucesión $\{x_n\}$ generada por el algoritmo converge débilmente a un minimizador de la función F. Además, se puede garantizar una cota superior para la diferencia entre el valor de la función en x_{n+1} y el valor óptimo F^*, "la cual está dada por la siguiente expresión

$$F(x_{n+1}) - F^* \leq \frac{d(x_0, X^*)^2}{2\sigma_n}$$

donde $d(x_0, X^*)$ denota la distancia entre el punto inicial x_0 y el conjunto de minimizadores X^*, mientras que σ_n es un parámetro dependiente de la iteración n". Esta desigualdad proporciona una medida de la velocidad de convergencia de la secuencia hacia el valor óptimo F.

"El objetivo de este trabajo es demostrar la efectividad y la convergencia del algoritmo de punto proximal en estos espacios, validando los resultados teóricos mediante una implementación detallada en Python". Además, se presentan diversas aplicaciones prácticas en Python del algoritmo en problemas de optimización.

Palabras clave: Convergencia del algoritmo de punto proximal, Espacios de Hilbert, Subdiferencial de Fréchet, Envolvente de Moreau-Yosida, Funciones convexas, Optimización convexa, Minimización de funciones, Python en optimización, Funciones no diferenciables.

ABSTRACT

Convergence of the Proximal Point Algorithm in Hilbert Spaces: Implementation and Simulations in Python

Miguel Francisco Túllume Lluén

2024

Advisor : Dr. Alex Armando, Cruz Huallpara

Degree Obtained : Bachelor in Mathematics

In this thesis work, the convergence of the proximal point algorithm in Hilbert spaces is studied using the definitions of subdifferentials and distances projected onto a nonempty, convex and closed subset for convex functions.

To analyze convergence in convex functions, the Frechet subdifferential is used, "defined as:

$$\widehat{\partial}F(x) = \left\{ v \in \mathbb{R}^n \;\middle|\; \liminf_{\substack{y \to x \\ y \neq x}} \frac{F(y) - F(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

In the case of differentiable functions, the Fréchet subdifferential coincides with the usual derivative (or the gradient in the case of several variables)".

In addition, the Moreau-Yosida regularization technique is applied, this regularization transforms an original function into a new (more tractable) smoothed function while maintaining many of the properties of the original function. Thanks to this regularization we can "define the Moreau-Yosida Envelope as

$$F_\lambda(x) = \min_{z \in \mathcal{H}} \{F(\lambda, x)(z)\} = \min_{z \in \mathcal{H}} \left\{ F(z) + \frac{1}{2\lambda}\|z - x\|^2 \right\} = F(x^*) + \frac{1}{2\lambda}\|x^* - x\|^2$$

It is especially useful in problems where the objective function has discontinuities or is difficult to minimize directly due to its non-differentiability".

In the study of convergence of optimization methods, a key result is that, under certain conditions, the sequences generated by the algorithm converge to a minimizer of the objective function. This type of result is particularly relevant when working with spaces where the set of minimizers X^* is non-empty.

Specifically, if the set of minimizers X^* is not empty, the sequence $\{x_n\}$ generated by the algorithm converges weakly to a minimizer of the function F. Moreover, an upper bound can be guaranteed for the difference between the function value at x_{n+1} and the optimal value F^*, "which is given by the following expression

$$F(x_{n+1}) - F^* \leq \frac{d(x_0, X^*)^2}{2\sigma_n}$$

where $d(x_0, X^*)$ denotes the distance between the initial point x_0 and the set of minimizers X^*, while σ_n is an iteration n dependent parameter". This inequality provides a measure of the speed of convergence of the sequence to the optimal value F.

"The objective of this paper is to demonstrate the effectiveness and convergence of the proximal point algorithm in these spaces, validating the theoretical results by means of a detailed Python implementation". In addition, several practical Python applications of the algorithm in optimization problems are presented.

Keywords: Convergence of the proximal point algorithm, Hilbert spaces, Fréchet subdifferential, Moreau-Yosida envelope, Convex functions, Convex optimization, Minimization of functions, Python in optimization, Non-differentiable functions.

Índice general

1 Introducción **11**

2 Definiciones Preliminares **13**

2.1 Espacios Vectoriales 14

2.2 Espacios Métricos 17

2.3 Sucesiones 18

2.4 Subsucesiones 20

2.5 Sucesiones de Cauchy 22

2.6 Espacios Normados 24

2.7 Espacios de Banach 26

2.8 Convergencia de Sucesiones es Espacios de Banach 28

2.9 Conjuntos Convexos 29

2.10 Límite Inferior 31

2.11 Semicontinuidad Inferior 33

2.12 Funciones Lipschitzianas 34

2.13 Análisis Convexo 36

2.14 El Teorema de Separación de Hahn-Banach 40

3 Espacios de Hilbert **42**

3.1 Convergencia de Sucesiones en Espacios de Hilbert 43

3.2 Función Convexa 43

3.3 Función Cuasiconvexa 46

3.4 Subdiferencial de Fréchet 52

3.5 Regla Generalización de Fermat para el subdiferencial de Fréchet 53

3.6 Teorema de Moreau–Rockafellar para el subdiferencial de Fréchet 54

3.7 Subdiferencial Limitante 56

3.8 Derivada direccional de Clarke 58

3.9 Subdiferencial de Clarke 61

3.10 Regla Generalización de Fermat para el subdiferencial de Clarke 74

3.11 Conjunto de Funciones Convexas en Espacios de Hilbert 76

4 Proyección en Espacios de Hilbert **79**

4.1 Proyección 79

4.2 Ortogonalidad 86

5 Moreau-Yosida **87**

5.1 Regularización de Moreau-Yosida 88

5.2 Envolvente de Moreau-Yosida 92

6 El Problema **97**

6.1 Formulación del Problema 97

7 Algoritmo **99**

7.1 Pseudocodigo 100

8 Convergencia del Algoritmo **104**

9 Aplicación del algoritmo en Phyton **115**

9.1 Herramientas y Bibliotecas 115

9.2 Ejemplos Prácticos 116

9.2.1 PPA para la Minimización de Funciones Absolutas 116

9.2.2 PPA para la Minimización de Funciones a Tramos 120

9.2.3 PPA para la Minimización de Funciones Logaritmicas 125

9.2.4 PPA para la Minimización de Función Suma de la norma euclidiana y el máximo 130

10 Conclusiones y Recomendaciones **136**

10.1 Conclusiones 136

10.2 Recomendaciones 137

1 Introducción

El presente trabajo de tesis aborda el estudio del algoritmo de punto proximal en el contexto de los espacios de Hilbert, con un énfasis particular en su implementación y simulaciones utilizando Python. El algoritmo de punto proximal es una herramienta fundamental en la teoría de optimización, especialmente en la resolución de problemas que involucran funciones no diferenciables. Este algoritmo ha demostrado ser eficaz en una amplia gama de aplicaciones, desde la optimización convexa hasta problemas de equilibrio y análisis variacional.

La estructura de la tesis está organizada de la siguiente manera: **Capítulo 2: Definiciones Preliminares**, se introducen las definiciones y conceptos básicos necesarios para el desarrollo del trabajo, incluyendo espacios vectoriales, métricos, normados y de Banach. El **Capítulo 3: Espacios de Hilbert**, profundiza en la teoría de los espacios de Hilbert, incluyendo la convergencia de sucesiones, funciones convexas y cuasiconvexas, y subdiferenciales. Mientras que en el **Capítulo 4: Proyección en Espacios de Hilbert**, se estudia la teoría de la proyección en espacios de Hilbert, fundamental para la implementación del algoritmo de punto proximal.

En el **Capítulo 5: Regularización de Moreau-Yosida**, se analiza la técnica de regularización y la envolvente de Moreau-Yosida, que juega un papel crucial en la convergencia del algoritmo. Posteriormente, en el **Capítulo 6: El Problema**, se formula el problema específico de optimización que se abordará utilizando el algoritmo de punto proximal. El **Capítulo 7: Algoritmo**, expone el pseudocódigo del algoritmo de punto proximal y se detalla su funcionamiento, explicando paso a paso cómo opera el algoritmo

El **Capítulo 8: Convergencia del Algoritmo**, ofrece una demostración rigurosa y detallada de la convergencia del algoritmo en espacios de Hilbert. A continuación, el **Capítulo 9: Aplicación del Algoritmo en Python**, describe las herramientas y bibliotecas utilizadas en la implementación del algoritmo en Python, y se presentan ejemplos prácticos de su aplicación. Finalmente en el **Capítulo 10: Conclusiones y Recomendaciones**, se resumen los hallazgos principales de la tesis y se discuten las implicaciones de los resultados obtenidos. Además, se ofrecen recomendaciones para trabajos futuros y posibles extensiones del estudio.

2 Definiciones Preliminares

En este capítulo destaca el valor de la investigación académica y la consulta de fuentes especializadas, como libros y papers científicos, para el desarrollo y avance de teorías, como Rojo (1973), Benazic (2000), Lages Lima (1981), Mas-Colell et al. (1995) y Lázaro Carrión (2002), con quienes se logró un amplio marco teorico en espacios vectoriales, métricos, normados, banach, sucesiones de cauchy, entre otros. Mientras que para teoria de convexidad de conjuntos y funciones, espacios de hilbert, subdiferencial de clarke, etc., se utilizaron las investigaciones de Clarke (1990), The Luc (1989), Martinet (1970), Mordukhovich et al. (2006), Rockafellar (1976) y Cano Lengua (2011).

Definición 2.0.1 *Consideraremos el conjunto de los reales extendidos, agregando el símbolo* $+\infty$*. Extendemos el orden total de los reales de la manera intuitiva:*

$$x < +\infty \text{ para todo } x \in \mathbb{R}$$

Definiré también las siguientes operaciones:

1. $x + (+\infty) = (+\infty) + x = (+\infty) + (+\infty) = +\infty$ *para todo* $x \in \mathbb{R}$.
2. $x \cdot (+\infty) = +\infty$ *para todo* $x > 0$.
3. $0 \cdot (+\infty) = 0$

Definición 2.0.2 (Funciones de los Reales Extendidos) *Son aquellas funciones de la forma* $F : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$*, donde* $\mathbb{X}$ *es un conjunto.*

Ejemplo 2.0.1 (Función Indicatriz) *"La función indicatriz de un subconjunto no vacío C*

de $\mathbb{X}$ *es*

$$\delta_C(x) = \begin{cases} 0 \quad , & x \in C \\ +\infty \ , & x \notin C \end{cases}$$

El conjunto de las funciones de $\mathbb{X}$ *en* $\mathbb{R} \cup \{+\infty\}$*, no es un espacio vectorial, pero si es un cono convexo".*

Definición 2.0.3 *Sea* $\mathbb{X}$ *un conjunto no vacío. Sea* $F : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$*, entonces*

1. *"*$Dom(F) = \{x \in \mathbb{X} \mid F(x) < +\infty\}$*"*
2. *"*$Graf(F) = \{(x,\xi) \in \mathbb{X} \times \mathbb{R} \mid F(x) = \xi\}$*"*
3. *"*$Epi(F) = \{(x,\xi) \in \mathbb{X} \times \mathbb{R} \mid F(x) \leq \xi\}$*"*

Definición 2.0.4 (Función Propia) *Diremos que* F *es propia si su dominio, denotado por* $dom(F) = \{x \in \mathcal{H} \mid F(x) < +\infty\}$*, no es vacío, es decir,* $dom(F) \neq \emptyset$*, que es equivalente a* $epi(F) \neq \emptyset$*.*

2.1 Espacios Vectoriales

Definición 2.1.1 *Sea* $\mathbb{X}$ *un conjunto no vacío, decimos que* $\mathbb{X}$ *es un espacio vectorial sobre un cuerpo* $\mathbb{K}$ *(que puede ser los números reales o complejos) si para cualquier par de vectores* $u,v \in \mathbb{X}$ *y cualquier par de escalares* $a,b \in \mathbb{K}$ *se cumplen las siguientes propiedades:*

1. *Cerradura bajo la suma:* $u + v \in \mathbb{X}$
2. *Cerradura bajo la multiplicación por escalar:* $au \in \mathbb{X}$
3. *Ley conmutativa de la suma:* $u + v = v + u$
4. *Ley asociativa de la suma:* $(u + v) + w = u + (v + w)$

5. *Existe un elemento en* $\mathbb{X}$*, llamado vector cero o elemento neutro de la suma, tal que* $u + 0 = u = 0 + u$ *para todo vector* $u \in \mathbb{X}$

6. *Para cada vector* $u \in \mathbb{X}$*, existe un vector opuesto* $-u \in \mathbb{X}$*, tal que* $u + (-u) = 0$

7. *Ley asociativa de la multiplicación por escalar:* $a(bu) = (ab)u$

8. *Multiplicación por el escalar unidad:* $1u = u$

Ejemplo 2.1.1 *"A continuación presentaré algunos espacios vectoriales que cumplen las propiedades antes mencionadas*

a) ***Espacio vectorial euclidiano:***

Un espacio vectorial euclidiano de dimensión n, denotado como $\mathbb{R}^n$*, es el conjunto de todos los vectores con n componentes reales. En* $\mathbb{R}^3$*, los vectores tienen la forma* (x, y, z)*, donde* $x, y, z \in \mathbb{R}$*. Las operaciones de suma y multiplicación por escalar están definidas de la siguiente manera:*

- ***Suma de vectores:*** *Si* $v = (v_1, v_2, v_3)$ *y* $w = (w_1, w_2, w_3)$*, entonces* $v + w = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$*.*
- ***Multiplicación por un escalar:*** *Si* $\alpha \in \mathbb{R}$ *y* $v = (v_1, v_2, v_3)$*, entonces* $\alpha v = (\alpha v_1, \alpha v_2, \alpha v_3)$*.*

b) ***Espacio vectorial de matrices:***

El conjunto de todas las matrices $n \times m$ *con entradas en un campo F forma un espacio vectorial. Las operaciones de suma y multiplicación por escalar se definen de la siguiente manera:*

- ***Suma de matrices:*** *Si* $A = [a_{ij}]$ *y* $B = [b_{ij}]$ *son matrices del mismo tamaño, entonces la suma* $A + B$ *es otra matriz del mismo tamaño con entradas* $a_{ij} + b_{ij}$*.*

- ***Multiplicación por un escalar:*** *Si* $\alpha \in F$ *y* $A = [a_{ij}]$ *es una matriz, entonces* αA *es una matriz con entradas* αa_{ij}.

c) ***Espacio vectorial de polinomios:***

El conjunto de todos los polinomios con coeficientes en un campo F *forma un espacio vectorial (*F *puede ser, por ejemplo* $\mathbb{R}$*), denotado como* $F[x]$*. Las operaciones están definidas de la siguiente manera:*

- ***Suma de polinomios:*** *Si* $p(x)$ *y* $q(x)$ *son polinomios, su suma es el polinomio cuyas coeficientes son la suma de los coeficientes correspondientes de* $p(x)$ *y* $q(x)$.
- ***Multiplicación por un escalar:*** *Si* $\alpha \in F$ *y* $p(x)$ *es un polinomio, entonces* $\alpha p(x)$ *es el polinomio cuyos coeficientes son multiplicados por* α.

d) ***Espacio vectorial de funciones:***

El conjunto de todas las funciones $f : X \to F$*, donde* X *es un conjunto y* F *es un campo, es un espacio vectorial. Las operaciones se definen de la siguiente manera:*

- ***Suma de funciones:*** *Si* f *y* g *son funciones de* X *a* F*, entonces* $(f + g)(x) = f(x) + g(x)$ *para todo* $x \in X$.
- ***Multiplicación por un escalar:*** *Si* $\alpha \in F$ *y* f *es una función de* X *a* F*, entonces* $(\alpha f)(x) = \alpha f(x)$ *para todo* $x \in X$.

e) ***Espacio vectorial de sucesiones:***

El conjunto de todas las sucesiones $a_1, a_2, a_3, \ldots$ *de elementos de un campo* F *forma un espacio vectorial. Las operaciones se definen de la siguiente manera:*

- ***Suma de sucesiones:*** *Si* $a = (a_n)$ *y* $b = (b_n)$ *son sucesiones, entonces* $a + b = (a_n + b_n)$*, es decir, la suma se define término a término.*
- ***Multiplicación por un escalar:*** *Si* $\alpha \in F$ *y* $a = (a_n)$ *es una sucesión, entonces* $\alpha a = (\alpha a_n)$*, es decir, se multiplica cada término de la sucesión por el escalar* α*".*

2.2 Espacios Métricos

Definición 2.2.1 *Un espacio métrico es un conjunto* $\mathbb{X}$ *junto con una función* $d : \mathbb{X}\times\mathbb{X} \longrightarrow \mathbb{R}$ *llamada métrica o distancia que "satisface las siguientes propiedades:*

1. *Identidad:* $d(x,y) \geq 0$ *y* $d(x,y) = 0$ *si y solo si* $x = y$ *para todo par de elementos* $x,y \in \mathbb{X}$.

2. *Simetría:* $d(x,y) = d(y,x)$, *para todo par de elementos* $x,y \in \mathbb{X}$.

3. *Desigualdad triangular:* $d(x,z) \leq d(x,y)+d(y,z)$, *para todo par de elementos* $x,y,z \in \mathbb{X}$*".*

Observación 2.2.1 *La métrica* $d(x,y)$ *es una función que mide la distancia o separación entre los puntos* x *e* y *del espacio métrico* $\mathbb{X}$. *Un ejemplo, es el espacio de las funciones continuas en un intervalo cerrado y acotado con la distancia del supremo, y el espacio de las secuencias de números reales con la distancia del máximo.*

Ejemplo 2.2.1 (La recta) *Es decir, "el conjunto* $\mathbb{R}$ *de los números reales, es el ejemplo más importante de espacios métricos. La distancia entre dos puntos* $x, y \in \mathbb{R}$ *es dada por* $d(x, y) = |x - y|$. *Las propiedades de* 1. *a* 3. *resultan inmediatas de las propiedades de valor absoluto de números reales. Esta es llamada* ***métrica usual*** *de la recta. A menos que se mencione explícitamente lo contrario, es a lo que nos referiremos siempre que consideremos* $\mathbb{R}$ *como espacio métrico".*

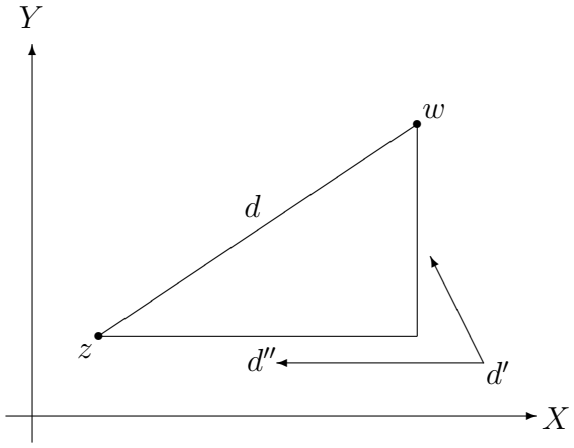
Ejemplo 2.2.2 (El Espacio Euclidiano $\mathbb{R}^n$) *Este ejemplo generaliza el ejemplo anterior. "Los puntos de* $\mathbb{R}^n$ *son de la forma* $x = (x_1, \ldots, x_n)$ *donde cada una de las* n *coordenadas* x_i *es un número real. Hay tres formas naturales de definir la distancia entre dos puntos en* $\mathbb{R}^n$. *Dados* $x = (x_1, \ldots, x_n)$ *y* $y = (y_1, \ldots, y_n)$, *escribiremos:*

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} = \left[\sum_{i=1}^{n}(x_i - y_i)^2\right]^{\frac{1}{2}}$$

$$d'(x,y) = |x_1 - y_1| + \ldots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i|$$

$$d''(x,y) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\} = \max_{1 \le i \le n} |x_i - y_i|$$

La figura siguiente de abajo proporciona una comparación entre las distancias $d(z,w)$, $d'(z,w)$ *y* $d''(z,w)$



Las funciones $d, d', d'' : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ *son métricas". De hecho, cumple obviamente las condiciones* ***1.***, ***2.*** *y* ***3.*** *Se puede obtener una interpretación intuitiva de la métrica* d', *en el caso* $n = 2$, *imaginando que el plano* $\mathbb{R}^2$ *es el plano de una ciudad cuyas calles son rectas paralelas a los ejes coordenados* $x = 0$ *y* $y = 0$. *Entonces el camino más corto que conecta* x *a* y *a través de las calles tiene una longitud igual a* $d'(z,w) = |x - u| + |y - u|$.

2.3 Sucesiones

Definición 2.3.1 *"Una sucesión en un conjunto* $\mathbb{X}$ *es una aplicación* $x : \mathbb{N} \longrightarrow \mathbb{X}$, *definida en el conjunto* $\mathbb{N} = \{1,2,3,\cdots,n,\cdots\}$. *El valor que la sucesión* x *asume en el número* $n \in \mathbb{N}$ *será indicado por* (x_n) *o* $\{x_n\}$, *en vez de* $x(n)$, *será llamado el n-ésimo termino de la sucesión".*

Ejemplo 2.3.1 *Un ejemplo de sucesión la podemos definir de la siguiente forma* $(x_n) = \dfrac{1}{n}$, *así*

tenemos $(x_n) = \left\{\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \cdots, \frac{1}{n}, \cdots\right\}$ *o lo que es equivalente a escribirlo de la siguiente forma* $(x_n) = \left\{1, 0.5, 0.3, 0.25, 0.2, \cdots, \frac{1}{n}, \cdots\right\}$.

Definición 2.3.2 (Límite de una Sucesión) *"Sea* (x_n) *una sucesión en un espacio métrico* $(\mathbb{X},\ d)$*. Se dice que punto* $a \in \mathbb{X}$ *es el límite de la sucesión* (x_n) *cuando, para todo número* $\varepsilon > 0$ *dado arbitrariamente, podemos obtener un* $n_0 \in \mathbb{N}$*", tal que*

$$n > n_0 \implies d(x_n, a) < \varepsilon$$

Entonces, podemos escribir $a = \lim(x_n), a = \lim\limits_{n}(x_n)$ *o* $a = \lim\limits_{n\in\mathbb{N}}(x_n)$*. Decimos también que* (x_n) *tiende hacia* a *y se escribe* $x_n \longrightarrow a$.

Definición 2.3.3 (Convergencia de Sucesiones) *"Cuando existe* $a = \lim(x_n)$ *en* $\mathbb{X}$*, se dice que la sucesión de puntos* $(x_n) \in \mathbb{X}$ *es convergente en* $\mathbb{X}$*, y converge hacia* a*". Si no existe* $\lim(x_n)$ *en* $\mathbb{X}$*, decimos que la sucesión es divergente en* $\mathbb{X}$.

Ejemplo 2.3.2 *Dada la sucesión* $(x_n) = \frac{1}{n}$ *que se dio en el* ***Ejemplo 2.3.1****, se tiene que* $\lim(x_n) = 0$.

"En efecto, dado cualquier $\varepsilon > 0$*, tomamos* $n_0 > \frac{1}{\varepsilon}$ *y vemos que*

$$n > n_0 \implies 0 < \frac{1}{n} < \varepsilon \implies \left|\frac{1}{n} - 0\right| < \varepsilon$$

por lo que tenemos

$$n > n_0 \implies d(x_n,0) < \varepsilon$$

En terminos más geometricos, para todo n suficiente grande, $\frac{1}{n}$ *pertenece al intervalo* $(-\varepsilon, \varepsilon)$*". Con lo que podemos decir que* (x_n) *converge a cero, en otros terminos* $x_n \longrightarrow 0$.

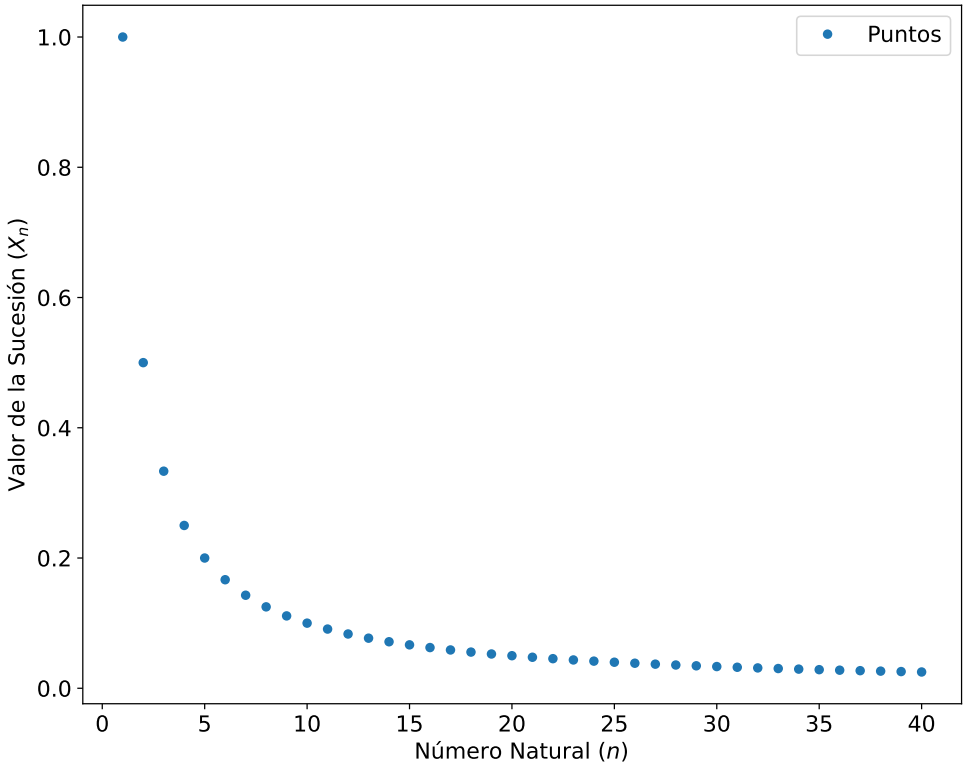


Figura 2.1: Gráfica de la Convergencia de la sucesión $(x_n) = \frac{1}{n}$.

2.4 Subsucesiones

Definición 2.4.1 *"Sea $(x_m) \subseteq \mathbb{X}$ y $k : \mathbb{N} \to \mathbb{N}$ una función creciente. La composición $x \circ k : \mathbb{N} \to \mathbb{X}$ que a cada número natural m le asocia el vector $x \circ k(m) = x_{k_m}$ es llamada subsucesión de (x_m)".*

Observación 2.4.1 *"De ahora en adelante, $(x_{k_m}) \subseteq (x_m)$ significará (x_{k_m}) es una subsucesión de (x_m)".*

El siguiente proposición establece que "cualquier subsucesión de una sucesión convergente es convergente".

Proposición 2.4.1 *"Sea* $(x_m) \subseteq \mathbb{X}$. *Si* $\lim_{m\longrightarrow\infty} x_m = L$ *entonces toda subsucesión de* (x_m) *converge hacia* L*".*

Prueba 2.4.1 *Sea* $\varepsilon > 0$, *por hipótesis* $\exists\, m_0 \in \mathbb{N}$ *tal que* $m \geq m_0$ *implica*

$$d(x_m, L) < \varepsilon$$

Dada $(x_{k_m}) \subseteq (x_m)$, *se tiene que si* $m \geq m_0$ *entonces* $k_m \geq k_{m_0} \geq m_0$, *luego*

$$d(x_{k_m}, L) < \varepsilon$$

Así queda demostrado que toda subsucesión de (x_m) *converge a* L. ■

Otro ejemplo muy interesante de analizar es el siguiente

Ejemplo 2.4.1 *Sea* $x_1 = 1$. *La sucesión* (x_n) *se define recursivamente como*

$$(x_{n+1}) = \frac{1}{2}\left(\frac{x_n^2 + a}{x_n}\right)$$

donde x_n *es el n-ésimo término de la sucesión y* a *es un parámetro positivo.*
Para entender la convergencia de la sucesión, supongamos que (x_n) *converge a un límite* L. *Entonces, si* $x_n \longrightarrow L$, *tenemos que* $x_{n+1} \longrightarrow L$ ***(Proposición (2.4.1))***. *Sustituyendo en la relación de recurrencia, obtenemos:*

$$L = \frac{1}{2}\left(\frac{L^2 + a}{L}\right)$$

Usando propiedades elementales del álgebra tenemos

$$L^2 = a$$

Por lo tanto,

$$L = \sqrt{a}$$

Dado que $a > 0$, el límite $L = \sqrt{a}$ es un número real positivo.

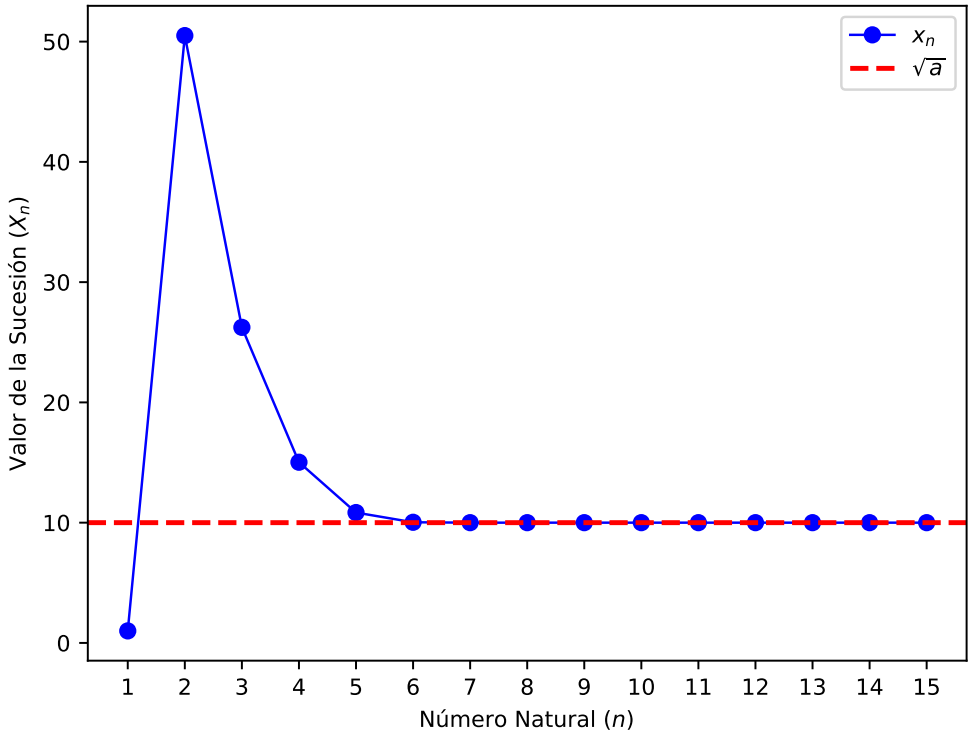


Figura 2.2: Gráfica de la Convergencia de la sucesión $(x_{n+1}) = \frac{1}{2}\left(\frac{x_n^2 + a}{x_n}\right)$, con $a = 100$. Se usa principalmente para calcular raíces cuadradas de números positivos mediante un método iterativo eficiente y de rápida convergencia. Esta técnica es útil en muchos campos de la ciencia y la ingeniería donde se requiere calcular raíces cuadradas con alta precisión.

2.5 Sucesiones de Cauchy

Definición 2.5.1 *Sea $(\mathbb{X},d)$ un espacio métrico, una sucesión (x_n) en $\mathbb{X}$ se dice sucesión de Cauchy si para cualquier número real positivo ε, existe un índice natural N tal que si $n,m > N$ entonces $d(x_n, x_m) < \varepsilon$.*

Observación 2.5.1 *En otras palabras, "la sucesión se acerca cada vez más a un límite a medida que aumenta el número de términos", y esta convergencia puede ser tan precisa como se desee, siempre y cuando se elija un ε suficientemente pequeño.*

Observación 2.5.2 (Propiedad La Arquimediana) *Sean a,$b \in \mathbb{R}^+ \Longrightarrow \exists n \in \mathbb{Z}^+$ tal que $a < nb$. Una consecuencia importante en sucesiones es:*

$$\forall \varepsilon > 0,\ \exists n \in \mathbb{Z}^+ \text{ tal que } \frac{1}{n} < \varepsilon$$

Ejemplo 2.5.1 *Veremos que la sucesión $(x_n) = \dfrac{1}{n}$ visto en el **Ejemplo 2.3.1** es de Cauchy, usando la definición.*

Prueba 2.5.1 *Tomamos $\varepsilon > 0$ por la **Propiedad La Arquimediana (2.5.2)** existe $N \in \mathbb{N}$ tal que*

$$\frac{1}{N} < \frac{\varepsilon}{2}$$

por lo tanto si n,$m \geq N$ tendremos que

$$\mid x_n - x_m \mid = \left| \frac{1}{n} - \frac{1}{m} \right| < \frac{1}{n} + \frac{1}{m} \leq \frac{1}{N} + \frac{1}{N} = \frac{2}{N} < \varepsilon$$

$$d(x_n,x_m) < \varepsilon$$

Así tenemos que (x_n) es de cauchy. ■

Proposición 2.5.1 *"Toda sucesión (x_n) Convergente en un espacio métrico $(\mathbb{X},d)$, es de Cauchy".*

Prueba 2.5.2 *"Sea $\lim x_n = a$ en un espacio métrico $(\mathbb{X},d)$, dado un $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Longrightarrow d(x_n,a) < \dfrac{\varepsilon}{2}$, si tomamos n,$m > n_0$ tenemos*

$$d(x_m,x_n) \leq d(x_m,a) + d(x_n,a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Así la sucesión (x_n) es de Cauchy". ■

Ejemplo 2.5.2 *Como se vio en el* ***Ejemplo 2.3.2****, la sucesión* $(x_n) = \frac{1}{n}$ *es convegente, entonces por* ***Proposición (2.5.1)*** *la sucesión* (x_n) *es de Cauchy.*

Observación 2.5.3 *Es importante destacar que en un* ***espacio métrico completo*** *(que más adelante se define detalladamente), "una sucesión de Cauchy converge a un límite bien definido; mientras que en un espacio métrico no completo, es posible tener sucesiones de Cauchy que no converjan a ningún límite en el espacio".*

2.6 Espacios Normados

Definición 2.6.1 *"Un espacio vectorial normado es un espacio vectorial* $\mathbb{X}$ *sobre un campo escalar* $\mathbb{K}$ *(que suele ser* $\mathbb{R}$ *o* $\mathbb{C}$*) que está equipado con una norma* $\| \cdot \|$*, es una función que asigna a cada vector* $x \in \mathbb{X}$ *un número real no negativo* $x \longmapsto \| x \|$*", satisfaciendo las siguientes propiedades:*

1. ***No negatividad:*** *"* $\| x \| \geq 0$ *para todo* $x \in \mathbb{X}$*,* $y \| x \| = 0$ *si y solo si* $x = 0$*".*
2. ***Homogeneidad absoluta:*** *"* $\| \alpha x \| = | \alpha | \| x \|$ *para todo* $x \in \mathbb{X}$ *y* $\alpha \in \mathbb{K}$*".*
3. ***Desigualdad del triángulo:*** *"* $\| x + y \| \leq \| x \| + \| y \|$ *para todo* x*,*$y \in \mathbb{X}$*".*

Observación 2.6.1 *Un espacio normado es un espacio vectorial con una noción de longitud o tamaño para los vectores, que cumple con ciertas propiedades algebraicas y geométricas. La norma es una herramienta importante para medir distancias entre vectores y para definir la convergencia en el espacio.*

Definición 2.6.2 (Distancia inducida) *Dado un espacio métrico* $(\mathbb{X},d)$*, la distancia inducida por una función de norma* $\| \cdot \| \colon \mathbb{X} \longrightarrow \mathbb{R}$ *se define como:*

$$d(x,y) = \| x - y \|$$

donde $x, y \in \mathbb{X}$ son dos elementos del espacio. "Esta definición de la distancia inducida se extiende naturalmente a cualquier espacio métrico $\mathbb{X}$ mediante la definición:

$$d(x,y) = \text{ínf} \left\{ \sum_{i=1}^{n} \parallel x_i - x_{i-1} \parallel \;\; \middle| \;\; n \in \mathbb{N},\ x_0 = x,\ x_n = y,\ x_i \in \mathbb{X};\ \textit{para } 1 \leq i \leq n \right\}$$

esta definición generaliza la noción de distancia en el espacio euclidiano", donde la distancia entre dos puntos x,$y \in \mathbb{X} = \mathbb{R}^n$ se define como:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Observación 2.6.2 *La distancia inducida es una forma de medir la separación entre dos elementos de un espacio métrico. Es importante señalar que, en general, la distancia inducida no necesariamente proviene de una norma, pero tiene propiedades similares a las de una distancia inducida por una norma, como la desigualdad triangular y la identidad de la distancia.*

Definición 2.6.3 *Sea $\mathbb{X}$ un espacio métrico con distancia (o métrica) d. El diámetro de un subconjunto $C \in \mathbb{X}$ es*

$$diam\ C = \sup_{(x,y) \in C \times C} d(x,y)$$

La distancia a un conjunto $C \in \mathbb{X}$ es la función

$$d_C : \mathbb{X} \longrightarrow [0, +\infty]$$
$$x \longmapsto \text{ínf}\ \ d(x,C)$$

Cuando $C = \emptyset$ tenemos que la distancia de cualquier punto a C es $d_C = +\infty$.

Ejemplo 2.6.1 *"Sea C un subconjunto no vacío de un espacio métrico $(\mathbb{X}, d)$". Entonces*

$$\forall x,y \in \mathbb{X}, \quad \Big| d_C(x) - d_C(y) \Big| \leq d(x,y)$$

Prueba 2.6.1 *"Tomando x,$y \in \mathbb{X}$. Entonces $\forall z \in \mathbb{X}$, $d(x,z) \leq d(x,y) + d(y,z)$, podemos tomar el ínfimo de $z \in C$ y obtener $d_C(x) \leq d(x,y) + d_C(y)$, por lo tanto*

$$d_C(x) - d_C(y) \leq d(x,y) \tag{6.1}$$

Intercambiando x e y, obtenemos $d_C(y) - d_C(x) \leq d(y,x)$, ahora multiplicamos por (-1) y usanado $d(x,y) = d(y,x)$" tenemos

$$-d(x,y) \leq d_C(x) - d_C(y) \tag{6.2}$$

*por último, de **(6.1)** Y **(6.2)***

$$-d(x,y) \leq d_C(x) - d_C(y) \leq d(x,y)$$

por propiedad de valor absoluto

$$\Big|d_C(x) - d_C(y)\Big| \leq d(x,y)$$

así obtenemos lo pedido. ■

2.7 Espacios de Banach

Definición 2.7.1 *"Un espacio de Banach es un espacio vectorial normado $(\mathbb{X}, \|\cdot\|)$ **completo**". Es decir, un espacio vectorial $\mathbb{X}$, equipado con una norma $\|\cdot\|$ que satisface las propiedades de los espacios normados y adicionalmente la siguiente propiedad:*

4. *"El espacio vectorial normado $\mathbb{X}$ es **completo**, es decir, toda sucesión de Cauchy en $\mathbb{X}$ converge a un elemento de $\mathbb{X}$". Como se detalla en **Observación 2.5.3**.*

Ejemplo 2.7.1 *Tenemos algunos ejemplos de Espacios de Banach*

1. *El cuerpo escalar $\mathbb{K}$, con la norma del valor absoluto o módulo, es un Espacio de Banach.*

2. *Fijando $n \in \mathbb{N}$, el espacio $\mathbb{K}^n$, con cada una de las normas*

$$\| x \|_p = \left\{ \sum_{i=1}^{n} | x_i |^p \right\}^{\frac{1}{p}} \quad con \quad x \in \mathbb{K}^n, \ 1 \leq p < \infty$$

$$\| x \|_\infty = \max_{1 \leq i \leq n} | x_i | \quad con \quad x \in \mathbb{K}^n$$

Es un Espacio de Banach.

3. *Para $1 \leq p < \infty$, el espacio de las sucesiones de escalares con potencia p-ésima absolutamente sumable,*

$$\ell^p = \left\{ x = \{x_n\}_{n=1}^{\infty} : x_n \in \mathbb{K} \ (n \in \mathbb{N}), \ \sum_{n=1}^{\infty} | x_n |^p < \infty \right\}$$

provisto de la norma

$$\| x \|_p = \left\{ \sum_{i=1}^{n} | x_n |^p \right\}^{\frac{1}{p}} \quad con \quad x \in \ell^p$$

Es un Espacio de Banach.

4. *Dado un espacio medible $(\mathbb{X}, \Sigma, \mu)$, donde $\mathbb{X}$ es el conjunto medible, Σ es una sigma-álgebra de conjuntos en $\mathbb{X}$ y μ es una medida positiva finita definida en Σ, el espacio L^p consiste en todas las funciones medibles $F : \mathbb{X} \longrightarrow \mathbb{R}$ (o $F : \mathbb{X} \longrightarrow \mathbb{C}$ en el caso complejo) que satisfacen la siguiente condición*

$$\| F \|_p = \left(\int_{\mathbb{X}} | F |^p \, d\mu \right)^{\frac{1}{p}} < \infty$$

Aquí, se cumple que p es un número real mayor que 1. La norma $\| F \|_p$ es conocida como la norma L_p y se calcula como la potencia p-ésima de la integral de valor absoluto

de F, elevada a $\frac{1}{p}$. El espacio L^p es un espacio de Banach.

Definición 2.7.2 (Bolas) *"Dado un espacio métrico ($\mathbb{X}$,d), una bola abierta centrada en un punto $x_0 \in \mathbb{X}$ y de radio $r > 0$ se define como el conjunto":*

$$B(x_0, r) = \{x \in \mathbb{X} : d(x,x_0) < r\}$$

Es decir, "la bola abierta es el conjunto de todos los puntos en $\mathbb{X}$ cuya distancia al punto x_0 es menor que r. La bola abierta de radio r centrada en x_0 se puede denotar como $B_r(x_0)$. Por otro lado, una bola cerrada centrada en un punto $x_0 \in \mathbb{X}$ y de radio $r > 0$ se define como el conjunto:

$$B[x_0, r] = \{x \in \mathbb{X} : d(x,x_0) \leq r\}$$

es decir, la bola cerrada es el conjunto de todos los puntos en $\mathbb{X}$ cuya distancia al punto x_0 es menor o igual que r". La bola cerrada de radio r centrada en x_0 se puede denotar como $\overline{B}_r(x_0)$.

Observación 2.7.1 *Cabe destacar que en algunos contextos, como el análisis funcional, la notación $B(x_0,r)$ se utiliza para denotar la bola cerrada, mientras que la notación $B(x_0,r)\backslash x_0$ se utiliza para denotar la bola abierta. Sin embargo, en este contexto, la notación descrita anteriormente es la más común en el análisis métrico.*

2.8 Convergencia de Sucesiones es Espacios de Banach

Proposición 2.8.1 *"Una sucesión (x_n) en un espacio de Banach $\mathbb{X}$ converge fuertemente a algún $x \in \mathbb{X}$", y escribimos $x_n \longrightarrow x$, cuando $n \longrightarrow \infty$, si y sólo si*

$$\lim_{n \longrightarrow \infty} \|x_n - x\| = 0 \ , \quad \forall x \in \mathbb{X}$$

El punto x es el límite de la sucesión. Por otro lado, (x_n) tiene la propiedad de Cauchy, o es una sucesión de Cauchy, si

$$\lim_{n,m\longrightarrow\infty} \|x_n - x_m\| = 0 \ , \quad \text{para } n,m > N$$

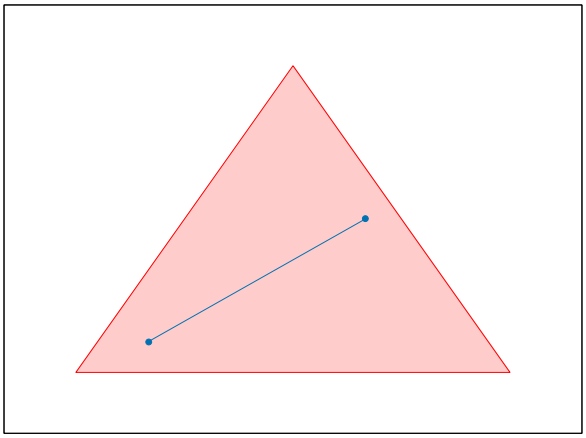
Toda sucesión convergente tiene la propiedad de Cauchy, y toda sucesión de Cauchy está acotada. "Un espacio de Banach es un espacio normado en el cual toda sucesión de Cauchy es convergente, una propiedad llamada **Completitud**".

2.9 Conjuntos Convexos

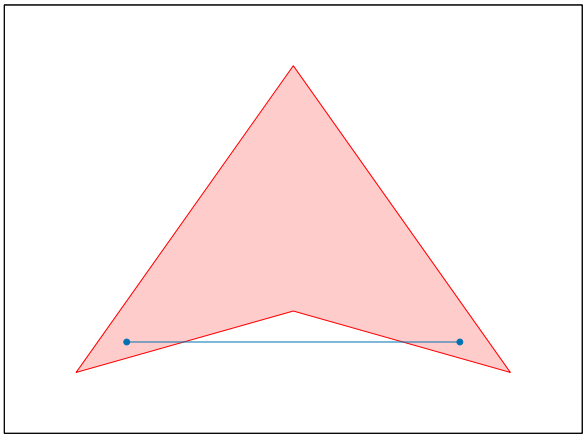
Definición 2.9.1 *"Un conjunto C en un espacio vectorial es convexo si para cualquier par de puntos $x, y \in C$, el segmento de línea que conecta a x e y también está contenido en C".*
Esto se puede expresar como:

$$\forall x,y \in C \implies tx + (1-t)y \in C, \ \forall t \in [0,1]$$

donde $tx + (1-t)y$ es el punto en el segmento de línea que conecta x e y correspondiente al valor de t del camino desde x hasta y.



(a) Conjunto Convexo



(b) Conjunto no Convexo

Si el conjunto no cumple con esta propiedad, es decir, si existe un par de puntos x,$y \in C$ tal que el segmento de línea que los conecta no está completamente contenido en C, entonces se dice que el conjunto es no convexo.

Ejemplo 2.9.1 *Probaremos que el conjunto definido como* $E = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ *es convexo.*

Prueba 2.9.1 *Sean* (a_1,a_2), $(b_1,b_2) \in E$ *y* $\forall t \in [0,1]$, *"entonces*

$$\| t(a_1,a_2) + (1-t)(b_1,b_2) \| \leq t \| (a_1,a_2) \| + (1-t) \| (b_1,b_2) \|$$

donde $\| \cdot \|$ *es la norma euclidiana en* $\mathbb{R}^2$. *Como* $(a_1,a_2), (b_1,b_2) \in E$, *se tiene:*

$$t \| (a_1,a_2) \| + (1-t) \| (b_1,b_2) \| \leq t(1) + (1-t)(1) = 1$$

Entonces obtenemos:

$$\| t(a_1,a_2) + (1-t)(b_1,b_2) \| \leq 1$$

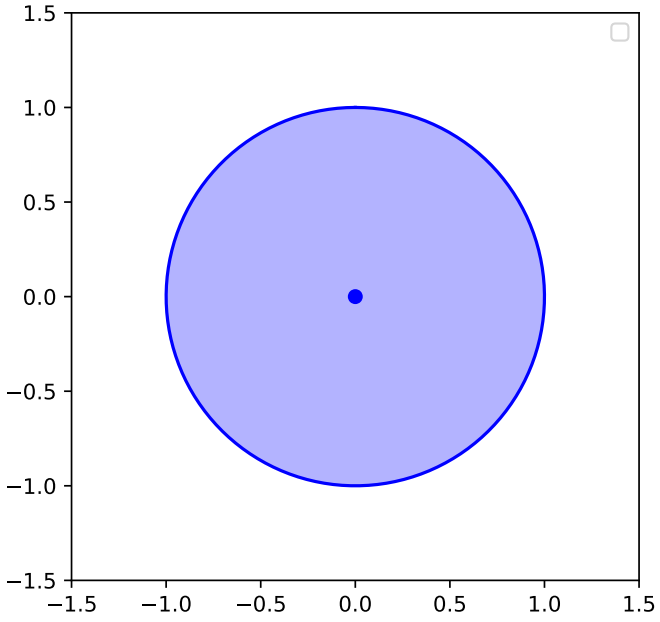


Figura 2.3: Es la gráfica del conjunto $E = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$.

Por lo tanto, concluimos que $t(a_1,a_2) + (1-t)(b_1,b_2) \in E$*".* ■

2.10 Límite Inferior

Definición 2.10.1 *El límite inferior de una función* $F : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *en* x_0 *es el valor en* $\mathbb{R} \cup \{+\infty\}$ *"definido por:*

$$\liminf_{x \longrightarrow x_0} F(x) := \lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} F(x) \right] = \sup_{\delta > 0} \left[\inf_{x \in B[x_0, \delta]} F(x) \right] = \sup_{V \in \mathcal{N}(x_0)} \left[\inf_{x \in V} F(x) \right]$$

donde $V \in \mathcal{N}(x_0)$ *es la colección de todos las vecindades de* x_0*. Entonces es una solución localmente óptima al problema de minimizar* F*". Por un vecindario de* x_0 *se entiende cualquier conjunto que contenga* x *en su interior, por ejemplo, una bola cerrada.*

Proposición 2.10.1 (máximo y mínimo de una suma) *"Para cualesquiera funciones* F_1

y F_2 con valores reales extendidos, se tiene (bajo la convención $\infty - \infty = \infty$) las estimaciones:

$$\inf_{x\in\mathbb{X}} [F_1(x) + F_2(x)] \geq \inf_{x\in\mathbb{X}} F_1(x) + \inf_{x\in\mathbb{X}} F_2(x),$$

$$\sup_{x\in\mathbb{X}} [F_1(x) + F_2(x)] \leq \sup_{x\in\mathbb{X}} F_1(x) + \sup_{x\in\mathbb{X}} F_2(x),$$

así como la regla":

$$\inf_{(x_1,x_2)\in\mathbb{X}_1\times\mathbb{X}_2} [F_1(x_1) + F_2(x_2)] = \inf_{x_1\in\mathbb{X}_1} F_1(x_1) + \inf_{x_2\in\mathbb{X}_2} F_2(x_2).$$

Proposición 2.10.2 (múltiplos escalares frente a máximo y mínimo) *"Para cualquier función F con valores reales extendidos, se tiene (bajo la convención $0 \cdot \infty = 0$) que:*

$$\inf_{x\in X} \lambda F(x) = \lambda \inf_{x\in X} F(x), \quad \sup_{x\in X} \lambda F(x) = \lambda \sup_{x\in X} F(x) \quad \textit{cuando } \lambda \geq 0,$$

mientras que", por otro lado:

$$\inf_{x\in X} F(x) = -\sup_{x\in X} \{-F(x)\}, \quad \sup_{x\in X} F(x) = -\inf_{x\in X} \{-F(x)\}.$$

Proposición 2.10.3 (límites inferiores de sumas) *"Para funciones arbitrarias con valores reales extendidos, se tiene que:*

$$\liminf_{x\longrightarrow x_0} [F_1(x) + F_2(x)] \geq \liminf_{x\longrightarrow x_0} F_1(x) + \liminf_{x\longrightarrow x_0} F_2(x)$$

si la suma en el lado derecho no es $\infty - \infty$". Por otro lado, siempre se tiene:

$$\liminf_{x\longrightarrow x_0} \lambda F(x) = \lambda \liminf_{x\longrightarrow x_0} F(x) \quad \textit{cuando } \lambda \geq 0.$$

Prueba 2.10.1 *A partir de la definición de* ***límite inferior (2.10.1)****, "calculamos:*

$$\liminf_{x \longrightarrow x_0} [F_1(x) + F_2(x)] = \lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} (F_1(x) + F_2(x)) \right]$$

por la ***Proposición (2.10.1)***

$$\lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} (F_1(x) + F_2(x)) \right] \geq \lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} F_1(x) + \inf_{x \in B[x_0, \delta]} F_2(x) \right]$$

por la linealidad de los límites:

$$\lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} F_1(x) + \inf_{x \in B[x_0, \delta]} F_2(x) \right] = \lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} F_1(x) \right] + \lim_{\delta \downarrow 0} \left[\inf_{x \in B[x_0, \delta]} F_2(x) \right]$$

así obtenemos lo siguiente

$$\liminf_{x \longrightarrow x_0} [F_1(x) + F_2(x)] \geq \liminf_{x \longrightarrow x_0} F_1(x) + \liminf_{x \longrightarrow x_0} F_2(x)$$

siempre que esta suma no sea $\infty - \infty$. La segunda fórmula de esta proposición es fácilmente verificable a partir de la regla de multiplicación escalar en la ***Proposición (2.10.2)*** *cuando $\lambda > 0$". Cuando $\lambda = 0$, es trivial porque ambos lados de la ecuación deben ser 0.* ■

2.11 Semicontinuidad Inferior

Definición 2.11.1 *"La función $F : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ es semicontinua inferiormente (lsc) en x_0", si se cumple:*

$$\liminf_{x \longrightarrow x_0} F(x) \geq F(x_0)$$

o equivalentemente:

$$\liminf_{x \longrightarrow x_0} F(x) = F(x_0)$$

y es semicontinua inferiormente en $\mathbb{X}$ *si esto se cumple para todo* $x_0 \in \mathbb{X}$.

Sea $F : C \subseteq \mathbb{X} \longrightarrow \mathbb{R}$ una función objetivo, con C un subconjunto cerrado y convexo del espacio de banach $\mathbb{X}$.

2.12 Funciones Lipschitzianas

Definición 2.12.1 *"F es* $\mathbb{R}-$*lipschitz en C, entonces F satisface la condición de Lipschitz, si existe un número* $L > 0$, *para todo x,y* $\in C$*" tal que*

$$\Big|F(x) - F(y)\Big| \leq L \parallel x - y \parallel$$

Diremos que F es Lipschitziana al rededor de p (o cerca de p o localmente Lipschitziana en p) si existe una vencidad de p, $B(p, \delta)$ *y un número* $L > 0$, *para todo x,y* $\in B(p, \delta)$ *tal que*

$$\Big|F(x) - F(y)\Big| \leq L \parallel x - y \parallel$$

Ejemplo 2.12.1 *La función norma* $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ *definida por* $F(\overline{x}) = \|a\overline{x} + b\|$ *es Lipschiziana.*

Prueba 2.12.1 *"En efecto,*

$$\Big|\ F(\overline{x}) - F(\overline{y})\ \Big| = \Big|\ \parallel a\overline{x} + b \parallel - \parallel a\overline{y} + b \parallel\ \Big|$$

$$\Big|\ F(\overline{x}) - F(\overline{y})\ \Big| \leq \parallel a\overline{x} + b - a\overline{y} - b \parallel$$

$$\Big|\ F(\overline{x}) - F(\overline{y})\ \Big| \leq \parallel a\overline{x} - a\overline{y} \parallel$$

$$\Big|\ F(\overline{x}) - F(\overline{y})\ \Big| \leq \mid a \mid \cdot \parallel \overline{x} - \overline{y} \parallel$$

Siendo la constante Lipschitziana $L = \mid a \mid$*".* ■

Ejemplo 2.12.2 *La función valor absoluto* $F : \mathbb{R} \longrightarrow \mathbb{R}$ *definida por* $F(x) = \mid ax + b \mid$ *es Lipschiziana.*

Prueba 2.12.2 *En efecto,*

$$\left| \; F(x) - F(y) \; \right| = \left| \; \mid ax + b \mid - \mid ay + b \mid \; \right|$$

$$\left| \; F(x) - F(y) \; \right| \leq \; \mid a \mid \cdot \mid x - y \mid$$

Siendo la constante Lipschitziana $L = \mid a \mid$. ■

Ejemplo 2.12.3 *Análogamente que en el* ***Ejemplo 2.12.2*** *se puede comprobar que la función* $F(x_1,x_2) = \mid ax_1 + bx_2 + c \mid$ *es Lipschiziana.*

Prueba 2.12.3 *Daremos forma a la función valor absoluto:*

$$F(\overline{x}) = |\langle \overline{e},\overline{x} \rangle + c \mid$$

Conideranda que $\overline{x} = (x_1,x_2)$ *y* $\overline{e} = (a,b)$

$$\left| F(\overline{x}) - F(\overline{y}) \right| \; = \; \left| \; \mid \langle \overline{e},\overline{x} \rangle + c \mid - \mid \langle \overline{e},\overline{y} \rangle + c \mid \; \right|$$

$$\left| F(\overline{x}) - F(\overline{y}) \right| \; \leq \; \left| \; \langle \overline{e},\overline{x} \rangle - \langle \overline{e},\overline{y} \rangle \; \right|$$

$$\left| F(\overline{x}) - F(\overline{y}) \right| \; \leq \; \left| \; \langle \overline{e},\overline{x} - \overline{y} \rangle \; \right|$$

$$\left| F(\overline{x}) - F(\overline{y}) \right| \; \leq \| \; \overline{e} \; \| \cdot \| \; \overline{x} - \overline{y} \; \|$$

Es decir,

$$\left| F(x_1,x_2) - F(y_1,y_2) \right| \; \leq \| \; (a,b) \; \| \cdot \| \; (x_1,x_2) - (y_1,y_2) \; \|$$

Siendo la constante Lipschitziana $L = \| \; (a,b) \; \|$. ■

2.13 Análisis Convexo

Definición 2.13.1 (Función Convexa) *"F es $\mathbb{R}-$convexa en C, si para x,$y \in C$ y para todo $t \in [0,1]$", se cumple:*

$$F\left((1-t)x+ty\right) \leq (1-t)F(x)+tF(y)$$

En otras palabras, el valor de la función en cualquier punto en la línea que une x e y es como máximo el valor promedio ponderado de la función en x e y.

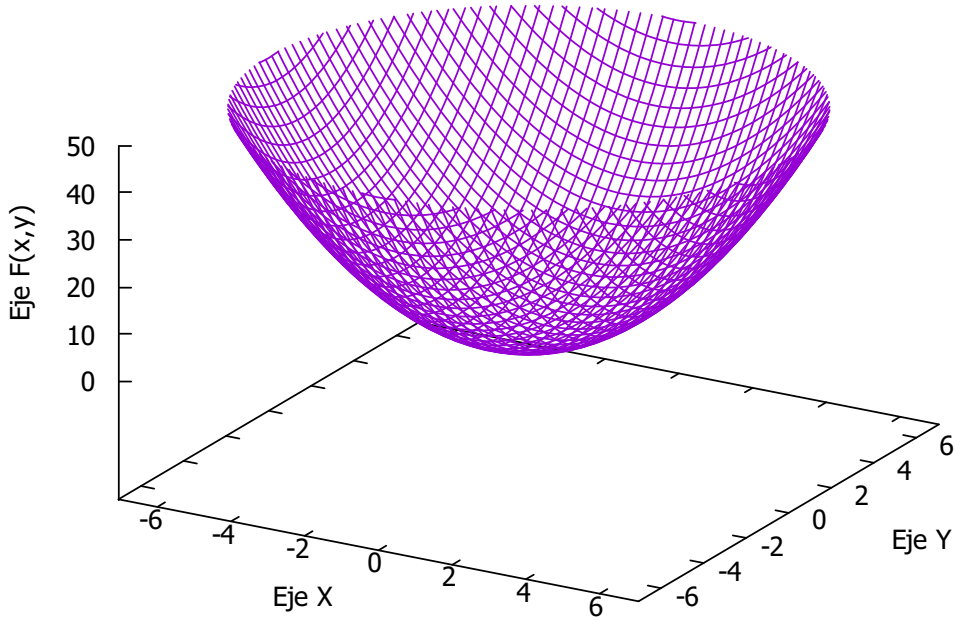


Figura 2.4: Gráfica de la función Convexa $F(x,y) = x^2 + y^2$.

Observación 2.13.1 *Las funciones convexas son muy relevantes en la optimización porque los problemas de optimización que involucran funciones convexas suelen ser más fáciles de*

resolver. Esto se debe al siguiente ***Corolario***

Corolario 2.13.1 *En una función convexa, cualquier mínimo local es también un mínimo global. Esto significa que, si F es convexa y x^* es un punto tal que $F(x^*) \leq F(x)$ para todos los x en un entorno de x^*, para todo x en el dominio de F.*

Definición 2.13.2 (Conjunto de Nivel) *"El conjunto de nivel de la función F está definida por":*

$$\underset{\alpha}{\text{Lev}}\, F(x) = \{x \in C \mid F(x) \leq \alpha,\ \alpha \in \mathbb{R}\}$$

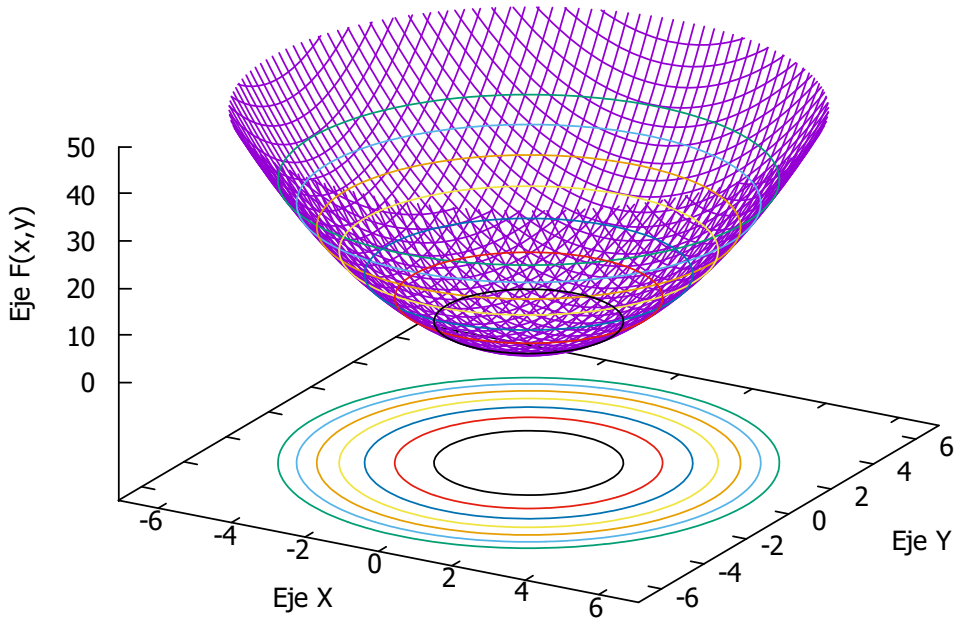


Figura 2.5: Es la gráfica de $F(x,y) = x^2 + y^2$ con su Conjunto de Nivel reflejado en el $EjeXY$.

Definición 2.13.3 *"El conjunto de nivel estricto de la función F está definida por":*

$$\underset{<\alpha}{\text{Lev}}(F) = \{x \in C \mid F(x) < \alpha,\ \alpha \in \mathbb{R}\}$$

Definición 2.13.4 (Argumento del mínimo) *El conjunto de minimizadores o Argumento del mínimo de F es*

$$Argmin(F) = \{\widehat{x} \in \mathbb{X} \mid F(\widehat{x}) \leq F(x) \ , \quad para\ todo\ x \in \mathbb{X}\}$$

Otra forma de representar el Argmin es en función del conjunto del nivel, de la siguiente forma

$$Argmin(F) = \bigcap_{\alpha > \text{ínf}(F)} \operatorname*{Lev}_{\alpha}(F)$$

Definición 2.13.5 *Si $x \in dom(F)$, entonces $x \in \operatorname*{Lev}_{\alpha}(F)$. Por lo tanto*

$$Dom(F) = \bigcup_{\alpha \in \mathbb{R}} \operatorname*{Lev}_{\alpha}(F)$$

"La $\mathbb{R}-$cuasiconvexidad de F en C no es más que una función caracterizada por la convexidad de su conjunto nivel $\operatorname*{Lev}_{\alpha} F(x)$". Tenemos por tanto la siguiente definición

Definición 2.13.6 (Función Cuasiconvexa Clasica) *F es $\mathbb{R}-$cuasiconvexa en C, si para x,$y \in C$ y para todo $t \in [0,1]$, se cumple*

$$F(tx + (1-t)y) \leq \text{máx}\,\{F(x),F(y)\}$$

donde el máximo se toma por componentes. O, equivalentemente

$$x,y \in C; \quad F(x) \leq F(y) \Longrightarrow F(tx + (1-t)y) \leq F(y), \quad \forall t \in [0,1]$$

Definición 2.13.7 (Función Cuasiconvexa por Conjunto de Nivel) *F es $\mathbb{R}-$cuasiconvexa en C, si $\operatorname*{Lev}_{\alpha} F(x)$ es convexa para cualquier $\alpha \in \mathbb{R}$.*

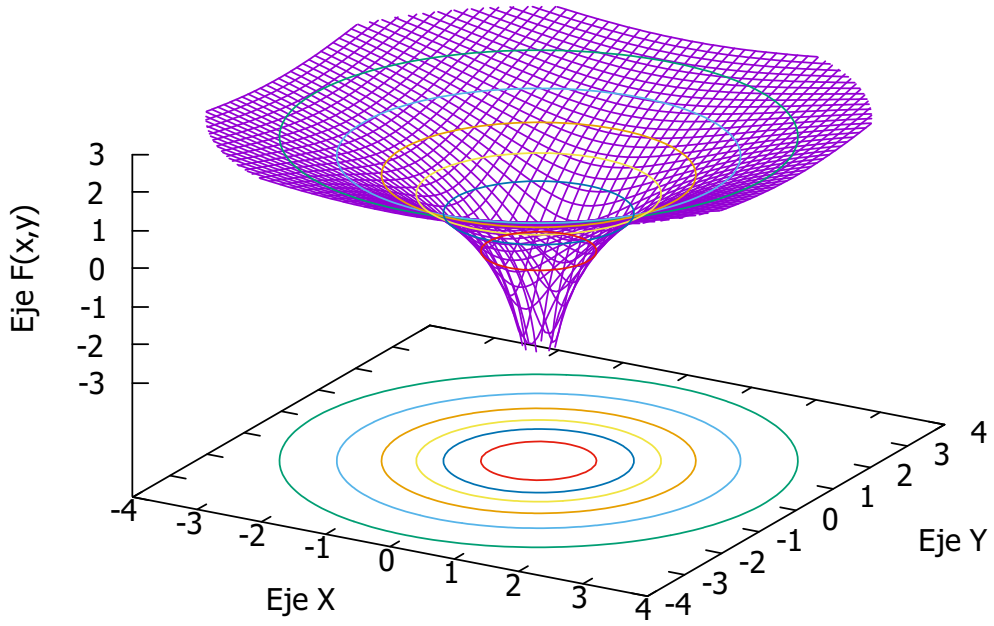


Figura 2.6: "Es la gráfica de la función Cuasiconvexa $F(x,y) = log(x^2 + y^2)$, conjunto de nivel $\underset{\alpha}{\text{Lev}}\, F(x)$ Convexa".

Observación 2.13.2 *Al igual que con las funciones convexas, si F es cuasiconvexa, entonces cualquier mínimo local de F es también un mínimo global. Sin embargo, esta propiedad no es tan fuerte como en el caso de funciones convexas debido a que la cuasiconvexidad no garantiza una forma* ***suave*** *o una curvatura específica.*

Observación 2.13.3 *Toda función convexa es también cuasiconvexa, pero no toda función cuasiconvexa es convexa. La cuasiconvexidad es una condición más débil que la convexidad (profundizaremos en el capitulo siguiente).*

2.14 El Teorema de Separación de Hahn-Banach

El Teorema de Separación de Hahn-Banach es fundamental en el análisis funcional y convexo, y tiene varias consecuencias importantes. El Teorema de Separación de Hahn-Banach constituye un principio fundamental en el ámbito del análisis funcional y convexo. Tal y como se demostrará, este teorema presenta una serie de implicaciones significativas (Peypouquet (2015)).

Teorema 2.14.1 (Teorema de Separación de Hahn-Banach) *Sean A y B subconjuntos convexos, disjuntos y no vacíos de un espacio normado $(X, \|\cdot\|)$*

i) *Si A es abierto, entonces existe $L \in X^* \setminus \{0\}$ tal que $\langle L, x\rangle < \langle L, y\rangle$ para cada $x \in A$ y $y \in B$.* *

ii) *Si A es compacto y B es cerrado, entonces existen $L \in X^* \setminus \{0\}$ y $\varepsilon > 0$ tales que $\langle L, x\rangle + \varepsilon \leq \langle L, y\rangle$ para cada $x \in A$ e $y \in B$.*

Teorema 2.14.2 *Sea C un subconjunto convexo, no vacío, de un espacio normado $(X, \|\cdot\|)$ que no contiene el origen.*

i) *Si C es abierto, entonces existe $L \in X^* \setminus \{0\}$ tal que $\langle L, x\rangle < 0$ para cada $x \in C$.*

ii) *Si C es cerrado, entonces existen $L \in X^* \setminus \{0\}$ y $\varepsilon > 0$ tales que $\langle L, x\rangle + \varepsilon \leq 0$ para cada $x \in C$.*

Resulta claro que el **Teorema (2.14.2)** es un caso específico del **Teorema (2.14.1)**. Para verificar su equivalencia, simplemente definimos $C = A - B$ y observamos que C es abierto si

*Donde X^* denota el **espacio dual topológico** de X. Este es el conjunto de todos los funcionales lineales acotados definidos en X. Es decir, $X^* = \mathcal{L}(X; \mathbb{R})$ (o $\mathcal{L}(X; \mathbb{C})$ si X es un espacio vectorial complejo). "Un funcional lineal acotado es una función $L : X \longrightarrow \mathbb{R}$ (o $\mathbb{C}$) que satisface las siguientes propiedades":

- **Linealidad**: $L(ax + by) = aL(x) + bL(y)$ para todos $x, y \in X$ y $a, b \in \mathbb{R}$ (o $\mathbb{C}$).
- **Acotación**: Existe una constante C tal que $|L(x)| \leq C\|x\|$ para todo $x \in X$.

A es abierto, y que C es cerrado si A es compacto y B es cerrado.

Además, la parte ii) del **Teorema (2.14.2)** se puede deducir fácilmente de la parte i) del **Teorema (2.14.1)** al considerar una bola abierta A lo suficientemente pequeña alrededor del origen (que no intersecta a C) y definiendo $B = C$.

Finalmente, en espacios de dimensión finita, la parte i) del **Teorema (2.14.2)** puede demostrarse sin necesidad de hacer suposiciones topológicas sobre los conjuntos involucrados. En particular, tenemos lo siguiente:

Proposición 2.14.1 *"Dado $N \geq 1$, sea C un subconjunto convexo no vacío de $\mathbb{R}^N$ que no contiene el origen. Entonces, existe $v \in \mathbb{R}^N \setminus \{0\}$ tal que $v \cdot x \leq 0$ para cada $x \in C$. En particular, si $N \geq 2$ y C es abierto, entonces*

$$V = \{x \in \mathbb{R}^N : v \cdot x = 0\}$$

es un subespacio no trivial de $\mathbb{R}^N$ que no intersecta a C".

Prueba 2.14.1 *Sea $(x_n) \subseteq C$ tal que el conjunto $\{x_n : n \geq 1\}$ es denso en C. Sea C_n la envolvente convexa del conjunto $\{x_k : k = 1, \ldots, n\}$ y sea p_n el elemento de menor norma de C_n. "Por convexidad, para cada $x \in C_n$ y $t \in (0,1)$, tenemos*

$$\|p_n\|^2 \leq \|p_n + t(x - p_n)\|^2 = \|p_n\|^2 + 2t\, p_n \cdot (x - p_n) + t^2\|x - p_n\|^2$$

por lo tanto,

$$0 \leq 2\|p_n\|^2 \leq 2\, p_n \cdot x + t\|x - p_n\|^2$$

Dejando $t \longrightarrow 0$, deducimos que $p_n \cdot x \geq 0$ para todo $x \in C_n$. A continuación podemos escribir $v_n = -p_n/\|p_n\|$". La sucesión (v_n) se encuentra en la esfera unitaria, que es compacta. Podemos extraer una subsecuencia que converge a algún $v \in \mathbb{R}^N$ con $\|v\| = 1$ (así $v \neq 0$) y $v \cdot x \leq 0$ para todo $x \in C$.

3 Espacios de Hilbert

"Sea $\mathcal{H}$ un espacio de Hilbert real con producto interno $\langle\cdot\rangle$. La norma asociada se denota por $\|\cdot\|$ y la distancia asociada por d", es decir

$$\forall x,y \in \mathcal{H}; \quad \|x\| = \sqrt{\langle x,x\rangle} \ \text{ y } \ d(x,y) = \|x-y\|$$

El operador de identidad en $\mathcal{H}$ se denota por Id.

Proposición 3.0.1 *Para cada $x,y \in \mathcal{H}$, se tiene:*

1. *La desigualdad de Cauchy-Schwarz: $|\langle x,y\rangle| \leq \|x\| \ \|y\|$*
2. *La desigualdad triangular: $\|x+y\| \leq \|x\| + \|y\|$*

Ejemplo 3.0.1 *Algunos espacios de Hilbert*

1. *El espacio Euclideano $\mathbb{R}^n$ con el producto interno: $\langle x,y\rangle = x \cdot y$.*
2. *El espacio $\ell^2(\mathbb{N};\mathbb{R})$ de sucesiones reales $x = (x_n)$ tales que $\sum_{n\in\mathbb{N}} x_n^2 < +\infty$, con el producto interno: $\langle x,y\rangle = \sum_{n\in\mathbb{N}} x_n y_n$.*
3. *Sea Ω un dominio acotado de $\mathbb{R}^n$. El espacio $\ell^2(\Omega;\mathbb{R}^n)$ de las funciones medibles $\phi : \Omega \longrightarrow \mathbb{R}^M$ tales que*

$$\int_\Omega \phi_m(\zeta)^2 d\zeta < +\infty \ , \quad \text{para } m = 1,2,\cdots,M$$

con el producto interno: $\langle \phi,\psi\rangle = \sum_{m=1}^{M} \int_\Omega \phi_m(\zeta)\psi_m(\zeta)d\zeta$

3.1 Convergencia de Sucesiones en Espacios de Hilbert

Proposición 3.1.1 *"Una sucesión (x_n) en un espacio de Hilbert $\mathcal{H}$ converge débilmente a algún $x \in \mathcal{H}$", si y sólo si*

$$\lim_{x \longrightarrow \infty} \langle x_n - x, y \rangle = 0 \ , \quad \forall y \in \mathcal{H}$$

La convergencia fuerte y la débil se relacionan de la siguiente forma:

Proposición 3.1.2 *Una sucesión (x_n) converge fuertemente a $\widehat{x}$ si, y sólo si, converge débilmente a $\widehat{x}$ y*

$$\limsup_{n \longrightarrow \infty} \parallel x_n \parallel \leq \parallel \widehat{x} \parallel$$

3.2 Función Convexa

Definición 3.2.1 *"Sea $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ una función propia. Decimos que F es función Convexa si su epígrafo $epi(F) = \{(x,\xi) \in \mathcal{H} \times \mathbb{R} \mid F(x) \leq \xi\}$ es un subconjunto convexo de $\mathcal{H} \times \mathbb{R}$".*

Proposición 3.2.1 *Sea $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ un función Convexa. Entonces su dominio $dom(F) < +\infty$ es Convexo.*

Proposición 3.2.2 *"Sea $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ una función propia". Entonces F es Convexa si y sólo si*

$$(\forall x,y \in dom(F)) \ (\forall \alpha \in \ ]0,1[)$$

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y) \tag{2.1}$$

Prueba 3.2.1 *Si $F \equiv +\infty \Longleftrightarrow epi(F) = \emptyset \Longleftrightarrow dom(F) = \emptyset$, en cuyo cado F es convexa y comple con **(2.1)**. Por lo tanto, asumiremos que $dom(F) \neq \emptyset$ y tomamos $(x,\xi) \in epi(F)$ y*

$\alpha \in]0,1[$. *"En primer lugar, supongamos que* F *es convexa, entonces* $\alpha(x,\xi) + (1-\alpha)(y,\eta) \in epi(F)$, *y por ello*

$$F(\alpha x + (1-\alpha)y) \leq \alpha\xi + (1-\alpha)\eta \tag{2.2}$$

Siendo $\xi \downarrow F(x)$ *y* $\eta \downarrow F(y)$ *en* ***(2.2)***, *obtenemos* ***(2.1)***. *Supongamos ahora que* F *satisface* ***(2.1)***, *entonces*

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y) \leq \alpha\xi + (1-\alpha)\eta$$

y por ello $\alpha(x,\xi) + (1-\alpha)(y,\eta) \in epi(F)$*".* ■

Definición 3.2.2 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia. Decimos que* F *es estrictamente convexa si*

$$(\forall x,y \in dom(F))\ (\forall \alpha \in]0,1[)$$

$$x \neq y \implies F(\alpha x + (1-\alpha)y) < \alpha F(x) + (1-\alpha)F(y) \tag{2.3}$$

Ahora sea C *un subconjunto no vacío de* $dom(F)$*". Decimos que* F *es Convexa en* C *si*

$$(\forall x,y \in C)\ (\forall \alpha \in]0,1[)$$

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y) \tag{2.4}$$

y F *es estrictamente Convexa en* C *si*

$$(\forall x,y \in C)\ (\forall \alpha \in]0,1[)$$

$$x \neq y \implies F(\alpha x + (1-\alpha)y) < \alpha F(x) + (1-\alpha)F(y) \tag{2.5}$$

Proposición 3.2.3 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia. Decimos que* F *es Convexa si y sólo si, para todas las familias finitas* $(\alpha_i)_{i\in I}$ *en* $]0,1[$ *tales que* $\sum_{i\in I} \alpha_i = 1$ *y*

$(x_i)_{i\in I}$ *en* $dom(F)$*", tenemos*

$$F\left(\sum_{i\in I}\alpha_i x_i\right) \leq \sum_{i\in I}\alpha_i F(x_i) \tag{2.6}$$

Prueba 3.2.2 *Supongamos que* F *es Convexa y fijemos familias finitas* $(x_i)_{i\in I}$ *en* $dom(F)$ *y* $(\alpha_i)_{i\in I}$ *en* $]0,1[$ *tales que* $\sum_{i\in I}\alpha_i = 1$*. Entonces* $(x_i, F(x_i))$ *está en* $epi(F)$*. Por tanto, por convexidad,* $\left(\sum_{i\in I}\alpha_i x_i, \sum_{i\in I}\alpha_i F(x_i)\right) \in conv\,(epi(F)) = epi(F)$*, lo que da* ***(2.6)****. La implicación inversa se deduce inmediatamente de la* ***proposición (2.1)****.* ■

Corolario 3.2.1 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R}\cup\{+\infty\}$ *una función propia", entonces lo siguiente es equivalente:*

1. F *es Convexa.*
2. *Para todas las familias finitas* $(\alpha_i)_{i\in I}$ *en* $]0,1[$ *tales que* $\sum_{i\in I}\alpha_i = 1$ *y* $(x_i)_{i\in I}$ *en* $dom(F)$*, se cumple*

$$F\left(\sum_{i\in I}\alpha_i x_i\right) \leq \sum_{i\in I}\alpha_i F(x_i)$$

3. *Para todo* $x, y \in \mathcal{H}$ *y para todo* $\alpha \in]0,1[$*, se cumple:*

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y).$$

Proposición 3.2.4 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R}\cup\{+\infty\}$ *una función propia", entonces* F *es estrictamente Convexa si y sólo si para todas las familias finitas* $(\alpha_i)_{i\in I}$ *en* $]0,1[$ *tales que* $\sum_{i\in I}\alpha_i = 1$ *y* $(x_i)_{i\in I}$ *en* $dom(F)$*, tenemos* $F\left(\sum_{i\in I}\alpha_i x_i\right) \leq \sum_{i\in I}\alpha_i F(x_i)$*, y la igualdad se cumple si y sólo si* $\{x_i\}_{i\in I}$ *es un singleton.*

3.3 Función Cuasiconvexa

Definición 3.3.1 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia. Decimos que* F *es Cuasiconvexa si, para todo* $\xi \in \mathbb{R}$*, el conjunto de nivel inferior* $\underset{\leq \xi}{\text{Lev}}(F) = \{x \in \mathcal{H} \mid F(x) \leq \xi\}$ *es Convexo".*

Proposición 3.3.1 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia"*

1. *Si* F *es Convexa, entonces* F *es Cuasiconvexa.*

2. *Si* F *es Creciente o Decreciente, entonces* F *es Cuasiconvexa.*

3. *Si* F *es Cuasiconvexa y sea* $\eta \in \mathbb{R}$*, entonces* $\min\{F, \eta\}$ *es Cuasiconvexa.*

Proposición 3.3.2 *Sea* $(F_i)_{i \in I}$ *una familia de funciones Cuasiconvexas de* $\mathcal{H}$ *a* $\mathbb{R} \cup \{+\infty\}$*. Entonces* $\sup_{i \in I}(F_i)$ *es Cuasiconvexa.*

Proposición 3.3.3 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia, entonces* F *es Cuasiconvexa si y solo si"*

$$(\forall x,y \in dom(F))\ (\forall \alpha \in\ ]0,1[)$$

$$F(\alpha x + (1-\alpha)y) \leq \max\{F(x), F(y)\} \tag{3.7}$$

Proposición 3.3.4 *"Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia", entonces* F *es*

1. *Estrictamente Cuasiconvexa si*

$$(\forall x,y \in dom(F))\ (\forall \alpha \in\ ]0,1[)$$

$$x \neq y \implies F(\alpha x + (1-\alpha)y) < \max\{F(x), F(y)\} \tag{3.8}$$

2. *Uniformemente Cuasiconvexa con módulo* $\phi : \mathbb{R}_+ \longrightarrow [0, +\infty]$, *si* ϕ *es creciente,* ϕ *desaparece sólo en 0, y*

$$(\forall x,y \in dom(F))\ (\forall \alpha \in\ ]0,1[)$$

$$F(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\phi\left(\parallel x - y \parallel\right) < \text{máx}\left\{F(x), F(y)\right\} \tag{3.9}$$

Lema 3.3.1 *Sea* $F : \mathbb{R} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *sea propia, y estrictamente creciente o estrictamente decreciente en domF. Entonces se deduce de* ***(3.8)*** *que F es estrictamente cuasiconvexa.*

Prueba 3.3.1 *Para mostrar que F es estrictamente cuasiconvexa, tomemos* $x, y \in \text{dom}(F)$ *con* $x \neq y$ *y consideremos* $\alpha \in]0,1[$. *El punto* $z = \alpha x + (1-\alpha)y$ *cumple* $x < z < y$

- ***Primero*** $(z < y)$***:*** *Como* $0 < \alpha < 1$ *y* $x < y$, *multiplicando ambos lados de* $x < y$ *por* α *y luego sumar* $(1-\alpha)y$, *obtenemos:*

$$\alpha x + (1-\alpha)y < \alpha y + (1-\alpha)y$$

$$\alpha x + (1-\alpha)y < y$$

$$z < y$$

- ***Segundo*** $(x < z)$***:*** *Como* $0 < \alpha < 1$ *y* $x < y$, *multiplicando ambos lados de* $x < y$ *por* $(1-\alpha)$ *y luego sumar* αx, *obtenemos*

$$(1-\alpha)x + \alpha x < (1-\alpha)y + \alpha x$$

$$x < \alpha x + (1-\alpha)y$$

$$x < z$$

De esta manera, hemos demostrado que $x < z < y$, *es decir, z es un punto estrictamente entre x y y. Sin pérdida de generalidad, tenemos que* $x < y$. *Dado que F es estrictamente creciente,*

tenemos:

$$F(x) < F(z) < F(y)$$

$$F(x) < F(\alpha x + (1-\alpha)y) < F(y)$$

por lo tanto, concluimos

$$F(\alpha x + (1-\alpha)y) < \text{máx}\{F(x), F(y)\} = F(y)$$

lo que prueba que F es estrictamente cuasiconvexa. Un razonamiento similar se aplica si F es estrictamente decreciente en $\text{dom}(F)$. ■

Proposición 3.3.5 (Convexidad Implica Cuasiconvexidad) *Vamos a demostrar por que toda función Convexa es también Cuasiconvexa.*

Prueba 3.3.2 *Tomemos F como una función convexa. Por la definición de convexidad, sabemos lo siguiente*

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y)$$

Ahora, consideremos los dos casos posibles para los valores de $F(x)$ y $F(y)$:
Caso 1: $F(x) \leq F(y)$

- *Multiplicando por α, tenemos $\alpha F(x) \leq \alpha F(y)$ y también $1-\alpha > 0$, formando la combinación convexa $\alpha F(x) + (1-\alpha)F(y)$*

$$\alpha F(x) + (1-\alpha)F(y) \leq \alpha F(y) + (1-\alpha)F(y)$$

$$\alpha F(x) + (1-\alpha)F(y) \leq F(y)$$

- *Por lo tanto:*

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y) \leq F(y) = \text{máx}\{F(x), F(y)\}$$

Caso 2: $F(x) > F(y)$***:***

- *Análogamente al caso anterior, tenemos*

$$\alpha F(x) + (1-\alpha)F(y) \leq \alpha F(x) + (1-\alpha)F(x)$$
$$\alpha F(x) + (1-\alpha)F(y) \leq F(x)$$

- *Por lo tanto:*

$$F(\alpha x + (1-\alpha)y) \leq \alpha F(x) + (1-\alpha)F(y) \leq F(x) = \text{máx}\{F(x), F(y)\}$$

En ambos casos, hemos demostrado que:

$$F(\alpha x + (1-\alpha)y) \leq \text{máx}\{F(x), F(y)\}$$

Esto es exactamente la definición de cuasiconvexidad, por lo que F *es cuasiconvexa.* ■

La convexidad es una propiedad más fuerte que la cuasiconvexidad, y cualquier función convexa será también cuasiconvexa porque la condición de convexidad impone que los valores intermedios entre dos puntos en el dominio de la función no superen los valores en los extremos.

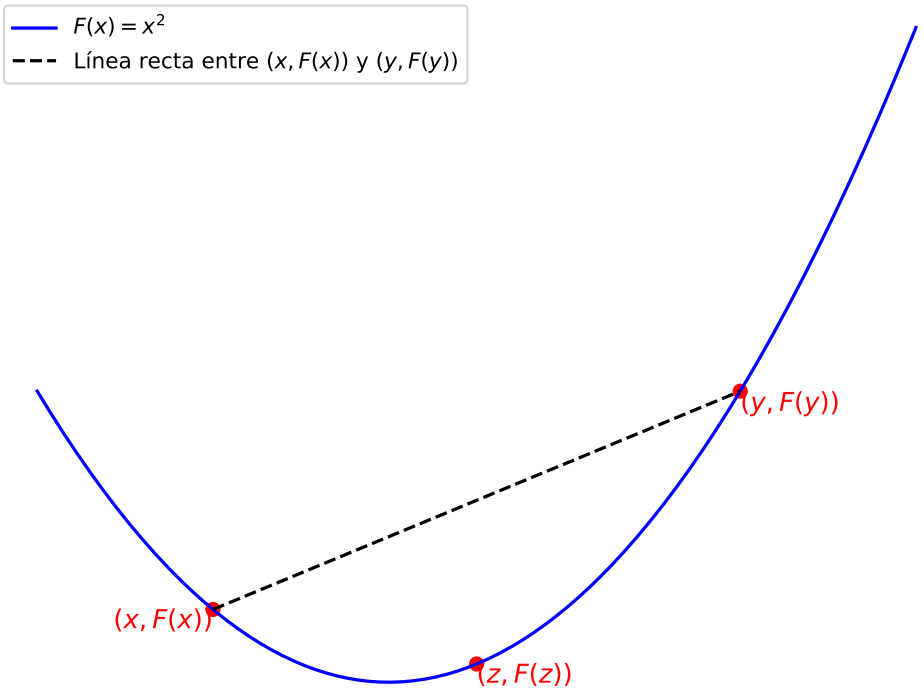


Figura 3.1: (Ejemplo) "La función $F(x) = x^2$ es convexa. Para cualquier $x, y \in \mathbb{R}$ y $\alpha \in [0, 1]$, entonces $F(\alpha x + (1 - \alpha)y) \leq \alpha F(x) + (1 - \alpha)F(y)$. Debido a que $F(x) = x^2$ es convexa, también es cuasiconvexa".

Proposición 3.3.6 (Cuasiconvexidad NO Implica Convexidad) *Un ejemplo clásico de una función cuasiconvexa que no es convexa es la función:*

$$F(x) = \begin{cases} x^2, & si\ x \leq 0, \\ 1, & si\ x > 0. \end{cases}$$

Esta función no es convexa, ya que no cumple la desigualdad de convexidad. Por ejemplo, al tomar $x_1 = 0$ y $x_2 = 1$, el punto medio $x = \frac{1}{2}$ da $F\left(\frac{1}{2}\right) = 1$, mientras que la combinación convexa de los valores $F(x_1)$ y $F(x_2)$ es $\frac{1}{2}$, y claramente $1 \not\leq \frac{1}{2}$.

Sin embargo, la función sí es cuasiconvexa, ya que F puede satisfacer la desigualdad ***(3.7)*** $F(\lambda x_1 + (1-\lambda)x_2) \leq \text{máx}(F(x_1), F(x_2))$ *para cualquier* x_1, x_2. *Esto se puede verificar en el mismo ejemplo, donde* $F\left(\frac{1}{2}\right) = 1 \leq \text{máx}\{0, 1\} = 1$.

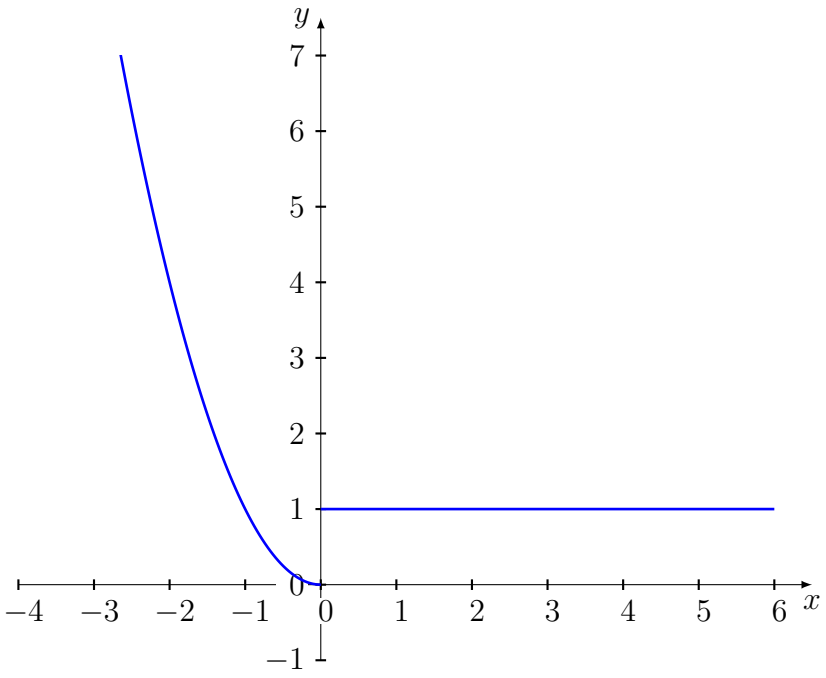


Figura 3.2: Gráfica de $F(x) = x^2$ para $x \leq 0$ y $F(x) = 1$ para $x > 0$

[Ver: Papa Quiroz (2019) para más detalles]

Proposición 3.3.7 *Sea* $F : \mathcal{H} \longrightarrow \mathbb{R}$ *y sea* $\phi : ran\ F \longrightarrow \mathbb{R}$ *creciente. Entonces se cumple lo siguiente:*

1. *Supongamos que F es estrictamente cuasiconvexa y que ϕ es estrictamente creciente. Entonces $\phi \circ F$ es estrictamente cuasiconvexa.*

2. *Supongamos que $\phi \circ F$ es estrictamente convexa. Entonces F es estrictamente cuasiconvexa.*

Prueba 3.3.3 *Supongamos que x e y son puntos distintos en $\mathcal{H} = domF = dom\,(\phi \circ F)$, y que $\alpha \in]0, 1[$.*

1. *"Teniendo en cuenta* ***(3.8)****,* $F(\alpha x + (1-\alpha)y) < \max\{F(x), F(y)\}$*, por lo tanto*

$$\begin{aligned}(\phi \circ F)(\alpha x + (1-\alpha)y) &= \phi\left(F(\alpha x + (1-\alpha)y)\right) \\ &< \phi\left(\max\{F(x), F(y)\}\right) = \max\{(\phi \circ F)(x), (\phi \circ F)(y)\}\end{aligned}$$

entonces", tenemos

$$(\phi \circ F)(\alpha x + (1-\alpha)y) < \max\{(\phi \circ F)(x), (\phi \circ F)(y)\} \tag{3.10}$$

2. *"De* ***(2.3)*** *se deduce que*

$$\begin{aligned}\phi\left(F(\alpha x + (1-\alpha)y)\right) &= (\phi \circ F)(\alpha x + (1-\alpha)y) \\ &< \alpha(\phi \circ F)(x) + (1-\alpha)(\phi \circ F)(y) \\ &\leq \max\{(\phi \circ F)(x), (\phi \circ F)(y)\} \\ &= \phi\left(\max\{F(x), F(y)\}\right)\end{aligned}$$

por lo tanto", obtenemos

$$F(\alpha x + (1-\alpha)y) < \max\{F(x), F(y)\} \tag{3.11}$$

Así quedan demostrados ambos items. ■

3.4 Subdiferencial de Fréchet

Definición 3.4.1 *"Sea* $F : C \subset \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *una función propia. Para cada* $x \in dom(F)$*, el conjunto de subgradientes regulares (también llamados subdiferenciales de Fréchet)*

de F en x, denotados por $\widehat{\partial}F(x)$ es el conjunto de vectores $v \in C \subset \mathcal{H}$, tal que

$$F(y) \geq F(x) + \langle v, y - x \rangle + O\left(\|y - x\|\right) \tag{4.12}$$

donde

$$\lim_{\substack{y \to x \\ y \neq x}} \frac{O\left(\|y - x\|\right)}{\|y - x\|} = 0$$

O equivalentemente,

$$\widehat{\partial}F(x) = \left\{ v \in \mathbb{R}^n \;\middle|\; \liminf_{\substack{y \to x \\ y \neq x}} \frac{F(y) - F(x) - \langle v, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

Si $x \neq dom(F)$ entonces $\widehat{\partial}F(x) = \emptyset$".

3.5 Regla Generalización de Fermat para el subdiferencial de Fréchet

Definición 3.5.1 *"Sea una función propia $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ tiene un mínimo local en el punto $\overline{x} \in dom(F)$", entonces*

$$0 \in \widehat{\partial}F(\overline{x})$$

Prueba 3.5.1 *Como F tiene un mínimo local en $\overline{x}$, entonces existe $\varepsilon_{\overline{x}} > 0$, tal que*

$$F(y) \geq F(\overline{x}) \;, \quad \forall y \in B(\overline{x}, \varepsilon_{\overline{x}}) \tag{5.13}$$

donde $B(\overline{x}, \varepsilon_{\overline{x}}) = \{y \in \mathcal{H} \mid \; \|y - \overline{x}\| < \varepsilon_{\overline{x}}\}$

*Definiendo $\theta\left(\| \, y - \overline{x} \, \|\right) = 0$, en la expresión **(5.13)** anterior tenemos que*

$$F(y) \geq F(\overline{x}) + \langle 0, y - \overline{x} \rangle + \theta\left(\|y - \overline{x}\|\right) \;, \quad \forall y \in B(\overline{x}, \varepsilon_{\overline{x}})$$

tal que

$$\lim_{\substack{y\to x\\ y\neq x}} \frac{\theta\left(\|y-\overline{x}\|\right)}{\|y-\overline{x}\|}=0$$

Entonces por las condiciones del Subdiferencial de Fréchet se puede concluir que $0\in\widehat{\partial}F(\overline{x})$.

■

3.6 Teorema de Moreau–Rockafellar para el subdiferencial de Fréchet

Teorema 3.6.1 *Sean* $F,G:\mathcal{H}\to\mathbb{R}\cup\{+\infty\}$ *funciones propias, convexas y semicontinuas inferiores. Entonces, para cada* $x\in\mathcal{H}$*, se cumple:*

$$\widehat{\partial}(F+G)(x)\subset\widehat{\partial}F(x)+\widehat{\partial}G(x)$$

Prueba 3.6.1 *"Sea* $v\in\widehat{\partial}(F+G)(x)$*. Nuestro objetivo es demostrar que existen* $v_F\in\widehat{\partial}F(x)$ *y* $v_G\in\widehat{\partial}G(x)$ *tales que* $v=v_F+v_G$*.*
Por la definición del subdiferencial de Fréchet, tenemos:

$$\liminf_{\substack{y\to x\\ y\neq x}}\frac{(F+G)(y)-(F+G)(x)-\langle v,y-x\rangle}{\|y-x\|}\geq 0$$

esto implica que, para todo $\varepsilon>0$*, existe* $\delta>0$ *tal que, si* $0<\|y-x\|<\delta$*, entonces:*

$$(F+G)(y)\geq(F+G)(x)+\langle v,y-x\rangle-\varepsilon\|y-x\| \tag{1}$$

como G *es convexa y semicontinua inferior, su subdiferencial de Fréchet en* x *no es vacío. Sea* $v_G\in\widehat{\partial}G(x)$*. Entonces, por definición, para todo* $\varepsilon>0$*, existe* $\delta'>0$ *tal que, si* $0<\|y-x\|<\delta'$*,*

entonces:

$$G(y) \geq G(x) + \langle v_G, y - x\rangle - \varepsilon\|y - x\| \tag{2}$$

elegimos $\delta'' = \min\{\delta, \delta'\}$ *y consideramos y tal que* $0 < \|y - x\| < \delta''$. *Restando* ***(2)*** *de* ***(1)****, obtenemos:*

$$F(y) + G(y) - G(y) \geq F(x) + G(x) - G(x) + \langle v, y - x\rangle - \langle v_G, y - x\rangle - 2\varepsilon\|y - x\|$$

lo que simplifica a:

$$F(y) \geq F(x) + \langle v - v_G, y - x\rangle - 2\varepsilon\|y - x\|. \tag{3}$$

definimos $v_F = v - v_G$. *La desigualdad* ***(3)*** *muestra que, para todo* $\varepsilon > 0$ *y y cercano a x:*

$$F(y) \geq F(x) + \langle v_F, y - x\rangle - 2\varepsilon\|y - x\|$$

dividiendo ambos lados por $\|y - x\|$ *y tomando el* $\liminf$ *cuando* $y \to x$, *obtenemos:*

$$\liminf_{\substack{y\to x\\ y\neq x}} \frac{F(y) - F(x) - \langle v_F, y - x\rangle}{\|y - x\|} \geq -2\varepsilon$$

dado que $\varepsilon > 0$ *es arbitrario, al hacer* $\varepsilon \to 0$, *se tiene:*

$$\liminf_{\substack{y\to x\\ y\neq x}} \frac{F(y) - F(x) - \langle v_F, y - x\rangle}{\|y - x\|} \geq 0,$$

lo que implica que $v_F \in \widehat{\partial} F(x)$.

Por lo tanto, hemos encontrado $v_F \in \widehat{\partial} F(x)$ *y* $v_G \in \widehat{\partial} G(x)$ *tales que* $v = v_F + v_G$. *Esto demuestra que:*

$$v \in \widehat{\partial} F(x) + \widehat{\partial} G(x).$$

Como $v \in \widehat{\partial}(F+G)(x)$ fue elegido arbitrariamente, concluimos que:

$$\widehat{\partial}(F+G)(x) \subset \widehat{\partial}F(x) + \widehat{\partial}G(x).$$

Esto concluye la demostración del Teorema". ■

3.7 Subdiferencial Limitante

Definición 3.7.1 (Subdiferencial de Funciones No Convexas) *"El conjunto de subgradientes generales (también llamados subdiferenciales limitantes) de F en $x \in C \subset \mathcal{H}$, denotado por $\partial^{Lim}F(x)$", se define como:*

$$\partial^{Lim}F(x) = \left\{ v \in C \mid \exists x^l \to x, F(x^l) \to F(x), v^l \in \widehat{\partial}F(x^l) \text{ y } v^l \to v \right\}$$

Proposición 3.7.1 *"Para una función $F : C \subset \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ y un punto $\overline{x} \in dom(F)$ en el cuál F es finito, los conjuntos subgradientes $\partial^{Lim}F(\overline{x})$ y $\widehat{\partial}F(\overline{x})$ son cerrados, con $\widehat{\partial}F(\overline{x})$ convexo y $\widehat{\partial}F(\overline{x}) \subset \partial^{Lim}F(\overline{x})$". (Obtenido de Cruzado Acuña (2018))*

Prueba 3.7.1 *Para la prueba de esta proposición necesitamos lo siguiente*

Observación 3.7.1 *"Un conjunto Z es cerrado, si y solo si, para todo $\{z^l\} \subset Z$ tal que $z^l \longrightarrow \overline{z}$ entonces $\overline{z} \in Z$". (Ver (Lages Lima (1981)))*

1. *Probaremos que $\widehat{\partial}F(\overline{x})$ es convexa. En efecto, sea $v^1, v^2 \in \widehat{\partial}F(\overline{x})$, $\forall y \in \mathbb{R}^n$, tenemos*

$$F(y) \geq F(\overline{x}) + \langle v^1, y - \overline{x} \rangle + \theta_1 \left(\|y - \overline{x}\| \right)$$

$$F(y) \geq F(\overline{x}) + \langle v^2, y - \overline{x} \rangle + \theta_2 \left(\|y - \overline{x}\| \right)$$

Entonces, para todo $\alpha \in [0,1]$ *se tiene que*

$$\alpha F(y) \geq \alpha F(\overline{x}) + \langle \alpha v^1, y - \overline{x} \rangle + \alpha\theta_1 (\|y - \overline{x}\|) \tag{7.14}$$

$$(1-\alpha) F(y) \geq (1-\alpha) F(\overline{x}) + \langle (1-\alpha) v^2, y - \overline{x} \rangle + (1-\alpha)\theta_2 (\|y - \overline{x}\|) \tag{7.15}$$

Sumando las desigualdades ***(7.14)*** *y* ***(7.15)****, obtenemos*

$$F(y) \geq F(\overline{x}) + \langle \alpha v^1 + (1-\alpha) v^2, y - \overline{x} \rangle + \theta (\|y - \overline{x}\|); \quad \forall y \in \mathbb{R}^n \tag{7.16}$$

Donde $\theta (\|y - \overline{x}\|) = \alpha\theta_1 (\|y - \overline{x}\|) + (1-\alpha)\theta_2 (\|y - \overline{x}\|)$*, dado que*

$$\lim_{\substack{y \to x \\ y \neq x}} \frac{\theta (\|y - x\|)}{\|y - x\|} = 0$$

Por ***(7.16)*** *tenemos* $\alpha v^1 + (1-\alpha) v^2 \in \widehat{\partial} f(\overline{x})$, $\forall \alpha \in [0,1]$.

2. *"Probaremos que* $\widehat{\partial} F(\overline{x})$ *es cerrado. Sea* $\left\{v^k\right\}$ *una sucesión en* $\widehat{\partial} F(\overline{x})$ *tal que* $v^k \longrightarrow v$*. Como* $v^k \in \widehat{\partial} F(\overline{x})$*, entonces*

$$F(y) \geq F(\overline{x}) + \langle v^k, y - \overline{x} \rangle + \theta (\|y - \overline{x}\|) \ , \quad \forall y \in \mathbb{R}^n$$

Calculando Límites cuando k *tiende al infinito, y teniendo en cuenta la continuidad del producto interno* $\langle \cdot, \cdot \rangle$ *tenemos*

$$F(y) \geq F(\overline{x}) + \langle v, y - \overline{x} \rangle + \theta (\|y - \overline{x}\|) \ , \quad \forall y \in \mathbb{R}^n$$

Luego, tenemos que $v \in \widehat{\partial} F(\overline{x})$*".*

3. *Ahora, "probaremos que* $\partial^{Lim} F(\overline{x})$ *es cerrado. Sea* $\left\{v^k\right\}$ *una sucesión en* $\partial^{Lim} F(\overline{x})$ *tal que* $v^k \longrightarrow v$*. Como* $v^k \in \partial^{Lim} F(\overline{x})$*, entonces:*

i) Existe una sucesión $\left\{x^{l_k}\right\}$ *tal que* $x^{l_k} \longrightarrow \overline{x}$

ii) También existe $v^{l_k} \in \widehat{\partial} F(\overline{x})$ *tal que* $v^{l_k} \longrightarrow v^k$

Ya que $v^{l_k} \in \widehat{\partial} F(\overline{x})$*, entonces tenemos*

$$F(y) \geq F(x^{l_k}) + \langle v^{l_k}, y - x^{l_k} \rangle + \theta\left(\|y - x^{l_k}\|\right) \ , \quad \forall y \in \mathbb{R}^n$$

haciendo que $l_k \longrightarrow +\infty$ *tenemos lo siguiente*

$$F(y) \geq F(\overline{x}) + \langle v^k, y - \overline{x} \rangle + \theta\left(\|y - \overline{x}\|\right) \ , \quad \forall y \in \mathbb{R}^n$$

Lo cual implica que $v^k \in \widehat{\partial} F(\overline{x})$*. Haciendo* $\left\{x^k\right\} = \overline{x}$ *tenemos* $x^k \longrightarrow \overline{x}$*, y teniendo en cuenta que* $v^k \in \widehat{\partial} F(x^k)$ *converge a* v*, se tiene que* $v \in \partial^{Lim} F(\overline{x})$*".*

4. *"Finalmente probaremos que* $\widehat{\partial} F(\overline{x}) \subset \partial^{Lim} F(\overline{x})$*. Sea* $v \in \widehat{\partial} F(\overline{x})$*, entonces tenemos*

$$F(y) \geq F(\overline{x}) + \langle v, y - \overline{x} \rangle + \theta\left(\|y - \overline{x}\|\right) \ , \quad \forall y \in \mathbb{R}^n$$

Luego, considerando sucesiones constantes $\overline{x} = \left\{x^k\right\}$ *y* $v = \left\{v^k\right\}$*, tenemos* $x^k \xrightarrow{F} \overline{x}$ *y* $v^k \longrightarrow v$*, entonces se tiene que* $v \in \partial^{Lim} F(\overline{x})$*".*

Asi queda demostrado la proposición. ■

3.8 Derivada direccional de Clarke

Para la definición de la Derivada Direccional de Clarke usaré (Fan et al. (2003))

Definición 3.8.1 *"Sea* F *una función localmente Lipschitz en* $x \in C \subset \mathcal{H}$ *con constante* $L > 0$ *y* $v \in C$*. La derivada direccional de Clarke de* F *en* x *en la dirección* v*, denotada por*

$F^{\circ}(x;v)$", se define como *

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

Proposición 3.8.1 *"Sean F,G* $: \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *funciones propias localmente Lipschitz en* $x \in dom(F) \cap dom(G)$*". Entonces, para todo* $d \in \mathbb{R}^n$*:*

1. *"*$(F+G)^{\circ}(x;d) \leq F^{\circ}(x;d) + G^{\circ}(x;d)$*"*
2. *"*$(\lambda F)^{\circ}(x;d) = \lambda F^{\circ}(x;d)\ , \quad \forall \lambda \geq 0$*"*
3. *"*$F^{\circ}(x;\lambda d) = \lambda F^{\circ}(x;d)\ , \quad \forall \lambda \geq 0$*"*

Prueba 3.8.1 *Resolveremos cada uno de los items*

1. *"Por definición de la Derivada direccional de Clarke*

$$(F+G)^{\circ}(x;d) = \lim_{\substack{y \to x \\ t \downarrow 0}} \sup \frac{(F+G)(y+td) - (F+G)(y)}{t}$$

$$(F+G)^{\circ}(x;d) = \inf_{\delta > 0} \sup_{y \in B(x,\delta)} \frac{(F+G)(y+td) - (F+G)(y)}{t}$$

$$(F+G)^{\circ}(x;d) = \inf_{\delta > 0} \sup_{y \in B(x,\delta)} \left(\frac{F(y+td) - F(y)}{t} + \frac{G(y+td) - G(y)}{t} \right)$$

Como $\sup(\varphi_1 + \varphi_2 + \cdots)(\cdot) \leq \sup(\varphi_1)(\cdot) + (\varphi_2)(\cdot) + \cdots$*, entonces:*

$$(F+G)^{\circ}(x;d) \leq \inf_{\delta > 0} \left(\sup_{y \in B(x,\delta)} \frac{F(y+td) - F(y)}{t} + \sup_{y \in B(x,\delta)} \frac{G(y+td) - G(y)}{t} \right)$$

$$(F+G)^{\circ}(x;d) \leq \inf_{\delta > 0} \sup_{y \in B(x,\delta)} \frac{F(y+td) - F(y)}{t} + \inf_{\delta > 0} \sup_{y \in B(x,\delta)} \frac{G(y+td) - G(y)}{t}$$

*En $t \downarrow 0$ tenemos que t esta en $]0, +\infty[$ y convege a 0.

$$(F+G)^{\circ}(x;d) \leq \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{F(y+td)-F(y)}{t} + \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{G(y+td)-G(y)}{t}$$

nuevamente por la definición de la Derivada direccional de Clarke"

$$(F+G)^{\circ}(x;d) \leq F^{\circ}(x;d) + G^{\circ}(x;d)$$

2. *"Por definición de la Derivada direccional de Clarke*

$$(\lambda F)^{\circ}(x;d) = \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{(\lambda F)(y+td)-(\lambda F)(y)}{t}$$

$$(\lambda F)^{\circ}(x;d) = \lambda \cdot \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{F(y+td)-F(y)}{t}$$

nuevamente por la definición de la Derivada direccional de Clarke"

$$(\lambda F)^{\circ}(x;d) = \lambda \cdot F^{\circ}(x;d)$$

3. *"Si $\lambda = 0$, la igualdad se verifica trivialmente, pero si $\lambda > 0$, por la definición de la Derivada direccional de Clarke, tenemos*

$$\lambda F^{\circ}(x;d) = \lambda \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{F(y+td)-F(y)}{t}$$

$$\lambda F^{\circ}(x;d) = \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{F\left(y+\dfrac{t}{\lambda}(\lambda d)\right)-F(y)}{\dfrac{t}{\lambda}}$$

$$\lambda F^{\circ}(x;d) = \limsup_{\substack{y \to x \\ \frac{t}{\lambda} \downarrow 0}} \frac{F\left(y + \frac{t}{\lambda}(\lambda d)\right) - F(y)}{\frac{t}{\lambda}}$$

nuevamente por la definición de la Derivada direccional de Clarke"

$$\lambda F^{\circ}(x;d) = F^{\circ}(x;\lambda d)$$

Así queda demostrada la Proposición. ■

3.9 Subdiferencial de Clarke

Definición 3.9.1 *"El subdiferencial de Clarke de F en x, denotada por $\partial^{\circ}F(x)$", se define como*

$$\partial^{\circ}F(x) = \{w \in C \mid F^{\circ}(x;v) \geq \langle w,v\rangle,\ v \in C\}$$

Proposición 3.9.1 *Si F es Localmente Lipschitziana en x, entonces para todo $\lambda \in \mathbb{R}$, se cumple*

$$\partial^{\circ}\left(\lambda F(x)\right) = \lambda \partial^{\circ}F(x)$$

Prueba 3.9.1 *Como F es de Lipschitz, λF también es de Lipschitz en x. Si $\lambda \geq 0$, entonces $(\lambda F)^{\circ} = \lambda F^{\circ}$.*
También $\partial^{\circ}\left(\lambda F(x)\right) = \lambda \partial^{\circ}F(x)$, "esto se prueba para $\lambda = -1$

$$d \in \partial^{\circ}\left(-F(x)\right) \Longleftrightarrow (-F)^{\circ}(x;v) \geq \langle d,v\rangle\ ,\quad \forall v \in \mathbb{R}^n$$

$$d \in \partial^{\circ}\left(-F(x)\right) \Longleftrightarrow F^{\circ}(x;-v) \geq \langle d,v\rangle\ ,\quad \forall v \in \mathbb{R}^n$$

$$d \in \partial^{\circ}\left(-F(x)\right) \Longleftrightarrow F^{\circ}(x;-v) \geq \langle -d,-v\rangle\ ,\quad \forall v \in \mathbb{R}^n$$

$$d \in \partial^\circ\left(-F(x)\right) \Longleftrightarrow -d \in \partial^\circ F(x)$$

$$d \in \partial^\circ\left(-F(x)\right) \Longleftrightarrow d \in -\partial^\circ F(x)$$

por lo tanto, $\partial^\circ\left(\lambda F(x)\right) = \lambda\partial^\circ F(x)$*".* ■

Teorema 3.9.1 (Moreau–Rockafellar para el subdiferencial de Clarke) *"Sean* $F, G : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *propias, convexas y semicontinuas inferiores. Para cada* $x \in \mathcal{H}$*", tenemos*

$$\partial^\circ(F+G)(x) \subset \partial^\circ F(x) + \partial^\circ G(x)$$

Prueba 3.9.2 *Consideremos* $\xi \in \partial^\circ(F+G)(x)$*. Por la definición del subdiferencial de Clarke, esto significa que:*

$$(F+G)^\circ(x;v) \geq \langle \xi, v\rangle \quad \textit{para todo } v \in \mathcal{H}$$

*aplicando la propiedad de subaditividad del límite superior (**(1)** de **(3.8.1)**), obtenemos:*

$$(F+G)^\circ(x;v) \leq F^\circ(x;v) + G^\circ(x;v).$$

por lo tanto, para cualquier $v \in \mathcal{H}$*, tenemos:*

$$\langle \xi, v\rangle \leq F^\circ(x;v) + G^\circ(x;v).$$

Ahora, veamos la relación con los subdiferenciales individuales:
Por definición del subdiferencial de Clarke para F *y* G*, sabemos que existen* $\xi_F \in \partial^\circ F(x)$ *y* $\xi_G \in \partial^\circ G(x)$ *tales que:*

$$F^\circ(x;v) \geq \langle \xi_F, v\rangle \quad y \quad G^\circ(x;v) \geq \langle \xi_G, v\rangle.$$

por lo tanto, podemos escribir:

$$\langle \xi, v\rangle \leq \langle \xi_F, v\rangle + \langle \xi_G, v\rangle = \langle \xi_F + \xi_G, v\rangle.$$

dado que esto es cierto para cualquier v, tenemos que:

$$\xi \in \partial^\circ F(x) + \partial^\circ G(x).$$

Por lo tanto hemos demostrado que si $\xi \in \partial^\circ(F+G)(x)$, entonces $\xi \in \partial^\circ F(x) + \partial^\circ G(x)$. Esto prueba la inclusión:

$$\partial^\circ(F+G)(x) \subset \partial^\circ F(x) + \partial^\circ G(x).$$

así queda demostrado el Teorema de Moreau–Rockafellar usando el Subdiferencial de Clarke.

■

Esta inclusión es un resultado clave en el análisis no suave, reflejando cómo la derivada direccional de Clarke y su subdiferencial permiten manejar la suma de funciones no diferenciables.

Proposición 3.9.2 *Si las funciones $F_i : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$, son localmente Lipschitz en x para $i = 1, \cdots, m$, entonces para escalares $\lambda_i \in \mathbb{R}$ tenemos*

$$\partial^\circ \left(\sum_{i=1}^{m} \lambda_i F_i \right)(x) \subset \sum_{i=1}^{m} \lambda_i \partial^\circ F_i(x)$$

Prueba 3.9.3 *Es suficiente para $m = 2$, el caso se generaliza por inducción. De la **Proposición (3.8.1) (1)** observamos que*

$$(F_1 + F_2)^\circ(x; d) \leq F_1^\circ(x; d) + F_2^\circ(x; d)$$

Por la definición de Subdiferenciabilidad

$$\partial^{\circ}\left(\lambda_1 F_1 + \lambda_2 F_2\right)(x) \subset \partial^{\circ}\left(\lambda_1 F_1(x)\right) + \partial^{\circ}\left(\lambda_2 F_2(x)\right)$$

Luego por la ***Proposición (3.9.1)*** *tenemos*

$$\partial^{\circ}\left(\lambda_1 F_1 + \lambda_2 F_2\right)(x) \subset \lambda_1 \partial^{\circ} F_1(x) + \lambda_2 \partial^{\circ} F_2(x)$$

Así queda demostrado para $m = 2$. ■

Ejemplo 3.9.1 *Calcular el Subdiferencial de Clarke de la función* $F : \mathbb{R} \longrightarrow \mathbb{R}$ *definida por*

$$F(x) = \mid ax + b \mid$$

Prueba 3.9.4 *Como ya se vio en el* ***Ejemplo 2.12.2****,* F *es localmente de Lipschitz, entonces procedemos a calcular* $\partial^{\circ} F(x)$*. Entonces, tenemos*

$$F(x) = \begin{cases} ax + b, \ si \ x \geq -\dfrac{b}{a} \\ -ax - b, \ si \ x < -\dfrac{b}{a} \end{cases}$$

Ahora consideraremos 3 *casos para* x*, en este ejemplo asumiremos que* $a > 0$*, para* $a < 0$ *es analogo.*

1. *Si* $x > -\dfrac{b}{a}$*, entonces*

$$F^{\circ}(x; v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{(a(y + tv) + b) - (ay + b)}{t}$$

$$F^{\circ}(x; v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{atv}{t}$$

$$F^{\circ}(x;v) = a.v$$

Así tenemos que $\partial^{\circ}F(x) = \{w \in \mathbb{R} \mid a.v \geq w.v \ , \ \forall v \in \mathbb{R}\}$. *Para valores de* v

i) si $v > 0$, *entonces*

$$\partial^{\circ}F(x) = \{w \in \mathbb{R} \mid a \geq w\}$$

ii) si $v < 0$, *entonces*

$$\partial^{\circ}F(x) = \{w \in \mathbb{R} \mid a \leq w\}$$

iii) si $v = 0$, *entonces*

$$w \in \mathbb{R}$$

Intersectando los datos obtenidos en ***i)***, ***ii)*** *y* ***iii)***, *obtenemos que*

$$\partial^{\circ}F(x) = \{a\}$$

2. *Si* $x < -\dfrac{b}{a}$, *entonces*

$$F^{\circ}(x;v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{(-a(y+tv) - b) - (-ay - b)}{t}$$

$$F^{\circ}(x;v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{-atv}{t}$$

$$F^{\circ}(x;v) = -a.v$$

Así tenemos que $\partial^{\circ}F(x) = \{w \in \mathbb{R} \mid \ -a.v \geq w.v \ , \ \forall v \in \mathbb{R}\}$. *Para valores de* v

i) si $v > 0$, *entonces*

$$\partial^{\circ}F(x) = \{w \in \mathbb{R} \mid \ -a \geq w\}$$

ii) si $v < 0$, entonces

$$\partial^\circ F(x) = \{w \in \mathbb{R} \mid -a \leq w\}$$

iii) si $v = 0$, entonces

$$w \in \mathbb{R}$$

*Intersectando los datos obtenidos en **i)**, **ii)** y **iii)**, obtenemos que*

$$\partial^\circ F(x) = \{-a\}$$

3. *"Si $x = -\dfrac{b}{a}$, tomando un y suficientemente cercano a x obtenemos los siguientes caso:*

   - *Si $y > -\dfrac{b}{a}$, tenemos*

   $$\frac{(a(y+tv)+b)-(ay+b)}{t} = \frac{atv}{t} = a.v$$

   - *Si $y < -\dfrac{b}{a}$, tenemos*

   $$\frac{(-a(y+tv)-b)-(-ay-b)}{t} = \frac{-atv}{t} = -a.v$$

   De los dos casos anteriores tenemos

   $$\sup \frac{F(y+tv)-F(y)}{t} = |a.v| \tag{9.17}$$

   Ahora por la definición de la Subdiferencial de Clarke

   $$\partial^\circ F\left(-\frac{b}{a}\right) = \{w \in \mathbb{R} \mid F^\circ(x;v) \geq w.v \ , \quad \forall v \in \mathbb{R}\}$$

$$\partial^\circ F\left(-\frac{b}{a}\right) = \left\{ w \in \mathbb{R} \;\middle|\; \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{F(y+tv) - F(y)}{t} \geq w.v \;,\quad \forall v \in \mathbb{R} \right\}$$

usando la ecuación ***(9.17)****", tenemos*

$$\partial^\circ F\left(-\frac{b}{a}\right) = \left\{ w \in \mathbb{R} \;\middle|\; \lim_{t \downarrow 0} |a.v| \geq w.v \;,\quad \forall v \in \mathbb{R} \right\}$$

$$\partial^\circ F\left(-\frac{b}{a}\right) = \left\{ w \in \mathbb{R} \;\middle|\; |a.v| \geq w.v \;,\quad \forall v \in \mathbb{R} \right\}$$

Para valores de v

i) si $v > 0$*, entonces*

$$\partial^\circ F\left(-\frac{b}{a}\right) = \{w \in \mathbb{R} \mid a \geq w\}$$

ii) si $v < 0$*, entonces*

$$\partial^\circ F\left(-\frac{b}{a}\right) = \{w \in \mathbb{R} \mid -a \leq w\}$$

iii) si $v = 0$*, entonces*

$$w \in \mathbb{R}$$

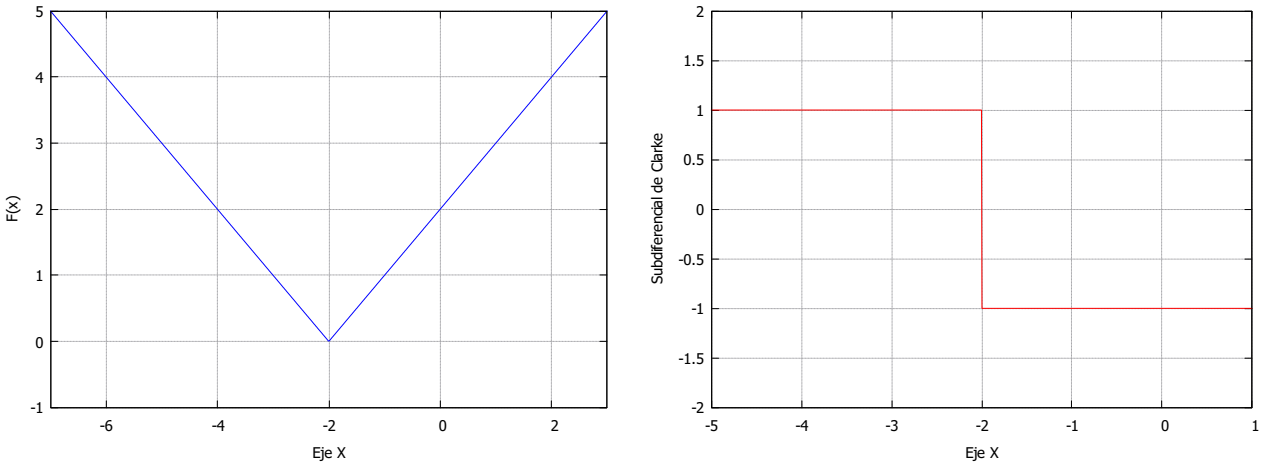
Intersectando los datos obtenidos en ***i)****,* ***ii)*** *y* ***iii)****, obtenemos que*

$$\partial^\circ F\left(-\frac{b}{a}\right) = [-a; a]$$

Concluyendo

$$\partial^\circ F(x) = \begin{cases} \{a\},\ si\ x > -\dfrac{b}{a} \\ [-a; a],\ si\ x = -\dfrac{b}{a} \\ \{-a\},\ si\ x < -\dfrac{b}{a} \end{cases}$$

Así se obtenemos $\partial^\circ F(x)$*.* ■



(a) Funcion $F(x)$ con $a = 1$ y $b = 2$ (b) Subdiferencial de Clarke $\partial^\circ F(x)$

Figura 3.3: Represención del Subdiferencial de Clarke de la función $F(x)$

Ejemplo 3.9.2 *Calcular el Subdiferencial de Clarke de la función* $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *definida por*

$$F(x_1,x_2) = \mid ax_1 + bx_2 + c \mid$$

Prueba 3.9.5 *Tenemos*

$$F(x_1,x_2) = \begin{cases} ax_1 + bx_2 + c, \ si \ \ ax_1 + bx_2 + c \geq 0 \\ -ax_1 - bx_2 - c, \ si \ \ ax_1 + bx_2 + c < 0 \end{cases}$$

Ahora consideraremos 3 *casos para* (x_1,x_2)*, en este ejemplo asumiremos que* $a > 0$ *y* $b > 0$*, para otros casos es analogo.*

1. *Si* $ax_1 + bx_2 + c > 0$*, entonces*

$$F^\circ\left((x_1,x_2);(v_1,v_2)\right) = \limsup_{\substack{(u_1,u_2)\to(x_1,x_2) \\ t\downarrow 0}} \frac{(a(u_1+tv_1)+b(u_2+tv_2)+c)-(au_1+bv_2+c)}{t}$$

$$F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) = \limsup_{\substack{(u_1,u_2)\to(x_1,x_2)\\ t\downarrow 0}} \frac{atv_1+btv_2}{t}$$

$$F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) = av_1+bv_2$$

Tenemos que

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid av_1+bv_2 \geq (w_1,w_2).(v_1,v_2)\ ,\ \forall(v_1,v_2)\in\mathbb{R}^2\right\}$$

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid av_1+bv_2 \geq w_1.v_1+w_2.v_2\ ,\ \forall(v_1,v_2)\in\mathbb{R}^2\right\}$$

Para valores de (v_1,v_2)

i) si $v_1 = 1$ *y* $v_2 = 0$, *entonces*

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid a \geq w_1\right\}$$

ii) si $v_1 = 0$ *y* $v_2 = 1$, *entonces*

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid b \geq w_2\right\}$$

iii) si $v_1 = -1$ *y* $v_2 = 0$, *entonces*

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid a \leq w_1\right\}$$

iv) si $v_1 = 0$ *y* $v_2 = -1$, *entonces*

$$\partial^{\circ}F(x_1,x_2) = \left\{(w_1,w_2)\in\mathbb{R}^2 \mid b \leq w_2\right\}$$

*Intersectando **i)** con **iii)** y **ii)** con **iv)** obtenemos que*

$$\partial^{\circ} F(x_1,x_2) = \{(a,b)\}$$

2. *Si* $ax_1 + bx_2 + c < 0$, *entonces*

$$F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) = \limsup_{\substack{(u_1,u_2)\to(x_1,x_2)\\ t\downarrow 0}} \frac{(-a(u_1+tv_1)-b(u_2+tv_2)-c)-(-au_1-bv_2-c)}{t}$$

$$F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) = \limsup_{\substack{(u_1,u_2)\to(x_1,x_2)\\ t\downarrow 0}} \frac{-atv_1-btv_2}{t}$$

$$F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) = -av_1 - bv_2$$

Tenemos que

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -av_1 - bv_2 \geq (w_1,w_2).(v_1,v_2) \ , \ \forall (v_1,v_2) \in \mathbb{R}^2\right\}$$

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -av_1 - bv_2 \geq w_1.v_1 + w_2.v_2 \ , \ \forall (v_1,v_2) \in \mathbb{R}^2\right\}$$

Para valores de (v_1,v_2)

i) si $v_1 = 1$ *y* $v_2 = 0$, *entonces*

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -a \geq w_1\right\}$$

ii) si $v_1 = 0$ *y* $v_2 = 1$, *entonces*

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -b \geq w_2\right\}$$

iii) si $v_1 = -1$ *y* $v_2 = 0$, *entonces*

$$\partial^\circ F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -a \leq w_1\right\}$$

iv) si $v_1 = 0$ *y* $v_2 = -1$, *entonces*

$$\partial^\circ F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -b \leq w_2\right\}$$

Intersectando ***i)*** *con* ***iii)*** *y* ***ii)*** *con* ***iv)*** *obtenemos que*

$$\partial^\circ F(x_1,x_2) = \{(-a,-b)\}$$

3. *Si* $ax_1 + bx_2 + c = 0$, *tomando un punto suficientemente cercano a* (x_1,x_2) *obtenemos los siguientes casos*

   - *Si* $ax_1 + bx_2 + c > 0$, *tenemos*

   $$\frac{(a(u_1 + tv_1) + b(u_2 + tv_2) + c) - (au_1 + bv_2 + c)}{t} = av_1 + bv_2$$

   - *Si* $ax_1 + bx_2 + c < 0$, *tenemos*

   $$\frac{(-a(u_1 + tv_1) - b(u_2 + tv_2) - c) - (-au_1 - bv_2 - c)}{t} = -av_1 - bv_2$$

   De los dos casos anteriores tenemos

$$\sup \frac{F\left((u_1,u_2) + t(v_1,v_2)\right) - F\left((u_1,u_2)\right)}{t} = |a.v_1 + b.v_2| \tag{9.18}$$

Ahora por la definición de la Subdiferencial de Clarke

$$\partial^{\circ} F\left((x_1,x_2)\right) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid F^{\circ}\left((x_1,x_2);(v_1,v_2)\right) \geq (w_1,w_2).(v_1,v_2) \ , \quad \forall v \in \mathbb{R}^2\right\}$$

$$\partial^{\circ} F\left((x_1,x_2)\right) = \left\{(w_1,w_2) \in \mathbb{R}^2 \ \middle| \ \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{F\left((u_1,u_2)+t(v_1,v_2)\right) - F\left((u_1,u_2)\right)}{t} \geq (w_1,w_2).(v_1,v_2) \ , \quad \forall v \in \mathbb{R}^2\right\}$$

Usando la ecuación ***(9.18)****, tenemos*

$$\partial^{\circ} F\left((x_1,x_2)\right) = \left\{(w_1,w_2) \in \mathbb{R}^2 \ \middle| \ \lim_{t \downarrow 0} |a.v_1 + b.v_2| \geq (w_1,w_2).(v_1,v_2) \ , \quad \forall v \in \mathbb{R}^2\right\}$$

$$\partial^{\circ} F\left((x_1,x_2)\right) = \left\{(w_1,w_2) \in \mathbb{R}^2 \ \middle| \ |a.v_1 + b.v_2| \geq w_1.v_1 + w_2.v_2 \ , \quad \forall v \in \mathbb{R}^2\right\}$$

Para valores de (v_1,v_2)

i) si $v_1 = 1$ *y* $v_2 = 0$*, entonces*

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid a \geq w_1\right\}$$

ii) si $v_1 = 0$ *y* $v_2 = 1$*, entonces*

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid b \geq w_2\right\}$$

iii) si $v_1 = -1$ *y* $v_2 = 0$*, entonces*

$$\partial^{\circ} F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -a \leq w_1\right\}$$

iv) si $v_1 = 0$ *y* $v_2 = -1$, *entonces*

$$\partial^\circ F(x_1,x_2) = \left\{(w_1,w_2) \in \mathbb{R}^2 \mid -b \leq w_2\right\}$$

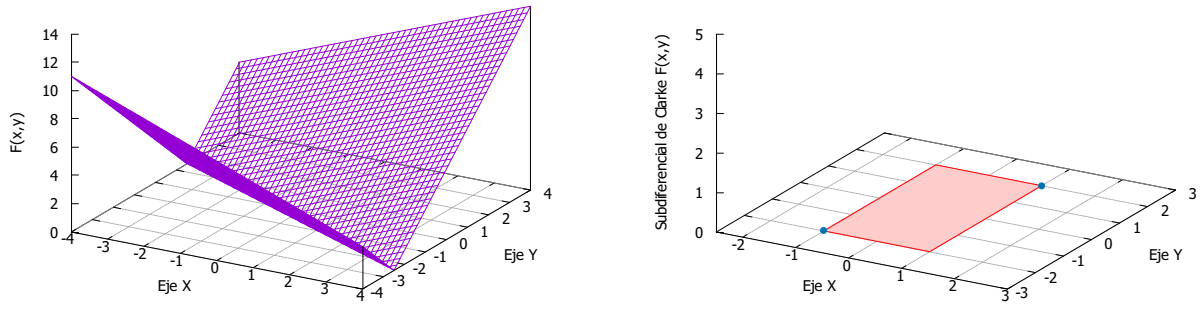
Intersectando ***i)*** *con* ***iii)*** *y* ***ii)*** *con* ***iv)*** *obtenemos que*

$$\partial^\circ F(x_1,x_2) = \{[-a,a] \times [-b,b]\}$$

Concluyendo

$$\partial^\circ F(x_1,x_2) = \begin{cases} \{(a,b)\}, \ si \ \ ax_1 + bx_2 + c > 0 \\ [-a,a] \times [-b,b], \ si \ \ ax_1 + bx_2 + c = 0 \\ \{(-a,-b)\}, \ si \ \ ax_1 + bx_2 + c < 0 \end{cases}$$

Así se obtenemos $\partial^\circ F(x_1,x_2)$. ■



(a) Funcion $F(x_1,x_2)$ con $a = 1$, $b = 2$ y $c = 1$ (b) Subdiferencial $\partial^\circ F(x_1,x_2)$

Figura 3.4: Represención del Subdiferencial de Clarke de la función $F(x_1,x_2)$

3.10 Regla Generalización de Fermat para el subdiferencial de Clarke

Teorema 3.10.1 (Regla de Fermat) *"Si $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ es localmente lipschitz en x y tiene un extremo x, es decir, x es mínimo local o máximo local", entonces*

$$0 \in \partial^{\circ} F(x)$$

Prueba 3.10.1 *"Supongamos que F alcanza su mínimo local en x, entonces $\exists \varepsilon$, tal que*

$$F(x+tv) - F(x) \geq 0 \ , \quad \forall \ t \in \ ]0,\varepsilon[\ y \ v \in \mathcal{H}$$

se tiene

$$F^{\circ}(x;v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{F(y+tv) - F(y)}{t} \geq \limsup_{t \downarrow 0} \frac{F(y+tv) - F(x)}{t} \geq 0$$

entonces obtenemos"

$$F^{\circ}(x;v) \geq 0 = \langle 0,v \rangle \ , \quad \forall \ v \in \mathcal{H}$$

Proposición 3.10.1 *"De las definiciones anteriores tenemos directamente que para todo $x \in dom(F)$", tenemos*

$$\widehat{\partial} F(x) \subset \partial^{Lim} F(x) \subset \partial^{\circ} F(x)$$

Prueba 3.10.2 *Ya se probo que $\widehat{\partial} F(x) \subset \partial^{Lim} F(x)$ en la **Proposición (3.7.1)**. Ahora lo que nos queda probar es $\partial^{Lim} F(x) \subset \partial^{\circ} F(x)$.*
Supongamos que $\xi \in \partial^{Lim} F(x)$. Entonces, por la definición del subdiferencial limitante, existen sucesiones

$$x_n \longrightarrow x \quad y \quad \xi_n \in \widehat{\partial} F(x_n)$$

tales que $\xi_n \longrightarrow \xi$. *Recordemos que si* $\xi_n \in \widehat{\partial} F(x_n)$, *entonces:*

$$\liminf_{y \longrightarrow x} \frac{F(y) - F(x) - \langle \xi_n, y - x \rangle}{\|y - x\|} \geq 0$$

Consideremos la expresión para $y = x + tv$, *donde* $t > 0$ *es pequeño:*

$$\frac{F(x + tv) - F(x) - \langle \xi_n, tv \rangle}{t} \geq \liminf_{y \longrightarrow x} \frac{F(y) - F(x) - \langle \xi_n, y - x \rangle}{\|y - x\|} \geq 0$$

por lo tanto:

$$\frac{F(x + tv) - F(x) - \langle \xi_n, tv \rangle}{t} \geq 0$$

usando álgebra elemental obtenemos

$$\frac{F(x + tv) - F(x)}{t} \geq \langle \xi_n, v \rangle$$

tomando el límite superior cuando t *tiende a 0, obtenemos:*

$$\limsup_{t \downarrow 0} \frac{F(x + tv) - F(x)}{t} \geq \langle \xi_n, v \rangle$$

"como $\xi_n \longrightarrow \xi$, *podemos escribir:*

$$F^{\circ}(x; v) \geq \limsup_{n \longrightarrow \infty} \langle \xi_n, v \rangle = \langle \xi, v \rangle$$

ya que $F^{\circ}(x; v) \geq \langle \xi, v \rangle$ *para todo* $v \in \mathcal{H}$, *concluimos que* $\xi \in \partial^{\circ} F(x)$*".* ■

La última proposición es la que refleja la jerarquía de estos subdiferenciales en el análisis no diferenciable (no suave).

3.11 Conjunto de Funciones Convexas en Espacios de Hilbert

Definición 3.11.1 *El conjunto de funciones convexas semicontinuas inferiores propias de* $\mathcal{H}$ *a* $]-\infty,+\infty]$ *es denotado por* $\Gamma_0(\mathcal{H})$.

Ejemplo 3.11.1 *Sea* $(e_i)_{i\in I}$ *una familia en* $\mathcal{H}$ *y sea* $(\phi_i)_{i\in I}$ *una familia en* $\Gamma_0(\mathbb{R})$ *tal que*

- $(\forall i \in I)\ \ \phi_i \geq \phi_i(0) = 0$
- *Conjunto de funciones de la forma*

$$F : \mathcal{H} \longrightarrow]-\infty,+\infty]$$
$$x \longmapsto \sum_{i\in I} \phi_i\left(\langle x \cdot e_i\rangle\right)$$

Entonces $F \in \Gamma_0(\mathcal{H})$.

Prueba 3.11.1 *Establecemos* $(\forall i \in I)$ *al conjunto de funciones*

$$F_i : \mathcal{H} \longrightarrow]-\infty,+\infty[$$
$$x \longmapsto \phi_i\left(\langle x \cdot e_i\rangle\right)$$

entonces, $F = \sum_{i\in I} F_i$ *y* $(\forall i \in I)\ \ 0 \leq F_i \in \Gamma_0(\mathbb{R})$, *por el* ***Corolario 9.4 (ii)*** *de (Bauschke and Combettes (2019)) se deduce que* $F \in \Gamma(\mathcal{H})$[†]. *Finalmente, dado que* $F(0) = 0$, F *es propia.* ■

Proposición 3.11.1 *Sea* $F \in \Gamma_0(\mathcal{H})$, $x \in \mathcal{H}$ *y sea* $y \in dom(F)$. *Para cada* $\alpha \in]0,1[$, *el conjunto* $x_\alpha = (1-\alpha)x + \alpha y$. *Entonces*

$$\lim_{\alpha\downarrow 0} F(x_\alpha) = F(x)$$

[†]El conjunto de funciones convexas semicontinuas inferiores de $\mathcal{H}$ a $]-\infty,+\infty]$ es denotado por $\Gamma(\mathcal{H})$.

Prueba 3.11.2 *La convexidad de F implica que, para cualquier $\alpha \in [0,1]$ y $x, y \in dom(F)$, se cumple la desigualdad de convexidad:*

$$F((1-\alpha)x + \alpha y) \leq (1-\alpha)F(x) + \alpha F(y)$$

En particular, esto implica que:

$$F(x_\alpha) = F((1-\alpha)x + \alpha y) \leq (1-\alpha)F(x) + \alpha F(y)$$

Queremos analizar el comportamiento de $F(x_\alpha)$ cuando $\alpha \downarrow 0$. Para esto, tomamos el límite superior en ambos lados de la desigualdad anterior:

$$F(x_\alpha) \leq (1-\alpha)F(x) + \alpha F(y).$$

Dado que $(1-\alpha)F(x) + \alpha F(y)$ es continua en α, evaluamos su límite cuando $\alpha \longrightarrow 0$:

$$\lim_{\alpha \downarrow 0} (1-\alpha)F(x) + \alpha F(y) = F(x)$$

ya que $\lim_{\alpha \longrightarrow 0}(1-\alpha) = 1$ y $\lim_{\alpha \longrightarrow 0} \alpha = 0$ Por lo tanto, obtenemos:

$$\limsup_{\alpha \downarrow 0} F(x_\alpha) \leq F(x)$$

La semicontinuidad inferior de F nos garantiza que:

$$F(x) \leq \liminf_{\alpha \downarrow 0} F(x_\alpha)$$

Es decir, el límite inferior de la sucesión $F(x_\alpha)$ no puede ser menor que $F(x)$. A partir de los

resultados obtenidos, tenemos:

$$F(x) \leq \liminf_{\alpha \downarrow 0} F(x_\alpha) \leq \limsup_{\alpha \downarrow 0} F(x_\alpha) \leq F(x)$$

Dado que el límite inferior y el límite superior están acotados por $F(x)$, podemos concluir que:

$$\lim_{\alpha \downarrow 0} F(x_\alpha) = F(x)$$

que es lo que queriamos demostrar. ■

4 Proyección en Espacios de Hilbert

"Para este capítulo usaré a Bauschke and Combettes (2019) y Peypouquet (2015), que consideran una propiedad importante de los espacios de Hilbert es que dado un subconjunto no vacío, cerrado y convexo C de $\mathcal{H}$ y un punto $x \notin C$, existe un único punto en C que es el más cercano a x".

4.1 Proyección

Definición 4.1.1 *"Sea C un subconjunto de $\mathcal{H}$, sea $x \in \mathcal{H}$, y sea $p \in C$. Entonces p es la mejor aproximación a x desde C (o una proyección de x sobre C)" si*

$$\| x - p \| = d_C(x)$$

Si cada punto de $\mathcal{H}$ tiene al menos una proyección sobre C, entonces C es proximal. "Si cada punto de $\mathcal{H}$ tiene exactamente una proyección sobre C, entonces C es un conjunto de Chebyshev. En este caso, el proyector (u operador de proyección) sobre C es el operador, denotado por P_C , que mapea cada punto de $\mathcal{H}$ a su única proyección sobre C".

Ejemplo 4.1.1 *Sea $\{e_i\}_{i \in I}$ un conjunto ortonormal* finito en $\mathcal{H}$, sea $V = span\{e_i\}_{i \in I}$, y sea*

*"Un conjunto es ortonormal si es un conjunto ortogonal y la norma (o módulo) de cada uno de sus elementos es uno. Esta definición sólo tiene sentido si los elementos pertenecen a un espacio en el que se define el producto interior, como es en el caso del espacio de Hilbert, donde el producto escalar se puede determinar en términos de distancia y proyección ortogonal".

$x \in \mathcal{H}$. *Entonces* V *es un conjunto de Chebyshev.*

$$P_V(x) = \sum_{i \in I} \langle x \cdot e_i \rangle e_i$$
$$d_V(x) = \sqrt{\| x \|^2 - \sum_{i \in I} \mid \langle x \cdot e_i \rangle \mid^2}$$

Prueba 4.1.1 *"Para cada familia* $(\alpha_i)_{i \in I}$*, tenemos*

$$\left\| x - \sum_{i \in I} \alpha_i e_i \right\|^2 = \|x\|^2 - 2 \left\langle x \cdot \sum_{i \in I} \alpha_i e_i \right\rangle + \left\| \sum_{i \in I} \alpha_i e_i \right\|^2$$

$$\left\| x - \sum_{i \in I} \alpha_i e_i \right\|^2 = \|x\|^2 - 2 \sum_{i \in I} \alpha_i \langle x \cdot e_i \rangle + \sum_{i \in I} \mid \alpha_i \mid^2$$
$$\left\| x - \sum_{i \in I} \alpha_i e_i \right\|^2 = \|x\|^2 - \sum_{i \in I} \mid \langle x \cdot e_i \rangle \mid^2 + \sum_{i \in I} \mid \alpha_i - \langle x \cdot e_i \rangle \mid^2$$

Por tanto, la función $(\alpha_i)_{i \in I} \longmapsto \left\| x - \sum_{i \in I} \alpha_i e_i \right\|^2$ *admite* $(\{x \cdot e_i\rangle)_{i \in I}$ *como único minimizador", y su valor mínimo es*

$$\|x\|^2 - \sum_{i \in I} \mid \langle x \cdot e_i \rangle \mid^2$$

Proposición 4.1.1 *Supongamos que* $\in \mathcal{H}$ *es finito-dimensional y que* C *es un subconjunto Chebyshev de* $\in \mathcal{H}$*. Entonces* $P_{_C}$ *es continua.*

Prueba 4.1.2 *La demostración la podemos encontrar en Bauschke and Combettes (2019)* ■

Corolario 4.1.1 *Supongamos que* $\mathcal{H}$ *es finito-dimensional. Sea* C *un subconjunto no vacío de* $\mathcal{H}$*. Entonces* C *es proximal si y sólo si es cerrado.*

Observación 4.1.1 *Todo conjunto de Chebyshev es un conjunto proximal. Sin embargo, un conjunto proximinal puede no ser un conjunto de Chebyshev.*

Una propiedad importante de los espacios de Hilbert es que dado un subconjunto no vacío, cerrado y convexo $C \in \mathcal{H}$ y un punto $x \in C$, hay un punto único en C que es el más cercano a x. Así, tenemos lo siguiente, según Peypouquet (2015):

Proposición 4.1.2 *"Sea C un subconjunto convexo, cerrado y no vacío de $\mathcal{H}$, y sea $x \in \mathcal{H}$. Entonces, existe un único punto $y^* \in C$" tal que*

$$\| x - y^* \| = \min_{y \in C} \| x - y \| \tag{1.1}$$

Además, es el único elemento de C tal que, para todo $y \in C$, se tiene

$$\langle x - y^*, y - y^* \rangle \leq 0 \tag{1.2}$$

El punto y^* es la proyección de x sobre C y se denota por $P_C(x)$.

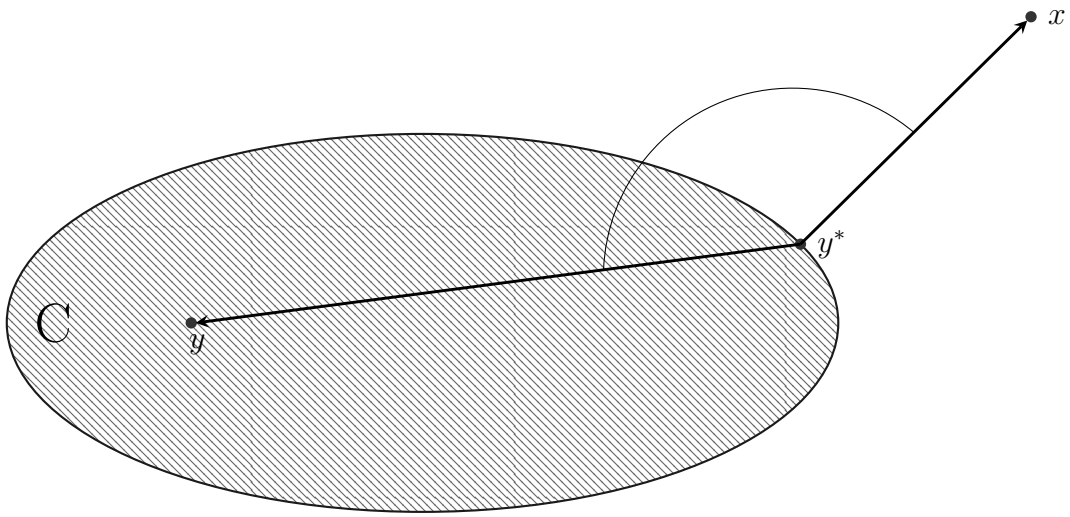


Figura 4.1: La caracterización **(1.2)** establece que $y^* \in C$ es la proyección de x sobre C si y solo si los vectores $x - y^*$ e $y - y^*$ forman un ángulo recto u obtuso para cada $y^* \in C$.

Prueba 4.1.3 *Probaremos la **Proposición (4.1.2)** en tres pasos:*

***1)** Verificamos que **(1.1)** tiene una solución: primero, establecemos $d = \inf_{y \in C} \| x - y \|$*

y consideremos una sucesión (y_n) *en* C *tal que* $\lim_{n\longrightarrow\infty} \| y_n - x \| = d$. *Por propiedades basicas de la norma y el producto interno, "tenemos*

$$\| y_n - y_m \|^2 = \| (y_n - x) + (x - y_m) \|^2$$

$$\| y_n - y_m \|^2 = \| y_n - x \|^2 + \| x - y_m \|^2 + 2\langle y_n - x, x - y_m \rangle$$

$$\| y_n - y_m \|^2 = \| y_n - x \|^2 + \| y_m - x \|^2 - 2\langle y_n - x, y_m - x \rangle$$

$$\| y_n - y_m \|^2 = 2 \| y_n - x \|^2 + 2 \| y_m - x \|^2 - \| y_n - x \|^2 - \| y_m - x \|^2 - 2\langle y_n - x, y_m - x \rangle$$

$$\| y_n - y_m \|^2 = 2 \left(\| y_n - x \|^2 + \| y_m - x \|^2 \right) - \left(\| y_n - x \|^2 + \| y_m - x \|^2 + 2\langle y_n - x, y_m - x \rangle \right)$$

$$\| y_n - y_m \|^2 = 2 \left(\| y_n - x \|^2 + \| y_m - x \|^2 \right) - \left(\| y_n - x + y_m - x \| \right)^2$$

$$\| y_n - y_m \|^2 = 2 \left(\| y_n - x \|^2 + \| y_m - x \|^2 \right) - \| (y_n + y_m) - 2x \|^2 \ ^{\dagger}$$

dado que C *es convexo, el punto medio entre* y_n *e* y_m *pertenece a* C*" y por la definición de* d *tenemos lo siguiente*

$$\| (y_n + y_m) - 2x \|^2 = 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2 \geq 4d^2$$

$$- \| (y_n + y_m) - 2x \|^2 \leq -4d^2$$

entonces, "deducimos que

$$\| y_n - y_m \|^2 \leq 2 \left(\| y_n - x \|^2 + \| y_m - x \|^2 \right) - 4d^2$$

$$\| y_n - y_m \|^2 \leq 2 \left(\| y_n - x \|^2 + \| y_m - x \|^2 - 2d^2 \right)$$

haciendo que m *y* n *tiendan a* ∞, *obtenemos que* $(y_n)_{n\in\mathbb{N}}$ *es una sucesión de Cauchy, y debe converger a algún* y^*, *que debe estar en la cerradura de* C*". La continuidad de la*

[†]Conocida como la identidad de Apolonio.

norma implica

$$d = \lim_{n \longrightarrow \infty} \| y_n - x \| = \| y^* - x \|$$

esto muestra la existencia de y^*.

2) *"Establecemos la equivalencia entre* ***(1.1)*** *y* ***(1.2)****: Luego, suponemos que* ***(1.1)*** *se cumple y sea* $y \in C$. *Dado que* C *es convexo, para cada* $\lambda \in (0,1)$ *el punto* $\lambda y + (1-\lambda)y^*$ *también pertenece a* C. *Por lo tanto*

$$\| x - y^* \|^2 \leq \| x - (\lambda y + (1-\lambda)y^*) \|^2$$

$$\| x - y^* \|^2 \leq \| x - y^* + \lambda y^* - \lambda y \|^2$$

$$\| x - y^* \|^2 \leq \| x - y^* \|^2 + 2\lambda \langle x - y^*, y^* - y \rangle + \lambda^2 \| y^* - y \|^2$$

esto implica

$$\langle x - y^*, y - y^* \rangle \leq \frac{\lambda}{2} \| y^* - y \|^2$$

haciendo $\lambda \longrightarrow 0$ *obtenemos*

$$\langle x - y^*, y - y^* \rangle \leq 0$$

ahora, suponemos que ***(1.2)*** *se cumple*

$$\| x - y \|^2 = \| x - y^* + y^* - y \|^2$$

$$\| x - y \|^2 = \| x - y^* \|^2 + 2\langle x - y^*, y^* - y \rangle + \| y^* - y \|^2$$

$$\| x - y \|^2 \geq \| x - y^* \|^2$$

para cada $y \in C$ *se cumple* $\min_{y \in C} \| x - y \| = \| x - y^* \|$*".*

3) *Comprobamos que* ***(1.2)*** *no puede tener más de una solución: Finalmente, si* $y_1^*, y_2^* \in C$

satisfacen ***(1.2)****, entonces*

$$\langle x - y_1^*, y_2^* - y_1^* \rangle \leq 0 \quad y \quad \langle x - y_2^*, y_1^* - y_2^* \rangle \leq 0$$

factorizando $(-)$ *tenemos*

$$\langle x - y_1^*, y_2^* - y_1^* \rangle \leq 0 \quad y \quad -\langle x - y_2^*, y_2^* - y_1^* \rangle \leq 0$$

luego sumamos, así tenemos

$$\langle x - y_1^* - x + y_2^*, y_2^* - y_1^* \rangle \leq 0$$

$$\langle y_2^* - y_1^*, y_2^* - y_1^* \rangle \leq 0$$

$$\| y_2^* - y_1^* \|^2 \leq 0$$

El único caso posible es la igualdad, entonces $y_2^* = y_1^*$. ■

Definición 4.1.2 *La caracterización de* $P_{_C}(x)$ *dada por* ***(1.2)*** *dice que para cada* $x \notin C$*, el conjunto* C *se encuentra en el semiespacio cerrado*

$$S = \{y \in \mathcal{H} \ \mid \ \langle x - P_{_C}(x), y - P_{_C}(x) \rangle \leq 0\}$$

Corolario 4.1.2 *"Sea* C *un subconjunto no vacío, cerrado y convexo de* $\mathcal{H}$*. Entonces*

$$C = \bigcap_{x \notin C} \{y \in \mathcal{H} \ \mid \ \langle x - P_{_C}(x), y - P_{_C}(x) \rangle \leq 0\}$$

La inversa, la intersección de semiespacios convexos cerrados es cerrada y convexa".

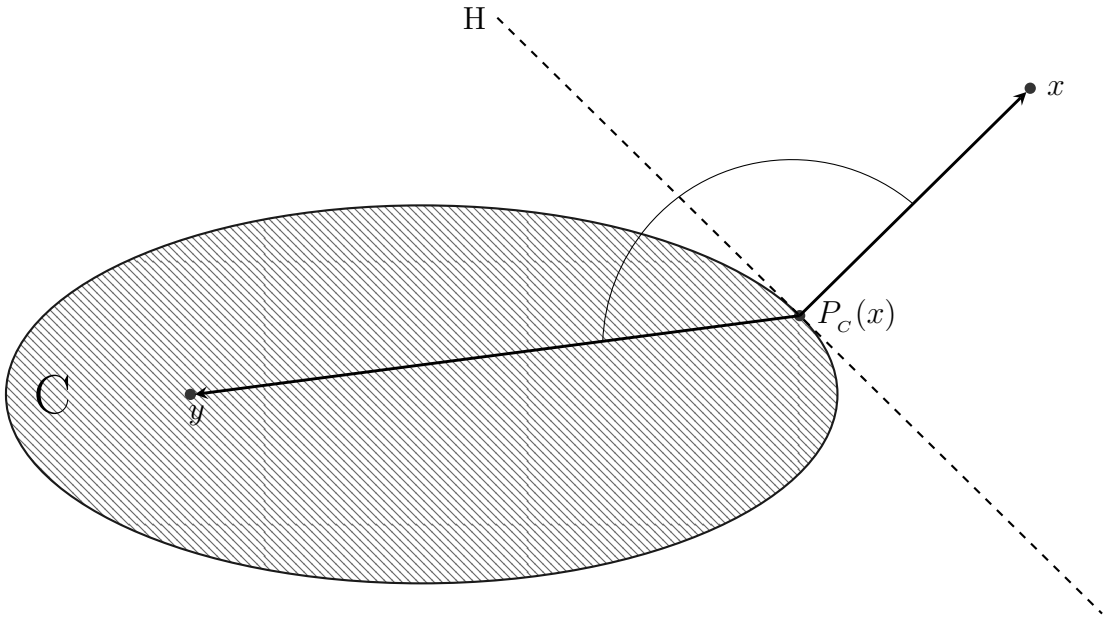


Figura 4.2: Se define un Semiespacio, que son todos los elementos (vectores) que forman un ángulo mayor a 90° respecto del vector $\overrightarrow{P_C(x)x}$. Es decir, es todo lo que esta a la izquierda del hiperplano H.

Otra propiedad de la proyección, con importantes consecuencias topológicas, es la siguiente:

Proposición 4.1.3 *"Sea C un subconjunto no vacío, cerrado y convexo de $\mathcal{H}$. La función $x \longrightarrow P_C(x)$ es no expansiva, es decir $1-$Lipschitz".*

Prueba 4.1.4 *Sea x_1,$x_2 \in \mathcal{H}$, entonces*

$$\langle x_1 - P_C(x_1), P_C(x_2) - P_C(x_1) \rangle \leq 0$$

y también

$$\langle x_2 - P_C(x_2), P_C(x_1) - P_C(x_2) \rangle \leq 0$$

"sumando estas dos desigualdades obtenemos

$$\| P_C(x_1) - P_C(x_2) \|^2 \leq \langle x_1 - x_2, P_C(x_1) - P_C(x_2) \rangle$$

Concluimos utilizando la desigualdad de Cauchy-Schwarz

$$\| P_C(x_1) - P_C(x_2) \|^2 \leq \| x_1 - x_2 \| \, . \, \| P_C(x_1) - P_C(x_2) \|$$

$$| P_C(x_1) - P_C(x_2) | \leq \| x_1 - x_2 \|$$

Con constante de lipschitz 1*".* ■

4.2 Ortogonalidad

Definición 4.2.1 *C es un subespacio afin si*

$$C = \lambda C + (1 - \lambda)C \tag{2.3}$$

Con $C \neq \emptyset$ *y* $\lambda \in \mathbb{R}$*.*

Las proyecciones en subespacios afines tienen propiedades adicionales, como la proyección ortogonal

Corolario 4.2.1 *Sea C un subespacio afín cerrado de $\mathcal{H}$. Se cumple lo siguiente:*

(i) *Sea $x, y^* \in \mathcal{H}$. Entonces $y^* = P_C(x)$ si y solo si*

$$\langle y - z, x - y^* \rangle = 0$$

Con $y^* \in C$ y para todo $y, z \in C$. En otras palabras $x - P_C(x) \perp C$.

(ii) *P_C es un operador afín.*

Prueba 4.2.1 *La prueba detallada la podemos encontrar en Peypouquet (2015).*

5 Moreau-Yosida

Presento una técnica de regularización para funciones definidas en un espacio de Hilbert. Al agregar un término cuadrático, se puede forzar la existencia y unicidad de un minimizador. Este hecho, es el núcleo de un importante método de minimización, conocido como algoritmo del punto proximal. Para todo ello estaré utilizando Peypouquet (2015).

Teorema 5.0.1 *Sea* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *propia, semicontinua inferior y inf-compacta*[*]. *Entonces* $\operatorname{argmin}(F)$ *es no vacío y compacto. Además,* $\text{ínf}(F) > -\infty$.

Prueba 5.0.1 *"La prueba se basa en considerar* (γ_n) *una sucesión no creciente tal que*

$$\lim_{n\longrightarrow\infty} \gamma_n = \text{ínf}(F)$$

(posiblemente $-\infty$*). Para cada* $n \in \mathbb{N}$ *considere el conjunto compacto* $K_n = \underset{\gamma_n}{\operatorname{Lev}}(F)$. *Entonces* $K_{n+1} \subseteq K_n$ *para cada* $n \in \mathbb{N}$ *y así*

$$\operatorname{argmin}(F) = \bigcap_{n\in\mathbb{N}} K_n$$

es no vacío y compacto". Claramente, $\text{ínf}(F) > -\infty$ *dado que la función* F *es Propia.* ■

Observación 5.0.1 *Consideremos lo siguiente:*

1) *Para que esta intersección sea no vacía y compacta, es necesario que el* $\text{ínf}(F)$ *sea finito. Esto se debe a que si* $\text{ínf}(F)$ *fuera* $-\infty$*, la sucesión* (γ_n) *podría no estar bien definida de manera que los conjuntos* K_n *permanezcan compactos y no vacíos.*

[*]Una función $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ es *inf-compacta* si, para cada $\gamma \in \mathbb{R}$, el conjunto de subnivel $\underset{\gamma}{\operatorname{Lev}}(F)$ es relativamente compacto. Claramente, una función $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ es propia, semicontinua inferior e *inf-compacta* si, y solo si, $\underset{\gamma}{\operatorname{Lev}}(F)$ es no vacío y compacto para cada $\gamma > \text{ínf}(F)$.

2) *La afirmación* $\inf(F) > -\infty$ *es crucial porque garantiza que* F *está acotada inferiormente, lo que es necesario para asegurar la existencia de un mínimo finito y bien definido. Esto es consistente con la definición de una función propia y es esencial para la demostración de que* $\operatorname{argmin}(F)$ *es no vacío y compacto.*

Teorema 5.0.2 *Si* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *es propia, convexa, coerciva y semicontinua inferior, entonces* $\operatorname{argmin}(F)$ *es no vacía y débilmente compacta. Además, si* F *es estrictamente convexa, entonces* $\operatorname{argmin}(F)$ *es un conjunto unitario.*

Prueba 5.0.2 *La función* F *cumple con las hipótesis del* ***Teorema (5.0.1)*** *para la topología débil. Claramente, una función estrictamente convexa no puede tener más de un minimizador.*

5.1 Regularización de Moreau-Yosida

La regularización de Moreau-Yosida convierte una función posiblemente no diferenciable F en una versión suavizada $F_{(\lambda,x)}$, agregando un término cuadrático de penalización controlado por λ. Este proceso facilita la minimización de F al hacerlo más estable y diferenciable, a la vez que preserva sus propiedades esenciales. Es un concepto clave en optimización convexa y tiene aplicaciones importantes en diversas áreas de la ciencia y la ingeniería.

Definición 5.1.1 *Sea* $\mathcal{H}$ *un espacio de Hilbert y* $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ *sea propia, convexa y semicontinua inferior. "Dado* $\lambda > 0$ *y* $x \in \mathcal{H}$*, la regularización de Moreau-Yosida de* F *con el parámetro* (λ,x) *es la función* $F_{(\lambda,x)} : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$*, definida por*

$$F_{(\lambda,x)}(z) = F(z) + \frac{1}{2\lambda} \parallel z - x \parallel^2$$

La siguiente proposición es crucial en la teoría de optimización convexa y análisis variacional, ya que proporciona una condición necesaria y suficiente para encontrar el punto que minimiza la regularización de una función".

Proposición 5.1.1 *Para cada $\lambda > 0$ y $x \in \mathcal{H}$, la función $F_{(\lambda,x)}$ tiene un minimizador único x^*. Por otra parte, x^* se caracteriza por la inclusión*

$$-\frac{x^* - x}{\lambda} \in \widehat{\partial} F(x^*)$$

Prueba 5.1.1 *Dado que $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ es propia, convexa, semicontinua inferior y coerciva*[†]*, entonces tenemos que*

(i) ***Propia:*** *Para que $F_{(\lambda,x)}$ sea propia, necesitamos que su dominio no esté vacío y que sus valores sean finitos para todos los puntos en su dominio.*

Dado que F es propia, $\mathrm{dom}(F) \neq \emptyset$ y $F(z) > -\infty$ para todo $z \in \mathrm{dom}(F)$. Además, el término cuadrático $\frac{1}{2\lambda}\|z - x\|^2$ es siempre finito y definido para todos los $z \in \mathcal{H}$. Por lo tanto, $\mathrm{dom}\left(F_{(\lambda,x)}\right) = \mathrm{dom}(F)$ y los valores de $F_{(\lambda,x)}$ son finitos. Por lo tanto, $F_{(\lambda,x)}$ es propia.

(ii) ***Estrictamente Convexa:*** *Para que $F_{(\lambda,x)}$ sea estrictamente convexa para todo $z_1, z_2 \in \mathrm{dom}\left(F_{(\lambda,x)}\right)$ y $\theta \in (0,1)$, se tiene que cumplir:*

$$F_{(\lambda,x)}(\theta z_1 + (1-\theta)z_2) \;<\; \theta F_{(\lambda,x)}(z_1) + (1-\theta)F_{(\lambda,x)}(z_2)$$

Considerando la regularización de Moreau-Yosida

$$F_{(\lambda,x)}(z) = F(z) + \frac{1}{2\lambda}\|z - x\|^2$$

El término $\frac{1}{2\lambda}\|z - x\|^2$ es una función cuadrática estrictamente convexa. La suma de una función estrictamente convexa y una función convexa resulta en una función estric-

[†]Sea $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ una función definida en un espacio de Hilbert $\mathcal{H}$. Decimos que F es coerciva si:

$$\lim_{\|x\| \longrightarrow +\infty} F(x) = +\infty$$

tamente convexa si la parte cuadrática domina.

En este caso, dado que $\frac{1}{2\lambda}\|z-x\|^2$ *es estrictamente convexa con* $\lambda > 0$ *y la función* F *es convexa, entonces la función* $F_{(\lambda,x)}(z)$ *es estrictamente convexa.*

(iii) ***Semicontinua Inferior:*** *Para demostrar que* $F_{(\lambda,x)}$ *es semicontinua inferior, debemos mostrar que para cualquier sucesión* $\{z_n\}$ *que converge a* z*, se tiene:*

$$F_{(\lambda,x)}(z) \leq \liminf_{n\longrightarrow\infty} F_{(\lambda,x)}(z_n)$$

Dado que F *es semicontinua inferior, tenemos:*

$$F(z) \leq \liminf_{n\longrightarrow\infty} F(z_n)$$

Y como el término $\frac{1}{2\lambda}\|z-x\|^2$ *es continuo (por ser una función cuadrática), se tiene:*

$$\frac{1}{2\lambda}\|z-x\|^2 \leq \liminf_{n\longrightarrow\infty} \frac{1}{2\lambda}\|z_n-x\|^2$$

Sumando estas dos desigualdades, obtenemos:

$$F(z)+\frac{1}{2\lambda}\|z-x\|^2 \leq \liminf_{n\longrightarrow\infty}\left(F(z_n)+\frac{1}{2\lambda}\|z_n-x\|^2\right)$$

Esto implica que $F_{(\lambda,x)}$ *es semicontinua inferior.*

(iv) ***Coerciva:*** $F_{(\lambda,x)}$ *es coercitiva si* $\|z\| \longrightarrow \infty$ *implica* $F_{(\lambda,x)}(z) \longrightarrow \infty$*.*

Dado que F *es propia, podemos asumir que* F *no decrece indefinidamente. Además, el término cuadrático* $\frac{1}{2\lambda}\|z-x\|^2$ *claramente crece a infinito cuando* $\|z\| \longrightarrow \infty$*. Por lo tanto, la suma* $F(z)+\frac{1}{2\lambda}\|z-x\|^2$ *también tiende a infinito cuando* $\|z\| \longrightarrow \infty$*.*

Por lo tanto, $F_{(\lambda,x)}$ *es coerciva.*

$F_{(\lambda,x)}$ *es propia, estrictamente convexa, semicontinua inferior y coerciva.*

La unicidad y la existencia de un minimizador x^* de $F_{(\lambda,x)}$, esta dado por ***Teorema (5.0.2)****.*

Luego, por el ***Teorema de Fermat (3.10.1)*** *tenemos*

$$0 \in \widehat{\partial} F_{(\lambda,x)}\left(x^*\right)$$
$$0 \in \widehat{\partial}\left(F\left(x^*\right) + \frac{1}{2\lambda}\|x^* - x\|^2\right)$$

Por el ***Teorema de Moreau-Rockafellar (3.6.1)*** *tenemos:*

$$0 \in \left(\widehat{\partial} F\left(x^*\right) + \widehat{\partial}\left(\frac{1}{2\lambda}\|x^* - x\|^2\right)\right)$$

Visto que $\frac{1}{2\lambda}\|z - x\|^2$ *es diferenciable, entonces el Subdiferencial de Fréchet coincide con el conjunto que contiene a su gradiente* [‡]

$$0 \in \widehat{\partial} F\left(x^*\right) + \frac{x^* - x}{\lambda}$$

Lo que es equivalente a

$$-\frac{x^* - x}{\lambda} \in \widehat{\partial} F(x^*)$$

que es el resultado que queriamos demostrar. ■

[‡]Para cualquier función que sea diferenciable en un punto, el subdiferencial de Fréchet en ese punto se reduce al conjunto que contiene únicamente el gradiente de la función en ese punto. Esto se debe a que no hay otros límites de gradientes que considerar, y el gradiente mismo ya proporciona toda la información necesaria.

$$\widehat{\partial} G(z) = \{\nabla G(z)\}$$

En el caso específico de $\frac{1}{2\lambda}\|z-x\|^2$, esta función es diferenciable en todos los puntos, por lo que su subdiferencial de Fréchet en cualquier punto z es simplemente:

$$\widehat{\partial}\left(\frac{1}{2\lambda}\|z - x\|^2\right)(z) = \left\{\nabla\left(\frac{1}{2\lambda}\|z - x\|^2\right)\right\} = \left\{\frac{z - x}{\lambda}\right\}$$

5.2 Envolvente de Moreau-Yosida

Definición 5.2.1 *Recordemos de la* ***Proposición (5.1.1)*** *que para cada $\lambda > 0$ y $x \in \mathcal{H}$, la Regularización de Moreau-Yosida $F(\lambda, x)$ de F tiene un minimizador único x^*. El envolvente de Moreau-Yosida de F con parámetro $\lambda > 0$ es la función $F_\lambda : \mathcal{H} \longrightarrow \mathbb{R}$ definida por*

$$F_\lambda(x) = \min_{z \in \mathcal{H}} \{F(\lambda, x)(z)\} = \min_{z \in \mathcal{H}} \left\{ F(z) + \frac{1}{2\lambda}\|z - x\|^2 \right\} = F(x^*) + \frac{1}{2\lambda}\|x^* - x\|^2$$

Observación 5.2.1 (Relación con la regularización de Moreau-Yosida) *La regularización de Moreau-Yosida transforma una función potencialmente no diferenciable en una función suavizada. Sin embargo, la envolvente de Moreau-Yosida es la evaluación del valor mínimo de esta regularización, lo que nos da una función que es más fácil de manejar numéricamente y proporciona mejores propiedades para los algoritmos de optimización.*

Ejemplo 5.2.1 *Consideremos la función $F(x) = |x|$. La regularización de Moreau-Yosida de F con parámetro $\lambda > 0$ está dada por:*

$$F_\lambda(x) = \inf_{y \in \mathbb{R}} \left\{ |y| + \frac{1}{2\lambda}(x - y)^2 \right\}$$

Para encontrar $F_\lambda(x)$, minimizamos la función dentro del ínfimo, consideramos dos casos: $y \geq 0$ e $y < 0$.

- ***Caso 1:*** $y \geq 0$

$$g(y) = y + \frac{1}{2\lambda}(x - y)^2$$

Derivamos respecto a y:

$$g'(y) = 1 - \frac{x - y}{\lambda}$$

Igualamos a cero para encontrar el mínimo:

$$1 - \frac{x-y}{\lambda} = 0 \implies y = x - \lambda$$

- ***Caso 2:*** $y < 0$

$$g(y) = -y + \frac{1}{2\lambda}(x-y)^2$$

Derivamos respecto a y:

$$g'(y) = -1 - \frac{x-y}{\lambda}$$

Igualamos a cero para encontrar el mínimo:

$$-1 - \frac{x-y}{\lambda} = 0 \implies y = x + \lambda$$

Evaluamos en ambos casos:

$$F_\lambda(x) = \begin{cases} x - \lambda + \frac{1}{2\lambda}\lambda^2, & si\ x > \lambda \\ 0 + \frac{1}{2\lambda}x^2, & si\ |x| \leq \lambda \\ -x - \lambda + \frac{1}{2\lambda}\lambda^2, & si\ x < -\lambda \end{cases}$$

Simplificando, obtenemos:

$$F_\lambda(x) = \begin{cases} x - \frac{\lambda}{2}, & si\ x > \lambda \\ \frac{x^2}{2\lambda}, & si\ |x| \leq \lambda \\ -x - \frac{\lambda}{2}, & si\ x < -\lambda \end{cases}$$

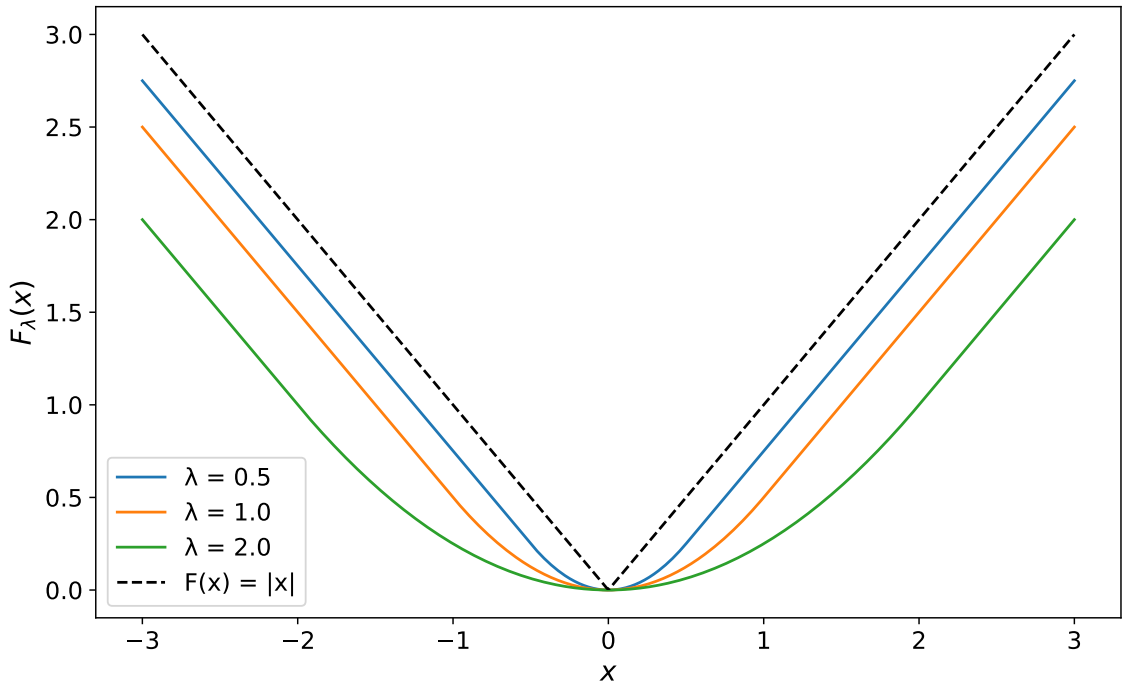


Figura 5.1: Grafica de la regularización de Moreau-Yosida de $F(x) = |x|$ para diferentes valores de λ.

Ejemplo 5.2.2 *Consideremos la función a tramos:*

$$F(x) = \begin{cases} 1 - x & si\ x < 1 \\ x^2 - 1 & si\ x \geq 1 \end{cases}$$

La regularización de Moreau-Yosida de F con parámetro $\lambda > 0$ está dada por:

$$F_\lambda(x) = \inf_{y \in \mathbb{R}} \left\{ F(y) + \frac{1}{2\lambda}(x - y)^2 \right\}$$

- ***Caso 1:*** $y < 1$

$$g(y) = 1 - y + \frac{1}{2\lambda}(x - y)^2$$

Derivada de $g(y)$:

$$g'(y) = -1 + \frac{1}{\lambda}(y - x)$$

Igualamos a cero para encontrar el punto crítico:

$$-1 + \frac{1}{\lambda}(y - x) = 0$$

$$y = x + \lambda$$

- ***Caso 2:*** $y \geq 1$

$$g(y) = y^2 - 1 + \frac{1}{2\lambda}(x - y)^2$$

Derivada de $g(y)$:

$$g'(y) = 2y - \frac{1}{\lambda}(x - y)$$

Igualamos a cero para encontrar el punto crítico:

$$2y - \frac{1}{\lambda}(x - y) = 0$$

$$y = \frac{x\lambda}{2\lambda + 1}$$

Verificando las soluciones dentro de las regiones respectivas

- ***Para*** $y < 1$***:*** *Si $y = x + \lambda < 1$, la solución $y = x + \lambda$ es válida:*

$$g(x + \lambda) = 1 - (x + \lambda) + \frac{1}{2\lambda}(x - (x + \lambda))^2$$

$$g(x + \lambda) = 1 - x - \frac{\lambda}{2}$$

- ***Para*** $y \geq 1$***:*** *Si* $y = \dfrac{x\lambda}{2\lambda+1} \geq 1$*, la solución es válida:*

$$g\left(\frac{x\lambda}{2\lambda+1}\right) = \left(\frac{x\lambda}{2\lambda+1}\right)^2 - 1 + \frac{1}{2\lambda}\left(x - \frac{x\lambda}{2\lambda+1}\right)^2$$

$$g\left(\frac{x\lambda}{2\lambda+1}\right) = \frac{x^2\lambda^2}{(2\lambda+1)^2} - 1 + \frac{1}{2\lambda}\left(x - \frac{x\lambda}{2\lambda+1}\right)^2$$

Así, tenemos la función de Regularización de Moreau-Yosida de F

$$F_\lambda(x) = \begin{cases} 1 - x - \dfrac{\lambda}{2} & si\ x + \lambda < 1 \\ \dfrac{x^2\lambda^2}{(2\lambda+1)^2} - 1 + \dfrac{1}{2\lambda}\left(x - \dfrac{x\lambda}{2\lambda+1}\right)^2 & si\ \dfrac{x\lambda}{2\lambda+1} \geq 1 \end{cases}$$

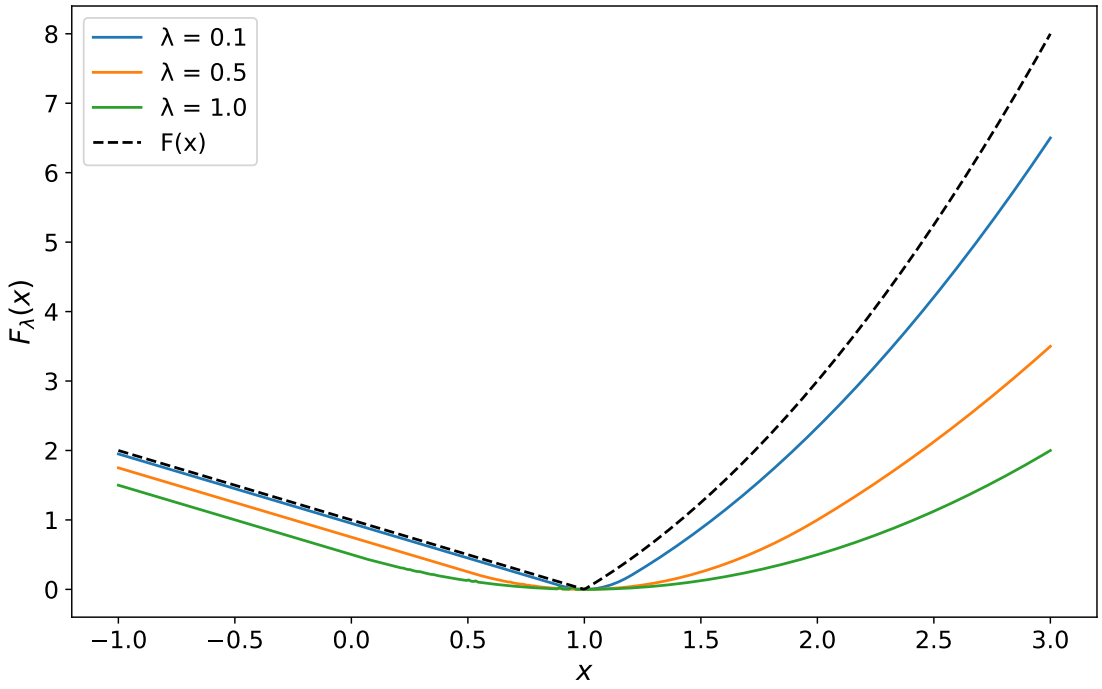


Figura 5.2: Grafica de la regularización de Moreau-Yosida de $F(x)$ para diferentes valores de λ.

6 El Problema

El objetivo de esta tesis es analizar la convergencia del Algoritmo de Punto Proximal (PPA) en espacios de Hilbert y desarrollar una implementación en Python. El PPA es un método iterativo ampliamente utilizado en la optimización de funciones convexas y cuasiconvexas, especialmente en casos donde la función objetivo no es diferenciable. Este algoritmo es útil para resolver problemas en diversas áreas de la ciencia y la ingeniería, donde la optimización juega un papel crucial.

6.1 Formulación del Problema

Consideremos una función $F : \mathcal{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$, donde $\mathcal{H}$ es un espacio de Hilbert.
Para aplicar el Algoritmo del Punto Proximal, la función objetivo F debe cumplir las siguientes condiciones:

- **Convexa**: Esta propiedad asegura que la función no tiene más de un mínimo local, lo que es esencial para la convergencia del algoritmo hacia un mínimo global.
- **Dominios Cerrados y Convexos**: El dominio de F debe ser un conjunto cerrado y convexo. Esta condición facilita la implementación del algoritmo y garantiza que las iteraciones permanezcan dentro de un conjunto factible.
- **Continuidad Lipschitz**: La función F debe ser Lipschitz continua en su dominio. Esta condición ayuda a controlar el comportamiento de F y asegura que no tenga cambios abruptos, lo cual es útil para la convergencia del algoritmo.
- **Función Propia**: La función objetivo F debe ser propia, es decir, no debe ser idénticamente $+\infty$ y debe ser finita en al menos un punto del dominio.

- **Semicontinuidad Inferior**: Esta propiedad asegura que los valores de la función no se reducen abruptamente en los puntos de acumulación, lo que contribuye a la estabilidad del algoritmo.

El problema consiste en minimizar F utilizando el PPA. La iteración del PPA se define como:

$$x_{k+1} = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \left\{ F(x) + \frac{1}{2\lambda_k} \|x - x_k\|^2 \right\},$$

donde $\lambda_k > 0$ es una sucesión de parámetros positivos y x_k es la iteración actual.
"La regularización de Moreau-Yosida de F con parámetro $\lambda > 0$ está dada por:

$$F_\lambda(x) = \underset{y \in \mathcal{H}}{\text{ínf}} \left\{ F(y) + \frac{1}{2\lambda} \|x - y\|^2 \right\}.$$

esta regularización suaviza la función F, facilitando su minimización mediante el PPA".

7 Algoritmo

"Algoritmo del Punto Proximal inicia con $x_0 \in \mathcal{H}$. Para cada $n \geq 0$, dado un paso λ_n y el estado x_n, defina x_{n+1} como el único minimizador de la Regularización de Moreau-Yosida F_{λ_n, x_n} de F, que es una función propia, semicontinua inferior y fuertemente convexa, definida como

$$F_{(\lambda_n, x_n)}(z) = F(z) + \frac{1}{2\lambda_n}\|z - x_n\|^2$$

Lo que se puede expresar (Según el **Capitulo 5**) como

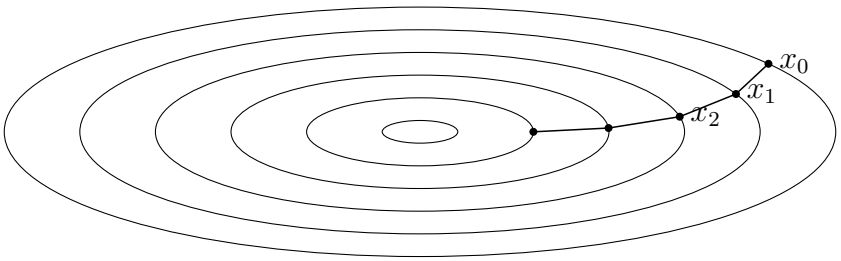
$$\{x_{n+1}\} = \underset{z \in \mathcal{H}}{\operatorname{argmin}} \left\{F_{(\lambda_n, x_n)}(z)\right\} = \operatorname{argmin}\left\{F(z) + \frac{1}{2\lambda_n}\|z - x_n\|^2 : z \in \mathcal{H}\right\} \tag{0.1}$$

También tenemos, por el **Teorema de Moreau-Rockafellar (3.6.1)**,

$$0 \in \widehat{\partial} F_{(\lambda_n, x_n)}(x_{n+1})$$
$$0 \in \widehat{\partial} F(x_{n+1}) + \frac{x_{n+1} - x_n}{\lambda_n},$$

o lo que es equivalente a",

$$-\frac{x_{n+1} - x_n}{\lambda_n} \in \widehat{\partial} F(x_{n+1}) \tag{0.2}$$



La inclusión de **(0.2)** se puede escribirse (según Peypouquet (2015)) como

$$x_{n+1} = \left(I + \lambda_n \widehat{\partial} F\right)^{-1}(x_n)$$

Una sucesión (x_n) generada por el algoritmo del punto proximal se considera una sucesión proximal. Los puntos donde esta sucesión se estabiliza son los minimizadores de la función objetivo F, ya que,

$$x_{n+1} = x_n \quad \text{si, y solo si,} \quad 0 \in \widehat{\partial} F(x_{n+1}).$$

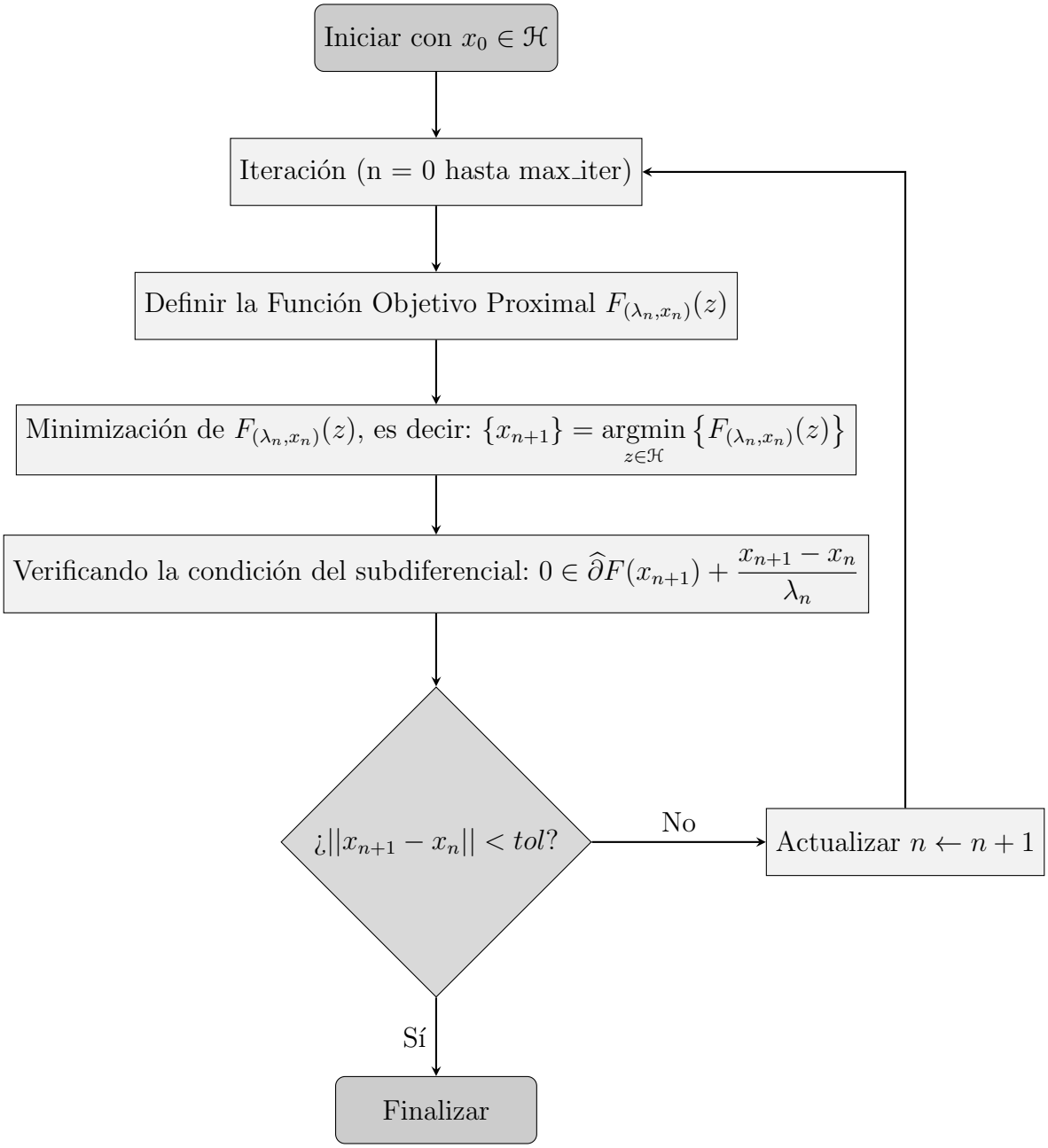
Según la definición del Algoritmo de Punto Proximal (PPA) usado en **(0.1)**, se establece que:

$$F(x_{n+1}) + \frac{1}{2\lambda_n}\|x_{n+1} - x_n\|^2 \leq F(x_n) \ , \quad \text{para cada } n \geq 0$$

Por lo tanto, la sucesión $(F(x_n))$ no es creciente, decrece estrictamente siempre que $x_{n+1} \neq x_n$. Eso quiere decir, en cada iteración, el punto x_{n+1} es distinto de x_n solo si se logra una reducción suficiente en el valor de la función objetivo.

7.1 Pseudocodigo

El algoritmo del punto proximal (PPA) es un método iterativo utilizado para resolver problemas de minimización en los cuales la función objetivo puede no ser diferenciable. Este método se adapta especialmente a funciones convexas utilizando el subdiferencial de Fréchet en espacios de Hilbert. A continuación, se presenta una descripción detallada del proceso implementado.



Se proporciona una explicación detallada del **Pseudocodigo**

- **Inicio**: El algoritmo comienza con un punto inicial x_0 en el espacio de Hilbert $\mathcal{H}$. Este es el punto de partida para la iteración.

- **Iteración**: Se inicia un ciclo de iteraciones desde $n = 0$ hasta un número máximo de iteraciones (max_iter). Este paso controla el número de iteraciones del algoritmo.

- **Definir la Función Objetivo Proximal**: Para cada iteración n, definir la función objetivo proximal $F_{(\lambda_n, x_n)}(z)$ como:

$$F_{(\lambda_n, x_n)}(z) = F(z) + \frac{1}{2\lambda_n}\|z - x_n\|^2$$

  Esta función incluye un término de penalización cuadrática.

- **Minimización de** $F_{(\lambda_n, x_n)}(z)$: Minimizar la función objetivo proximal para encontrar el siguiente punto x_{n+1}:

$$x_{n+1} = \underset{z \in \mathcal{H}}{\operatorname{argmin}} \left\{ F_{(\lambda_n, x_n)}(z) \right\}$$

  Este paso encuentra el punto que minimiza la función objetivo proximal en la iteración actual.

- **Verificar la Condición del Subdiferencial**: Verificar que la condición del subdiferencial se cumple:

$$0 \in \widehat{\partial} F(x_{n+1}) + \frac{x_{n+1} - x_n}{\lambda_n}$$

  Esto asegura que el punto x_{n+1} es un minimizador adecuado de la función objetivo proximal.

- **Convergencia**: Comprobar si la diferencia entre x_{n+1} y x_n es menor que una tolerancia predefinida (tol):

$$\|x_{n+1} - x_n\| < \text{tol}$$

  Si la condición se cumple, el algoritmo ha convergido a una solución suficientemente precisa.

- **Finalizar**: Si la condición de convergencia se cumple, el algoritmo termina. Esto detiene el algoritmo una vez que ha convergido.

- **Actualizar** $n \leftarrow n + 1$: Si la condición de convergencia no se cumple, actualizar el

contador de iteraciones n y continuar con la siguiente iteración. Este paso prepara para la próxima iteración del ciclo.

Algorithm 1 Algoritmo del Punto Proximal

Inicio con $x_0 \in \mathcal{H}$
for $n = 0$ to max_iter **do**
    Definir la Función Objetivo Proximal $F_{(\lambda_n, x_n)}(z)$ como:

$$F_{(\lambda_n, x_n)}(z) = F(z) + \frac{1}{2\lambda_n}\|z - x_n\|^2$$

    Minimizar $F_{(\lambda_n, x_n)}(z)$ para obtener x_{n+1}:

$$x_{n+1} = \arg\min_{z \in \mathcal{H}} \left\{ F_{(\lambda_n, x_n)}(z) \right\}$$

    Verificar la condición del subdiferencial:

$$0 \in \widehat{\partial} F(x_{n+1}) + \frac{x_{n+1} - x_n}{\lambda_n}$$

    if $\|x_{n+1} - x_n\| < tol$ **then**
        Finalizar
    else
        Actualizar $n \leftarrow n + 1$
    end if
end for

8 Convergencia del Algoritmo

Se estudia el comportamiento de x_n y $y_n = \dfrac{x_n - x_{n+1}}{\lambda_n}$. La inclusión **(0.2)** y la definición del **Subdiferencial de Fréchet (4.12)** (teniendo en cuenta la **Proposición (3.10.1)** juntas implican que para cada $u \in \mathcal{H}$,

$$F(u) \geq F(x_{n+1}) + \left\langle -\frac{x_{n+1} - x_n}{\lambda_n}, u - x_{n+1} \right\rangle$$

El siguiente Lema establece que si la norma de $\|y_n\|$ es decreciente, sugiere que la magnitud de los cambios en las iteraciones del algoritmo se está reduciendo, lo cual es un indicio de convergencia. En otras palabras, el algoritmo se está estabilizando alrededor de un punto, y las iteraciones sucesivas se están acercando a una solución del problema que se está resolviendo. Este comportamiento es fundamental en la convergencia de métodos iterativos, como el algoritmo de punto proximal, que depende de la reducción progresiva de las diferencias entre iteraciones para garantizar que el método alcanzará el punto óptimo.

Lema 8.0.1 *La sucesión dada por* $\{\|y_n\|\}_{n=1}^{\infty}$ *es una sucesión decreciente.*

Prueba 8.0.1 *Dado que* $y_n \in \widehat{\partial} F(x_n)$, $y_{n+1} \in \widehat{\partial} F(x_{n+1})$, *y el* $\widehat{\partial} F$ *es un operador monótono, tenemos:*

$$\langle y_{n+1} - y_n, x_{n+1} - x_n \rangle \geq 0$$

"En el contexto del Algoritmo del Punto Proximal (PPA), tenemos que $y_{n+1} = \dfrac{x_n - x_{n+1}}{\lambda_{n+1}}$, *así obtenemos:*

$$\left\langle \frac{x_n - x_{n+1}}{\lambda_{n+1}} - y_n, x_{n+1} - x_n \right\rangle \geq 0$$

$$-\frac{1}{\lambda_{n+1}} \langle x_n - x_{n+1} - y_n \lambda_{n+1}, x_n - x_{n+1} \rangle \geq 0$$

$$\langle x_n - x_{n+1} - y_n\lambda_{n+1}, x_n - x_{n+1}\rangle \leq 0$$

$$\langle x_n - x_{n+1}, x_n - x_{n+1}\rangle - \lambda_{n+1}\langle y_n, x_n - x_{n+1}\rangle \leq 0$$

$$\langle x_n - x_{n+1}, x_n - x_{n+1}\rangle \leq \lambda_{n+1}\langle y_n, x_n - x_{n+1}\rangle$$

sabemos que $\langle x_n - x_{n+1}, x_n - x_{n+1}\rangle = \|x_n - x_{n+1}\|^2$*", entonces*

$$\|x_n - x_{n+1}\|^2 \leq \lambda_{n+1}\langle y_n, x_n - x_{n+1}\rangle$$

se cumple $y_{n+1} = \dfrac{x_n - x_{n+1}}{\lambda_{n+1}} \Longrightarrow \|y_{n+1}\| = \dfrac{\|x_n - x_{n+1}\|}{\lambda_{n+1}} \Longrightarrow \lambda_{n+1}\|y_{n+1}\| = \|x_n - x_{n+1}\|$*, es así que tenemos*

$$(\lambda_{n+1}\|y_{n+1}\|)^2 \leq \lambda_{n+1}\langle y_n, x_n - x_{n+1}\rangle$$

$$\|y_{n+1}\|^2 \leq \frac{1}{\lambda_{n+1}}\langle y_n, x_n - x_{n+1}\rangle$$

$$\|y_{n+1}\|^2 \leq \left\langle y_n, \frac{x_n - x_{n+1}}{\lambda_{n+1}}\right\rangle$$

$$\|y_{n+1}\|^2 \leq \langle y_n, y_{n+1}\rangle \tag{0.1}$$

por la derecha se cumple la desigualdad de Cauchy-Schwarz [*]

$$\langle y_n, y_{n+1}\rangle \leq \|y_n\| \cdot \|y_{n+1}\| \tag{0.2}$$

juntando ***(0.1)*** *y* ***(0.2)*** *tenemos*

$$\|y_{n+1}\|^2 \leq \langle y_n, y_{n+1}\rangle \leq \|y_n\| \cdot \|y_{n+1}\|$$

[*]Para dos vectores u y v en un espacio vectorial con producto interno, la desigualdad de Cauchy-Schwarz se expresa como:

$$|\langle u, v\rangle| \leq \|u\|\|v\|,$$

donde $\langle u, v\rangle$ denota el producto interno de u y v, y $\|u\|$ y $\|v\|$ son las normas de los vectores u y v, respectivamente.

$$\|y_{n+1}\|^2 \leq \|y_n\| \cdot \|y_{n+1}\|$$

$$\|y_{n+1}\| \leq \|y_n\|$$

Lo que demuestra que $\{\|y_n\|\}_{n=1}^{\infty}$ *es una sucesión decreciente.* ■

El siguiente Lema proporciona una cota superior para la diferencia entre $F(x_n)$ y $F(u)$, lo que implica que $F(x_n)$ no puede ser arbitrariamente grande en comparación con $F(u)$. En otras palabras, a medida que el algoritmo avanza (es decir, conforme n aumenta), la diferencia entre $F(x_n)$ y $F(u)$ se reduce, lo que sugiere que x_n está convergiendo hacia u, y $F(x_n)$ está convergiendo hacia $F(u)$. Los términos en el lado derecho de la desigualdad reflejan la progresiva aproximación del algoritmo al punto óptimo: a medida que n crece, $\|x_n - u\|$ y $\|y_n\|$ deben disminuir, lo que indica que el algoritmo está funcionando correctamente y acercándose a la solución.

Lema 8.0.2 *Consideremos una sucesión arbitraria de números positivos* $\{\lambda_j\}_{j=1}^{\infty}$. *Supongamos que el Algoritmo del Punto Proximal (PPA) comienza en* x_0 *y genera la sucesión* $\{x_n\}_{n=0}^{\infty}$, *donde* $x_n = (I + \lambda_n F)^{-1} x_{n-1} \equiv J_{\lambda_n}(x_{n-1})$. *Entonces, para cualquier* $u \in \mathcal{H}$, *se cumple que*

$$F(x_n) - F(u) \leq \frac{\|u - x_0\|^2}{2\sigma_n} - \frac{\|u - x_n\|^2}{2\sigma_n} - \frac{\sigma_n}{2}\|y_n\|^2$$

Prueba 8.0.2 *"Como ya sabemos* $y_k = \dfrac{x_{k-1} - x_k}{\lambda_k} \in \widehat{\partial} F(x_k)$

$$F(u) - F(x_k) \geq \langle y_k, u - x_k \rangle$$

$$F(u) - F(x_k) \geq \left\langle \frac{x_{k-1} - x_k}{\lambda_k}, u - x_k \right\rangle \tag{0.3}$$

por lo tanto,

$$2\lambda_k \left[F(u) - F(x_k)\right] \geq 2 \langle x_{k-1} - x_k, u - x_k \rangle$$

$$2\lambda_k \left[F(u) - F(x_k)\right] \geq 2 \langle x_{k-1}, u \rangle + 2 \langle x_k, x_k \rangle - 2 \langle x_k, u \rangle - 2 \langle x_{k-1}, x_k \rangle$$

dando forma, multiplicando por (-1)*, sumando y restando los terminos* $\langle x_{k-1},x_{k-1}\rangle$ *y* $\langle u,u\rangle$

$$2\lambda_k\left[F(x_k)-F(u)\right]\leq\left(\langle u,u\rangle-2\langle x_{k-1},u\rangle+\langle x_{k-1},x_{k-1}\rangle\right)$$
$$-\left(\langle u,u\rangle-2\langle x_k,u\rangle+\langle x_k,x_k\rangle\right)-\left(\langle x_{k-1},x_{k-1}\rangle-2\langle x_k,x_{k-1}\rangle+\langle x_k,x_k\rangle\right)$$

$$2\lambda_k\left[F(x_k)-F(u)\right]\leq\|u-x_{k-1}\|^2-\|u-x_k\|^2-\|x_{k-1}-x_k\|^2$$

como ya vimos $\lambda_k\|y_k\|=\|x_{k-1}-x_k\|$*" entonces*

$$2\lambda_k\left[F(x_k)-F(u)\right]\leq\|u-x_{k-1}\|^2-\|u-x_k\|^2-\left(\lambda_k\|y_k\|\right)^2 \tag{0.4}$$

sumando ***(0.4)*** *para* $k=1,\ldots,n$*, obtenemos*

$$2\sum_{k=1}^{n}\lambda_k\left[F(x_k)-F(u)\right]\leq\|u-x_0\|^2-\|u-x_n\|^2-\sum_{k=1}^{n}\left(\lambda_k\|y_k\|\right)^2$$

$$2\sum_{k=1}^{n}\lambda_kF(x_k)-2\sum_{k=1}^{n}\lambda_kF(u)\leq\|u-x_0\|^2-\|u-x_n\|^2-\sum_{k=1}^{n}\left(\lambda_k\|y_k\|\right)^2 \tag{0.5}$$

multiplicando la desigualdad por (-1)

$$2\sigma_nF(u)-2\sum_{k=1}^{n}\lambda_kF(x_k)\geq\sum_{k=1}^{n}\lambda_k^2\|y_k\|^2+\|u-x_n\|^2-\|u-x_0\|^2 \tag{α}$$

por otro lado, sustituimos x_{k-1} *por* u *en* ***(0.3)****, "obtenemos*

$$F(x_{k-1})-F(x_k)\geq\left\langle\frac{x_{k-1}-x_k}{\lambda_k},x_{k-1}-x_k\right\rangle$$

multiplicando y dividiendo por λ_k

$$F(x_{k-1})-F(x_k)\geq\left\langle\frac{x_{k-1}-x_k}{\lambda_k},\frac{x_{k-1}-x_k}{\lambda_k}\right\rangle\cdot\lambda_k$$

$$F(x_{k-1}) - F(x_k) \geq \|y_k\|^2 \lambda_k \tag{0.6}$$

recordemos que $\sigma_k = \sum_{j=1}^{k} \lambda_j$, *para* $k \geq 1$". *Multiplicando* ***(0.6)*** *por* $\sigma_{k-1} = \sigma_k - \lambda_k$, *obtenemos*

$$\sigma_{k-1} F(x_{k-1}) - (\sigma_k - \lambda_k) F(x_k) \geq \sigma_{k-1} \|y_k\|^2 \lambda_k$$

$$\sigma_{k-1} F(x_{k-1}) - \sigma_k F(x_k) + \lambda_k F(x_k) \geq \sigma_{k-1} \|y_k\|^2 \lambda_k$$

Sumando la última desigualdad para $k = 1, \ldots, n$ *y notando que* $\sigma_0 = 0$, *obtenemos*

$$-\sigma_n F(x_n) + \sum_{k=1}^{n} \lambda_k F(x_k) \geq \sum_{k=2}^{n} \sigma_{k-1} \lambda_k \|y_k\|^2 \tag{β}$$

"*sumando dos veces* (β) *a* (α)

$$\begin{cases} -2\sigma_n F(x_n) + 2\sum_{k=1}^{n} \lambda_k F(x_k) \geq 2\sum_{k=2}^{n} \sigma_{k-1} \lambda_k \|y_k\|^2 \\ 2\sigma_n F(u) - 2\sum_{k=1}^{n} \lambda_k F(x_k) \geq \sum_{k=1}^{n} \lambda_k^2 \|y_k\|^2 + \|u - x_n\|^2 - \|u - x_0\|^2 \end{cases}$$

obtenemos

$$2\sigma_n (F(u) - F(x_n)) \geq 2\sum_{k=2}^{n} \sigma_{k-1} \lambda_k \|y_k\|^2 + \sum_{k=1}^{n} \lambda_k^2 \|y_k\|^2 + \|u - x_n\|^2 - \|u - x_0\|^2$$

$$2\sigma_n (F(u) - F(x_n)) \geq \left(\sum_{k=1}^{n} \lambda_k^2 + 2\sum_{k=2}^{n} \sigma_{k-1} \lambda_k \right) \|y_n\|^2 + \|u - x_n\|^2 - \|u - x_0\|^2.$$

$$2\sigma_n (F(u) - F(x_n)) \geq \sigma_n^2 \|y_n\|^2 + \|u - x_n\|^2 - \|u - x_0\|^2$$

por lo que tenemos

$$F(x_n) - F(u) \leq \frac{\|u - x_0\|^2}{2\sigma_n} - \frac{\|u - x_n\|^2}{2\sigma_n} - \frac{\sigma_n}{2} \|y_n\|^2$$

que es lo que queriamos probar". ■

El siguiente teorema describe las propiedades de convergencia del Algoritmo del Punto Proximal (PPA) bajo las suposiciones más generales.

Teorema 8.0.1 *Supongamos que $\{x_n\}_{n=0}^{\infty}$ es la sucesión generada por un PPA. Para cualquier $u \in \mathcal{H}$, se cumple la siguiente estimación global de convergencia:*

$$F(x_{n+1}) - F(u) \leq \frac{\|u - x_0\|^2}{2\sigma_n} \tag{0.7}$$

Por lo tanto, si $\sigma_n \longrightarrow \infty$, entonces [†] $F(x_n) \downarrow F^* = \inf_{x \in \mathcal{H}} F(x)$. *Si X^* no es vacío, entonces x_n converge débilmente a un minimizador de F. Además,*

$$F(x_{n+1}) - F^* \leq \frac{d(x_0, X^*)^2}{2\sigma_n} \tag{0.8}$$

Prueba 8.0.3 *"Desarrollando $y_k = \dfrac{x_k - x_{k+1}}{\lambda_k} \in \widehat{\partial} F(x_k)$, tenemos*

$$F(u) - F(x_{k+1}) \geq -\frac{1}{\lambda_k} \langle x_{k+1} - x_k, u - x_{k+1} \rangle$$

$$\lambda_k \left[F(x_{k+1}) - F(u) \right] \leq \langle x_{k+1} - x_k, u - x_{k+1} \rangle$$

$$2\lambda_k \left[F(x_{k+1}) - F(u) \right] \leq 2 \langle x_{k+1} - x_k, u - x_{k+1} \rangle$$

$$2\lambda_k \left[F(x_{k+1}) - F(u) \right] \leq 2 \langle x_{k+1}, u \rangle - 2 \langle x_{k+1}, x_{k+1} \rangle - 2 \langle x_k, u \rangle + 2 \langle x_k, x_{k+1} \rangle$$

dando forma, sumando y restando los terminos $\langle x_k, x_k \rangle$ y $\langle u, u \rangle$ tenemos:

$$2\lambda_k \left[F(x_{k+1}) - F(u) \right] \leq \langle x_k, x_k \rangle - 2 \langle x_k, u \rangle + \langle u, u \rangle$$
$$- \left(\langle x_{k+1}, x_{k+1} \rangle - 2 \langle x_{k+1}, u \rangle + \langle u, u \rangle \right) - \left(\langle x_{k+1}, x_{k+1} \rangle - 2 \langle x_{k+1}, x_k \rangle + \langle x_k, x_k \rangle \right)$$

[†] En $F(x_n) \downarrow F^*$ tenemos que $F(x_n)$ esta en $]F^*, +\infty[$ y convege a F^*.

lo que es igual a

$$2\lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_k - u\|^2 - \|x_k - x_{k+1}\|^2 - \|x_{k+1} - u\|^2$$

lo que a su vez es menor o igual a

$$2\lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_k - u\|^2 - \|x_k - x_{k+1}\|^2 - \|x_{k+1} - u\|^2 \leq \|x_k - u\|^2 - \|x_{k+1} - u\|^2$$

por lo que tenemos"

$$2\lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_k - u\|^2 - \|x_{k+1} - u\|^2$$

sumando para $k = 0, \ldots, n$, deducen que

$$2\sum_{k=0}^{n} \lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_0 - u\|^2 + \|x_1 - u\|^2 + \|x_2 - u\|^2 + \cdots$$
$$\cdots - \|x_1 - u\|^2 - \|x_2 - u\|^2 - \cdots - \|x_{n+1} - u\|^2$$

simplificando, obtenemos lo siguiente

$$2\sum_{k=0}^{n} \lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_0 - u\|^2 - \|x_{n+1} - u\|^2$$

"lo cual es menor o igual a

$$2\sum_{k=0}^{n} \lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_0 - u\|^2 - \|x_{n+1} - u\|^2 \leq \|x_0 - u\|^2$$

$$2\sum_{k=0}^{n} \lambda_k \left[F(x_{k+1}) - F(u)\right] \leq \|x_0 - u\|^2$$

para cada $n \in \mathbb{N}$. Dado que la sucesión $\{F(x_n)\}$ no es creciente, deducimos que

$$2\sigma_n \left[F(x_{n+1}) - F(u)\right] \leq \|x_0 - u\|^2$$

*así, se puede obtener **(0.7)**"*

$$F(x_{n+1}) - F(u) \leq \frac{\|u - x_0\|^2}{2\sigma_n}$$

*La fórmula **(0.8)** se obtiene al reemplazar x^* por u en **(0.7)**, donde x^* es el punto de X^* más próximo a x.* ■

Observación 8.0.1 *La condición $\sigma_n \longrightarrow \infty$ es la más débil necesaria para garantizar que $F(x_n) \downarrow F^*$. Si $\sigma_n \longrightarrow \sigma < \infty$, entonces x_n siempre converge de manera fuerte:*

$$\|x_{n+p} - x_n\| \leq \|x_n - x_{n+1}\| + \|x_{n+1} - x_{n+2}\| + \cdots + \|x_{n+p-1} - x_{n+p}\|$$

$$\|x_{n+p} - x_n\| \leq \sum_{j=n+1}^{n+p} \|x_{j-1} - x_j\|$$

como ya hemos visto $y_i = \dfrac{x_{j-1} - x_j}{\lambda_j}$, entonces, $\lambda_j \|y_j\| = \|x_{j-1} - x_j\|$

$$\|x_{n+p} - x_n\| \leq \sum_{j=n+1}^{n+p} \lambda_j \|y_j\|$$

*como la sucesión $\|y_j\|$ es decreciente por el **Lema (8.0.1)**, tenemos*

$$\sum_{j=n+1}^{n+p} \lambda_j \|y_j\| \leq \sum_{j=n+1}^{n+p} \lambda_j \|y_{j-1}\|$$

por lo que tenemos

$$\|x_{n+p} - x_n\| \leq \sum_{j=n+1}^{n+p} \lambda_j \|y_{j-1}\| \tag{0.9}$$

ya que $\sum_{j=n+1}^{n+p} \lambda_j = \sigma_n \longrightarrow \sigma$, *la desigualdad* ***(0.9)*** *muestra que* x_n *es una sucesión de Cauchy, y por lo que tenemos entonces, que converge de manera fuerte a algún punto* x^∞, *incluso si* F *no tiene un argumento minimizador. Aún cuando* $X^* \neq \emptyset$, *tenemos*

$$\|x - x^\infty\| \leq \sum_{j=1}^{\infty} \|x_{j-1} - x_j\| = \sum_{j=1}^{\infty} \lambda_j \|y_j\|$$

$$\|x - x^\infty\| \leq \sigma \|y_1\| \tag{0.10}$$

de manera que $d(x^\infty, X^*) \geq d(x, X^*) - \|x - x^\infty\| \geq d(x, X^*) - \sigma\|y_1\|$. *Si* σ *es pequeña, entonces* $d(x^\infty, X^*) > 0$, *y* x^∞ *no pertenece a* X^*.

Al hacer el cambio x_0 en lugar de u en la desigualdad del **Lema (8.0.2)**, obtenemos el siguiente resultado.

Corolario 8.0.1 *En un algoritmo proximal, se verifica la siguiente estimación:*

$$F(x_n) \leq F(x_0) - \frac{\|x_0 - x_n\|^2}{2\sigma_n}$$

Teorema 8.0.2 *Sea* $A = \partial F$ *y* $u \in D(A)$. *En el contexto de un algoritmo de punto proximal, se verifican las siguientes desigualdades:*

$$\|y_n\| \leq \frac{\|x_n - x_0\|}{\sigma_n} \tag{0.11}$$

$$\|y_n\| \leq \|A^0 u\| + \frac{\|u - x_0\|}{\sigma_n} \tag{0.12}$$

Prueba 8.0.4 *La desigualdad* ***(0.11)*** *se obtiene al sustituir* x_n *por* u *en la desigualdad del* ***Lema (8.0.2)***

$$F(x_n) - F(x_n) \leq \frac{\|x_n - x_0\|^2}{2\sigma_n} - \frac{\|x_n - x_n\|^2}{2\sigma_n} - \frac{\sigma_n}{2}\|y_n\|^2$$

$$0 \leq \frac{\|x_n - x_0\|^2}{2\sigma_n} - 0 - \frac{\sigma_n}{2}\|y_n\|^2$$

$$\|y_n\| \leq \frac{\|x_n - x_0\|}{\sigma_n}$$

De la convexidad de F, podemos obtener [‡]

$$F(x_n) \geq F(u) + \langle A^0 u, x_n - u\rangle$$

$$F(x_n) - F(u) \geq \langle A^0 u, x_n - u\rangle$$

usando la desigualdad de Cauchy-Schwarz [§] *al lado derecho de la desigualdad implica que*

$$F(x_n) - F(u) \geq -\|A^0 u\| \cdot \|x_n - u\|$$

"de la desigualdad del ***Lema (8.0.2)***

$$-\|A^0 u\| \cdot \|x_n - u\| \leq F(x_n) - F(u) \leq \frac{\|u - x_0\|^2}{2\sigma_n} - \frac{\|u - x_n\|^2}{2\sigma_n} - \frac{\sigma_n}{2}\|y_n\|^2$$

$$-2\sigma_n\|A^0 u\| \cdot \|x_n - u\| \leq \|u - x_0\|^2 - \|u - x_n\|^2 - \sigma_n^2\|y_n\|^2$$

$$\sigma_n^2\|y_n\|^2 \leq \|u - x_0\|^2 - \|u - x_n\|^2 + 2\sigma_n\|A^0 u\| \cdot \|x_n - u\|$$

[‡]Si $A(x) \neq \emptyset$, entonces $A^0(x)$ denota el elemento de norma mínima de $A(x)$. Esto significa que $A^0(x)$ es el elemento en $A(x)$ tal que su norma $\|A^0(x)\|$ es la más pequeña posible entre todos los elementos de $A(x)$.

[§]Por la desigualdad de Cauchy-Schwarz:

$$\langle A^0 u, x_n - u\rangle \geq -\|A^0 u\| \cdot \|x_n - u\|$$

usanado la desigualdad de Young[¶]*"*

$$\sigma_n^2\|y_n\|^2 \leq \|u-x_0\|^2 - \|u-x_n\|^2 + (\sigma_n^2\|A^0u\|^2 + \|x_n-u\|^2)$$

$$\sigma_n^2\|y_n\|^2 \leq \|u-x_0\|^2 + \sigma_n^2\|A^0u\|^2$$

$$\sigma_n^2\|y_n\|^2 \leq \|u-x_0\|^2 + \sigma_n^2\|A^0u\|^2 \leq \|u-x_0\|^2 + 2\|u-x_0\|\cdot\sigma_n\|A^0u\| + \sigma_n^2\|A^0u\|^2$$

$$\sigma_n^2\|y_n\|^2 \leq \big(\|u-x_0\| + \sigma_n\|A^0u\|\big)^2$$

lo que implica

$$\sigma_n\|y_n\| \leq \|u-x_0\| + \sigma_n\|A^0u\|$$

entonces tenemos

$$\|y_n\| \leq \|A^0u\| + \frac{\|u-x_0\|}{\sigma_n}$$

y así, el teorema queda demostrado. ■

[¶]"Desigualdad de Young: Sean $p,q>1$ tales que $\dfrac{1}{p}+\dfrac{1}{q}=1$, y sean $a,b\geq 0$". Entonces

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

teniendo en cuenta: $a=\sigma_n\|A^0u\|$ y $b=\|x_n-u\|$ tenemos:

$$\sigma_n\|A^0u\|\cdot\|x_n-u\| \leq \frac{\big(\sigma_n\|A^0u\|\big)^2}{2} + \frac{(\|x_n-u\|)^2}{2}$$

9 Aplicación del algoritmo en Phyton

En este capítulo, exploraremos la implementación y aplicación del algoritmo del punto proximal utilizando el lenguaje de programación Python. Python es ampliamente utilizado en la comunidad de científicos de datos y en la industria debido a su simplicidad y la extensa colección de bibliotecas disponibles para el análisis y la optimización.

Primero, presentaremos una breve introducción a las herramientas y bibliotecas de Python que se utilizarán en esta implementación. A continuación, se detallará la estructura del código y se proporcionarán ejemplos prácticos de cómo aplicar el algoritmo del punto proximal a problemas de optimización reales. Finalmente, analizaremos los resultados obtenidos y discutiremos posibles extensiones y mejoras al método.

9.1 Herramientas y Bibliotecas

En esta sección, introducimos las bibliotecas de Python necesarias para implementar el algoritmo del punto proximal. Algunas de las bibliotecas clave incluyen:

- **NumPy**: Para operaciones matemáticas y manipulación de matrices.
- **SciPy**: Para optimización y otras funciones científicas.
- **Matplotlib**: Para visualización de datos y gráficos.

9.2 Ejemplos Prácticos

Aquí, proporcionamos ejemplos detallados de cómo aplicar el algoritmo a problemas específicos de optimización, mostrando el código completo y explicando cada parte del proceso.

9.2.1 PPA para la Minimización de Funciones Absolutas

"Se aplica el Algoritmo del Punto Proximal para minimizar la función objetivo valor absoluto $F(x) =\mid ax + b \mid$. Este algoritmo es especialmente útil para resolver problemas de minimización donde la función objetivo no es diferenciable".

- **Implementación del Algoritmo**

  A continuación se presenta el código en Python utilizado para implementar el Algoritmo del Punto Proximal:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize

# Definir la función objetivo F(x) = |ax + b|
def F(x, a, b):
    return np.abs(a * x + b)

# Definir la función proximal g(y) = F(y) + (1/(2*lambda)) * (x - y)^2
def g(y, x, a, b, lambda_):
    return F(y, a, b) + (1 / (2 * lambda_)) * (x - y)**2

# Implementar el Algoritmo del Punto Proximal
```

```
def proximal_point_algorithm(x0, a, b, lambda_, tol=1e-6, max_iter=100):
    x_n = x0
    trajectory = [x_n]
    for _ in range(max_iter):
        result = minimize(g, x0=x_n, args=(x_n, a, b, lambda_),
            method='BFGS')
        x_n1 = result.x[0]
        trajectory.append(x_n1)
        if np.linalg.norm(x_n1 - x_n) < tol:
            break
        x_n = x_n1
    return x_n, trajectory

# Parámetros de la función y del algoritmo
a = 1
b = -1
x0 = 3.5
lambda_ = 0.3
max_iter = 100
tol = 1e-6

# Ejecutar el algoritmo del punto proximal
x_min, trajectory = proximal_point_algorithm(x0, a, b, lambda_, tol,
    max_iter)

# Graficar la trayectoria
```

```
x_values = np.linspace(-5, 5, 400)
y_values = F(x_values, a, b)


plt.figure(figsize=(10, 6))
plt.plot(x_values, y_values, label='$F(x) = |ax + b|$')
plt.plot(trajectory, F(np.array(trajectory), a, b), 'ro-',
    label='Trayectoria Proximal')
plt.xlabel('x')
plt.ylabel('F(x)')
plt.title('Algoritmo del Punto Proximal para $F(x) = |ax + b|$')
plt.legend()
plt.grid(False)
plt.show()


# Imprimir resultados
print(f'Minimizador: x = {x_min}')
print(f'Trayectoria: {trajectory}')
```

- **Parámetros del Algoritmo:**

  Los parámetros de la función son $a = 1$ y $b = -1$, el punto inicial es $x_0 = 3.5$ y el parámetro de regularización es $\lambda = 0.3$.

- **Resultados:**

  El resultado del algoritmo incluye el minimizador de la función y la trayectoria seguida por el algoritmo. Estos resultados se imprimen en la consola y se visualizan en el gráfico generado.

```
Minimizador: x = 1.0000006316175676
```

```
Trayectoria: [2.0, 1.8999999925350972, 1.7999999850731778,
    1.699999977313681, 1.5999999702001944, 1.499999963131728,
    1.3999999559252851, 1.2999999482834192, 1.1999999406117943,
    1.099999933377657, 1.0000006316175676, 0.9999999986453384]
```

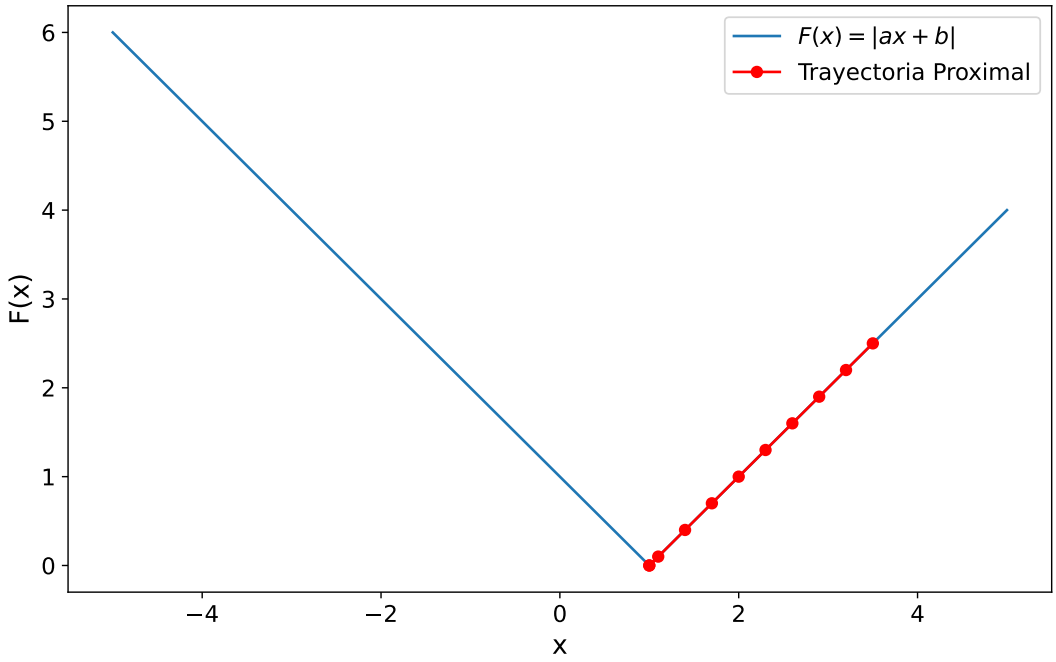


Figura 9.1: Se ejecuta el algoritmo para $F(x) =\mid ax + b \mid$ y se obtiene la grafica de la trayectoria de puntos generados durante la minimización.

A continuación tenemos el gráfico que muestra la velocidad de convergencia del Algoritmo del Punto Proximal para la función $F(x) =\mid ax + b \mid$ con diferentes valores de (λ). El gráfico muestra cómo varía el valor de x a lo largo de las iteraciones para cada valor de (λ).

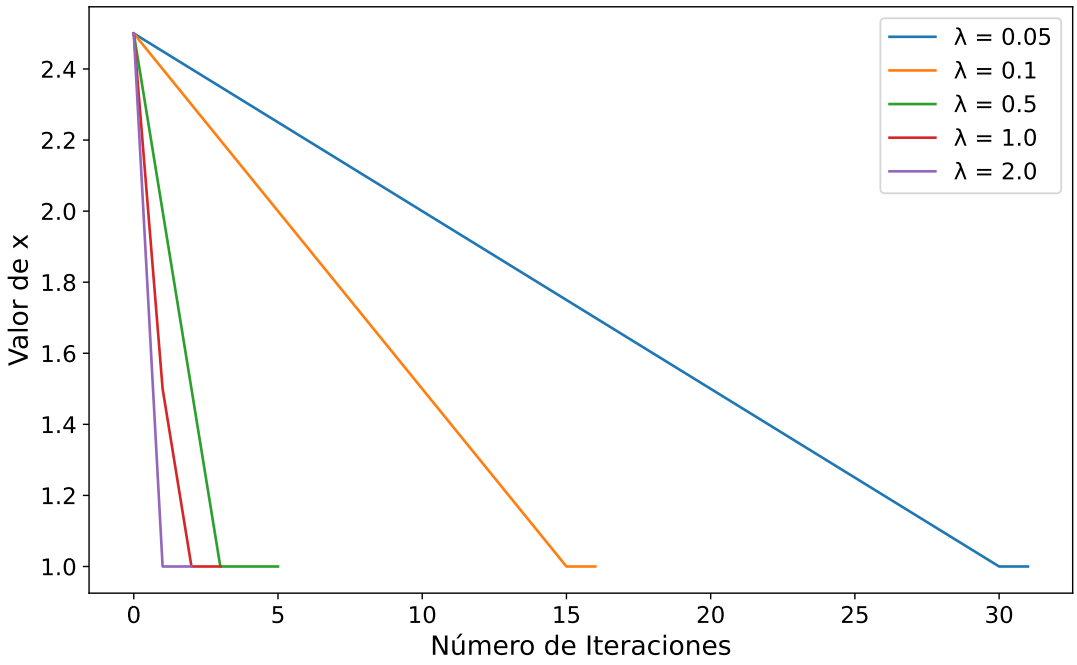


Figura 9.2: Velocidad de Convergencia para diferentes parámetros de λ, de la función $F(x) =\mid ax + b \mid$.

- Parámetro pequeño ($\lambda = 0.05$): Converge lentamente.
- Parámetro moderado ($\lambda = 0.1$): Converge más rápidamente.
- Parámetro grande ($\lambda = 2.0$): Converge rápidamente, pero puede ser inestable dependiendo de la función.

9.2.2 PPA para la Minimización de Funciones a Tramos

"Se aplica el Algoritmo del Punto Proximal para minimizar la función objetivo a tramos:

$$F(x) = \begin{cases} 1 - x & \text{si } x < 1 \\ x^2 - 1 & \text{si } x \geq 1 \end{cases}$$

este algoritmo es especialmente útil para resolver problemas de minimización donde la función objetivo puede no ser diferenciable en ciertos puntos".

- **Implementación del Algoritmo**

  A continuación se presenta el código en Python utilizado para implementar el Algoritmo del Punto Proximal:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize

# Definir la función objetivo F(x)
def F(x):
    if x < 1:
        return 1 - x
    else:
        return x**2 - 1

# Definir la función proximal g(y) = F(y) + (1/(2*lambda)) * (x - y)^2
def g(y, x, lambda_):
    if y < 1:
        return 1 - y + (1 / (2 * lambda_)) * (x - y)**2
    else:
        return y**2 - 1 + (1 / (2 * lambda_)) * (x - y)**2

# Implementar el Algoritmo del Punto Proximal
def proximal_point_algorithm(x0, lambda_, tol=1e-6, max_iter=100):
    x_n = x0
```

```
    trajectory = [x_n]
    for _ in range(max_iter):
        result = minimize(g, x0=x_n, args=(x_n, lambda_), method='BFGS')
        x_n1 = result.x[0]
        trajectory.append(x_n1)
        if np.linalg.norm(x_n1 - x_n) < tol:
            break
        x_n = x_n1
    return x_n, trajectory

# Parámetros del algoritmo
x_values = np.linspace(-1, 3, 400)
lambda_ = 0.05
x0 = 2.5
max_iter = 100
tol = 1e-6

# Ejecutar el algoritmo del punto proximal
x_min, trajectory = proximal_point_algorithm(x0, lambda_, tol, max_iter)

# Graficar la trayectoria
y_values = [F(x) for x in x_values]

plt.figure(figsize=(10, 6))
plt.plot(x_values, y_values, label='$F(x)$')
plt.plot(trajectory, [F(x) for x in trajectory], 'ro-',
```

```
    label='Trayectoria Proximal')
plt.xlabel('x')
plt.ylabel('F(x)')
plt.title('Algoritmo del Punto Proximal para la Función a Tramos
    $F(x)$')
plt.legend()
plt.grid(True)
plt.show()

# Imprimir resultados
print(f'Minimizador: x = {x_min}')
print(f'Trayectoria: {trajectory}')
```

- **Parámetros del Algoritmo**

  Los parámetros incluyen la tasa de regularización $\lambda = 0.05$, el punto inicial $x_0 = 2.5$, un máximo de iteraciones $max_iter = 100$, y una tolerancia $tol = 1 \times 10^{-6}$.

- **Resultados**

  El resultado del algoritmo incluye el minimizador de la función y la trayectoria seguida por el algoritmo. Estos resultados se imprimen en la consola y se visualizan en el gráfico generado.

```
Minimizador: x = 0.9999999965253955
Trayectoria: [2.5, 2.272727274206066, 2.0661157032480464,
    1.8782870023136418, 1.707533630198212, 1.5523032919899995,
    1.4111848026126241, 1.2828952670377594, 1.166268417649464,
    1.0602440084188274, 0.9999999965253955, 0.9999999965253955]
```

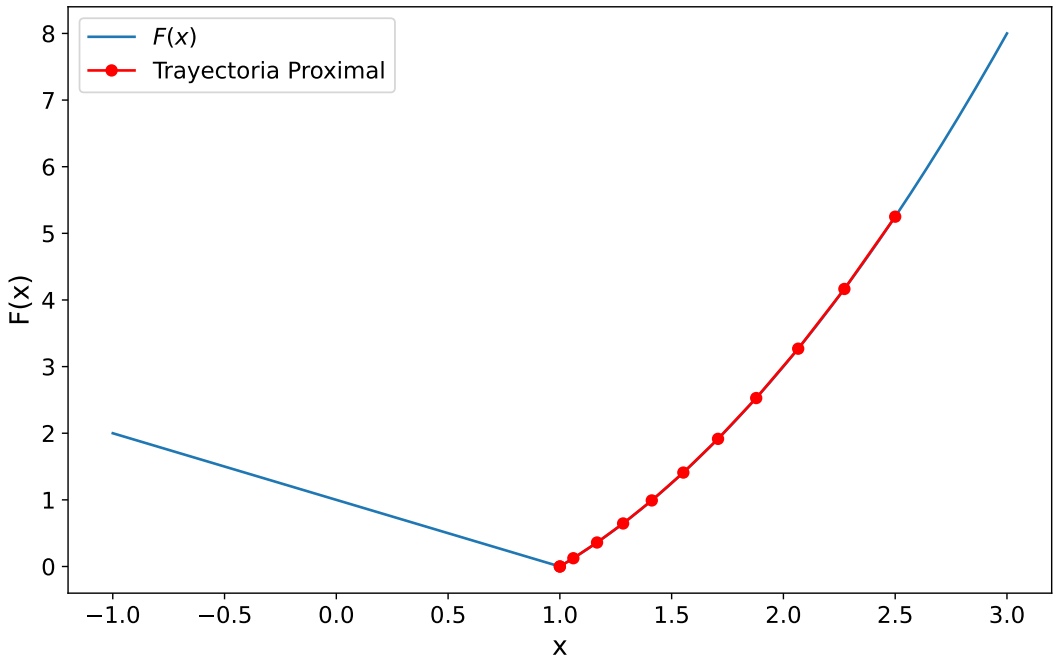


Figura 9.3: Se ejecuta el algoritmo para la función a tramos $F(x)$ y se obtiene la grafica de la trayectoria de puntos generados durante la minimización.

"Así tenemos el gráfico que muestra la velocidad de convergencia del Algoritmo del Punto Proximal para la función

$$F(x) = \begin{cases} 1 - x & \text{si } x < 1 \\ x^2 - 1 & \text{si } x \geq 1 \end{cases}$$

con diferentes valores de (λ). El gráfico muestra cómo varía el valor de x a lo largo de las iteraciones para cada valor de (λ)".

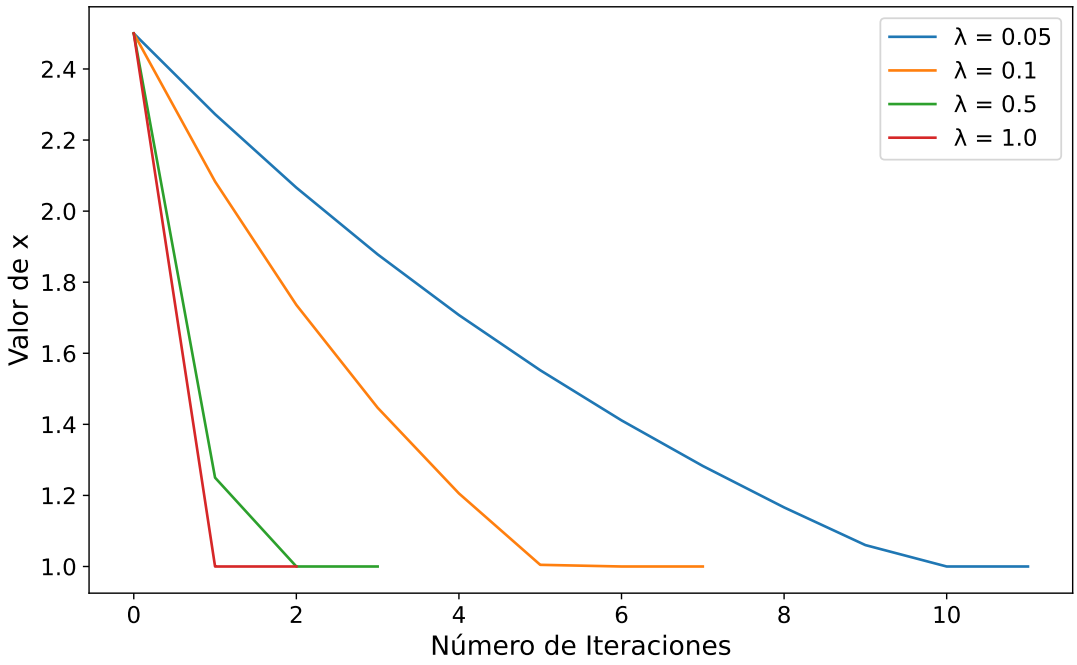


Figura 9.4: Velocidad de Convergencia para diferentes parámetros de λ, de la función a tramos $F(x)$.

- Parámetro pequeño ($\lambda = 0.05$): Converge lentamente.
- Parámetro moderado ($\lambda = 0.1$): Converge más rápidamente.
- Parámetro grande ($\lambda = 1.0$): Converge rápidamente.

9.2.3 PPA para la Minimización de Funciones Logaritmicas

"En este ejemplo, se aplicará el Algoritmo del Punto Proximal para minimizar la función objetivo $F(x,y) = \log(1 + x^2 + y^2)$. Este método es especialmente adecuado para funciones cuasiconvexas, como la función objetivo que se analiza en este ejemplo. El objetivo es encontrar el punto (x,y) que minimice la función $F(x,y)$. El Algoritmo del Punto Proximal logra esto al iterar sobre sucesivos puntos (x_n,y_n) generados mediante la minimización de una función proximal que incluye tanto la función objetivo original como un término de penalización cuadrática".

- **Implementación del Algoritmo**

  Se presenta el código en Python utilizado para implementar el Algoritmo del Punto Proximal:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize

# Definir la función objetivo F(x, y)
def F(x, y):
    return np.log(1 + x**2 + y**2)

# Definir la función proximal
# g(z) = F(z) + (1/(2*lambda)) * ||(x, y) - z||^2
def g(z, x, y, lambda_):
    return F(z[0], z[1]) + (1 / (2 * lambda_)) * ((x - z[0])**2 +
        (y - z[1])**2)

# Implementar el Algoritmo del Punto Proximal
def proximal_point_algorithm(x0, y0, lambda_, tol=1e-6, max_iter=100):
    x_n, y_n = x0, y0
    trajectory = [(x_n, y_n)]
    for _ in range(max_iter):
        result = minimize(g, x0=[x_n, y_n], args=(x_n, y_n, lambda_),
            method='BFGS')
        x_n1, y_n1 = result.x
        trajectory.append((x_n1, y_n1))
```

```
        if np.linalg.norm([x_n1 - x_n, y_n1 - y_n]) < tol:
            break
        x_n, y_n = x_n1, y_n1
    return (x_n, y_n), trajectory

# Parámetros del algoritmo
x0, y0 = 2.0, 2.0
lambda_ = 0.5
max_iter = 100
tol = 1e-6

# Ejecutar el algoritmo del punto proximal
min_point, trajectory = proximal_point_algorithm(x0, y0, lambda_, tol,
    max_iter)

# Crear una malla para graficar la función
x_values = np.linspace(-3, 3, 100)
y_values = np.linspace(-3, 3, 100)
X, Y = np.meshgrid(x_values, y_values)
Z = np.log(1 + X**2 + Y**2)

# Graficar la superficie de la función objetivo
fig = plt.figure(figsize=(12, 8))
ax = fig.add_subplot(111, projection='3d')
ax.plot_surface(X, Y, Z, cmap='viridis', alpha=0.6)
```

```
# Graficar la trayectoria del algoritmo del punto proximal
traj_x = [point[0] for point in trajectory]
traj_y = [point[1] for point in trajectory]
traj_z = [F(x, y) for x, y in trajectory]

ax.plot(traj_x, traj_y, traj_z, 'ro-', label='Trayectoria Proximal')
ax.set_xlabel('X')
ax.set_ylabel('Y')
ax.set_zlabel('F(x, y)')
ax.set_title('Algoritmo del Punto Proximal para
    $F(x, y) = \log(1 + x^2 + y^2)$')
plt.legend()
plt.show()

# Imprimir resultados
print(f'Minimizador: (x, y) = {min_point}')
print(f'Trayectoria: {trajectory}')
```

- **Parámetros del Algoritmo**

  Los parámetros incluyen la tasa de regularización $\lambda = 0.5$, el punto inicial $(x_0,y_0) = (2.0,2.0)$, un máximo de iteraciones $max_iter = 100$, y una tolerancia $tol = 1 \times 10^{-6}$.

- **Resultados**

  El resultado del algoritmo incluye el minimizador de la función y la trayectoria seguida por el algoritmo. Estos resultados se imprimen en la consola y se visualizan en el gráfico generado.

```
Minimizador: (x, y) =  (-2.94800399844182e-06, -1.9720746130889366e-06)
```

```
Trayectoria: [(2.0, 2.0), (valores intermedios),
     (-2.94800399844182e-06, -1.9720746130889366e-06)]
```

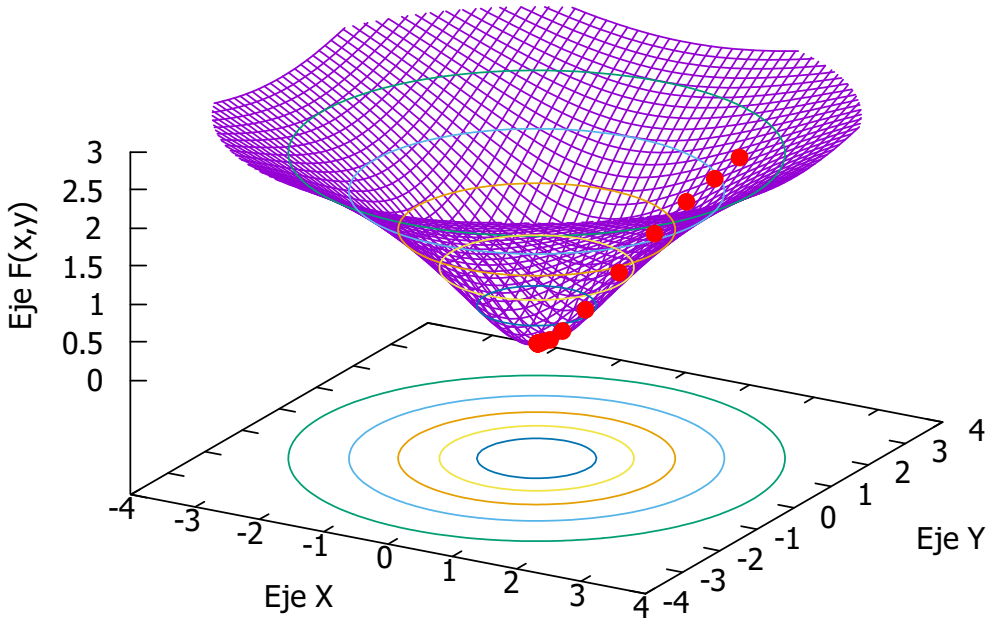


Figura 9.5: Se ejecuta el algoritmo para la función Logaritmica $F(x,y)$ y se obtiene la gráfica de la trayectoria de puntos generados durante la minimización.

A continuación tenemos el gráfico que muestra la velocidad de convergencia del Algoritmo del Punto Proximal para la función logaritmica $F(x,y) = \log(1 + x^2 + y^2)$ con diferentes valores de (λ). El gráfico muestra cómo varía el valor de (x,y) a lo largo de las iteraciones para cada valor de (λ).

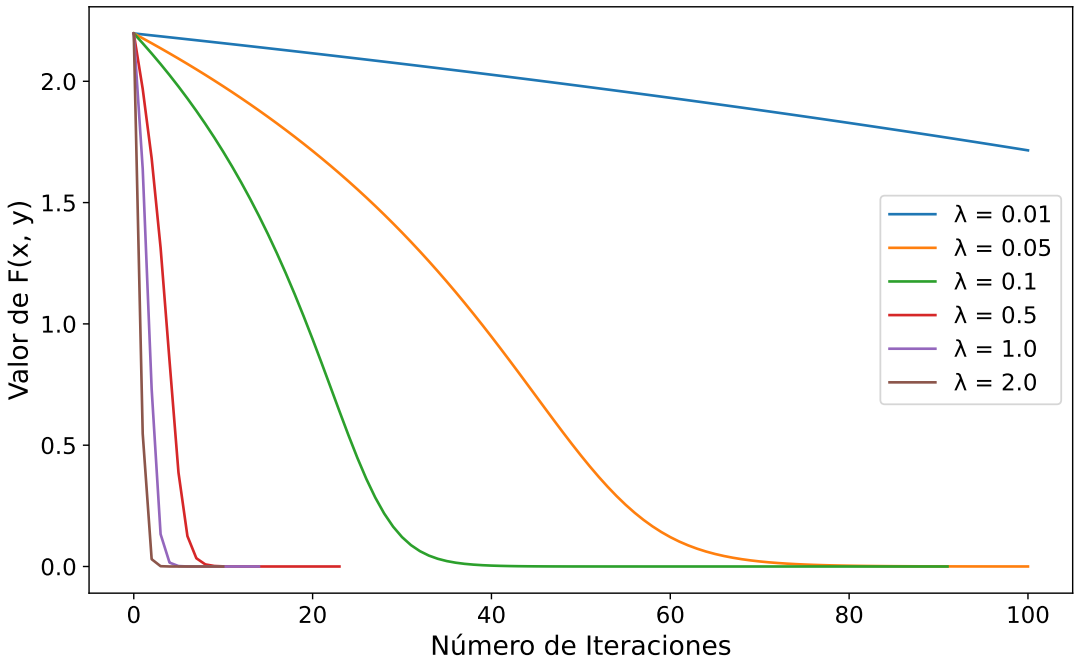


Figura 9.6: Velocidad de Convergencia para diferentes parámetros de λ, de la función logaritmica $F(x,y)$.

- Parámetro pequeño ($\lambda = 0.01$): Converge muy lentamente.

  Iteraciones: 100

- Parámetro moderado ($\lambda = 0.1$): Converge más rápidamente.

  Iteraciones: 91

- Parámetro grande ($\lambda = 2.0$): Converge rápidamente.

  Iteraciones: 10

9.2.4 PPA para la Minimización de Función Suma de la norma euclidiana y el máximo

"En este ejemplo, se aplicará el Algoritmo del Punto Proximal para minimizar la función objetivo $F(x, y) = \sqrt{x^2 + y^2} + \max\{x, y\}$. El Algoritmo del Punto Proximal es una técnica iterativa utilizada para resolver problemas de minimización de funciones, especialmente aque-

llas que pueden no ser diferenciables en ciertos puntos. Este método es particularmente útil en la optimización para este tipo de funciones y no lineales. Esta función combina una norma euclidiana con la función máxima, lo que la hace no diferenciable en ciertos puntos. Sin embargo, su estructura permite aplicar el Algoritmo del Punto Proximal de manera efectiva".

- **Implementación del Algoritmo**

  Se presenta el código en Python utilizado para implementar el Algoritmo del Punto Proximal:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import minimize

# Definir la función objetivo F(x, y)
def F(x):
    return np.sqrt(x[0]**2 + x[1]**2) + np.maximum(x[0], x[1])

# Definir la función proximal g(z) = F(z) + (1/(2*lambda)) * ||x - z||^2
def g(z, x, lambda_):
    return F(z) + (1 / (2 * lambda_)) * np.sum((x - z)**2)

# Implementar el Algoritmo del Punto Proximal
def proximal_point_algorithm(x0, lambda_, tol=1e-6, max_iter=100):
    x_n = x0
    trajectory = [x_n]
    for _ in range(max_iter):
        result = minimize(g, x0=x_n, args=(x_n, lambda_), method='BFGS')
        x_n1 = result.x
```

```
        trajectory.append(x_n1)
        if np.linalg.norm(x_n1 - x_n) < tol:
            break
        x_n = x_n1
    return x_n, trajectory

# Parámetros del algoritmo
x0 = np.array([2.0, 2.0])
lambda_ = 0.1
max_iter = 100
tol = 1e-6

# Ejecutar el algoritmo del punto proximal y obtener la trayectoria
min_point, trajectory = proximal_point_algorithm(x0, lambda_, tol,
    max_iter)

# Crear una malla para graficar la función
x_values = np.linspace(-3, 3, 100)
y_values = np.linspace(-3, 3, 100)
X, Y = np.meshgrid(x_values, y_values)
Z = np.sqrt(X**2 + Y**2) + np.maximum(X, Y)

# Graficar la superficie de la función objetivo y la trayectoria
fig = plt.figure(figsize=(12, 8))
ax = fig.add_subplot(111, projection='3d')
ax.plot_surface(X, Y, Z, cmap='viridis', alpha=0.6)
```

```
# Graficar la trayectoria del algoritmo del punto proximal
trajectory = np.array(trajectory)
traj_z = [F(x) for x in trajectory]
ax.plot(trajectory[:, 0], trajectory[:, 1], traj_z, 'ro-',
    label='Trayectoria Proximal')

ax.set_xlabel('X')
ax.set_ylabel('Y')
ax.set_zlabel('F(x, y)')
ax.set_title('Trayectoria del Algoritmo del Punto Proximal')
plt.legend()
plt.show()

# Imprimir resultados
print(f'Minimizador: (x, y) = {min_point}')
print(f'Trayectoria: {trajectory}')
```

- **Parámetros del Algoritmo**

  Los parámetros incluyen la tasa de regularización $\lambda = 0.1$, el punto inicial $(x_0, y_0) = (2.0, 2.0)$, un máximo de iteraciones $max_iter = 100$, y una tolerancia $tol = 1 \times 10^{-6}$.

- **Resultados**

  El gráfico generado muestra la trayectoria del Algoritmo del Punto Proximal en el espacio tridimensional. La función objetivo $F(x, y) = \sqrt{x^2 + y^2} + \text{máx}\{x, y\}$ se representa como una superficie, y la trayectoria del algoritmo se muestra en rojo, indicando cómo converge al punto de mínimo.

```
Minimizador: (x, y) = (-7.62328830e-09, -7.62329811e-09)
Trayectoria: [(2.0, 2.0), (valores intermedios),
     (-7.62328830e-09, -7.62329811e-09)]
```

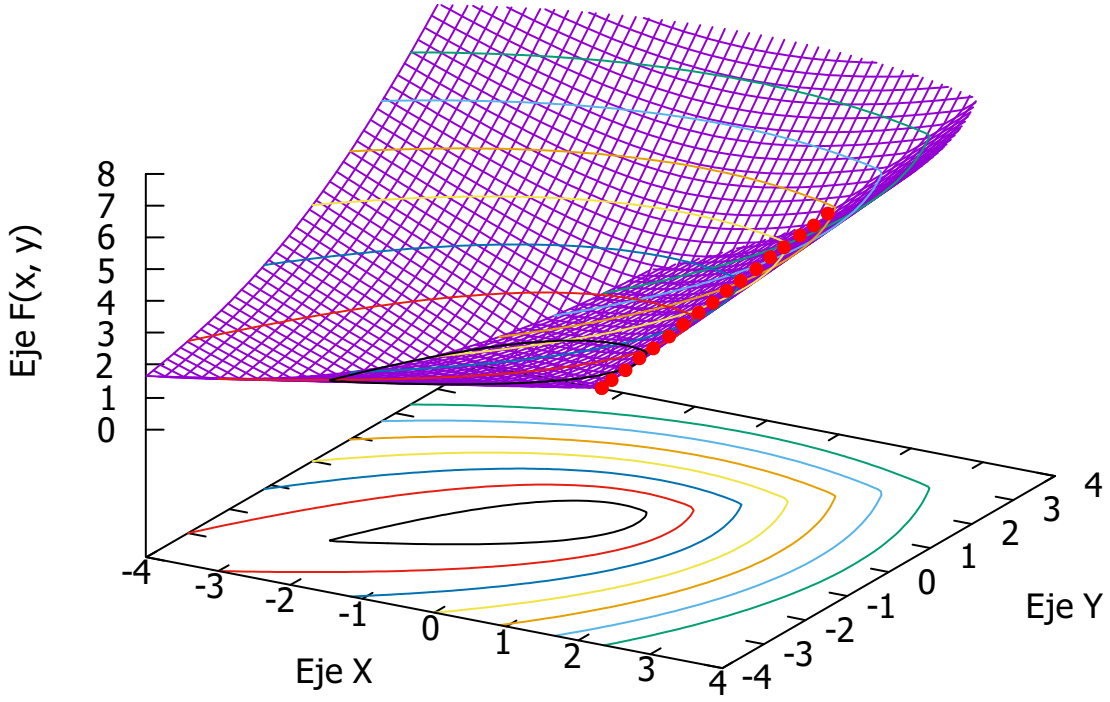


Figura 9.7: Se ejecuta el algoritmo para la función $F(x,y)$ y se obtiene la gráfica de la trayectoria de puntos generados durante la minimización.

A continuación tenemos el gráfico que muestra la velocidad de convergencia del Algoritmo del Punto Proximal para la función $F(x,y) = \sqrt{x^2 + y^2} + \text{máx}\{x, y\}$ con diferentes valores de (λ). El gráfico muestra cómo varía el valor de (x,y) a lo largo de las iteraciones para cada valor de (λ).

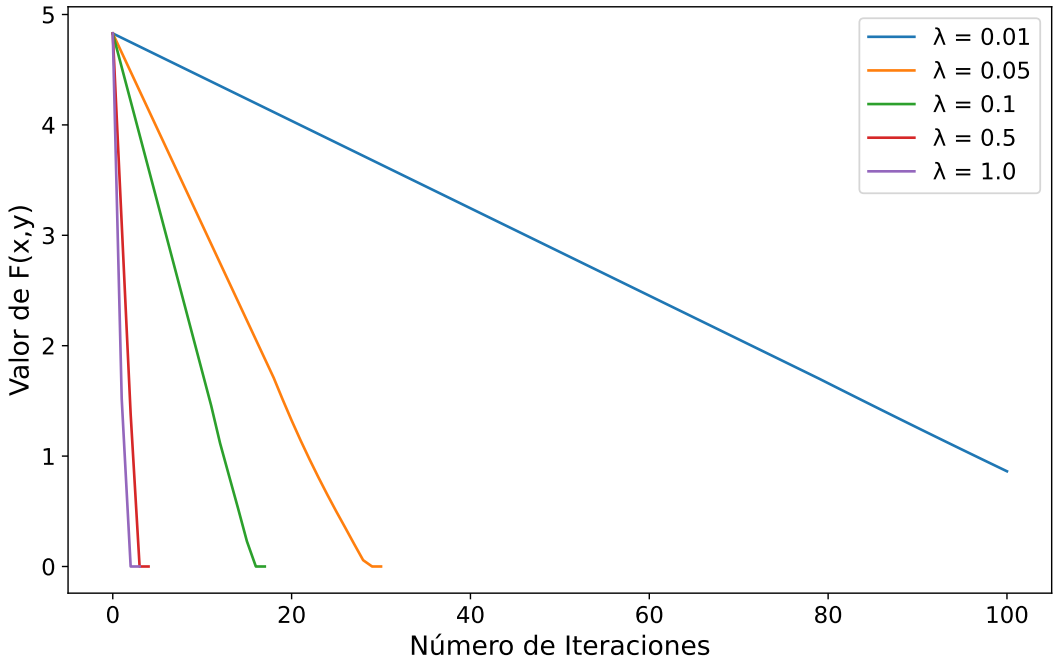


Figura 9.8: Velocidad de Convergencia para diferentes parámetros de λ, de la función Cuasiconvexa $F(x,y)$.

- Parámetro pequeño ($\lambda = 0.01$): Converge muy lentamente.

  Iteraciones: 100

- Parámetro moderado ($\lambda = 0.1$): Converge más rápidamente.

  Iteraciones: 17

- Parámetro grande ($\lambda = 1.0$): Converge rápidamente.

  Iteraciones: 3

10 Conclusiones y Recomendaciones

10.1 Conclusiones

1. **Convergencia del Algoritmo de Punto Proximal**: Se ha demostrado que el algoritmo de punto proximal converge de manera efectiva en espacios de Hilbert cuando se aplican las condiciones adecuadas. La implementación en Python ha permitido validar teóricamente estos resultados, proporcionando un marco práctico para su aplicación.

2. **Eficiencia y Aplicabilidad**: La simulación de diferentes escenarios en Python mostró que el algoritmo no solo es teóricamente sólido, sino también eficiente en la práctica. Esto se evidencia en la capacidad del algoritmo para manejar funciones convexas y otros tipos de funciones dentro de un rango de parámetros definidos.

3. **Flexibilidad del Algoritmo**: Una de las conclusiones más destacadas es la flexibilidad del algoritmo de punto proximal. Puede ser adaptado a una variedad de problemas de optimización, lo que lo hace una herramienta valiosa tanto en el ámbito académico como en el industrial.

4. **Impacto de los Parámetros de Regularización**: Los parámetros de regularización juegan un papel crucial en la convergencia del algoritmo. La elección adecuada de estos parámetros puede acelerar significativamente el proceso de convergencia, mientras que una elección inapropiada puede llevar a una convergencia lenta o incluso a la no convergencia.

5. **Desempeño Computacional**: La implementación en Python, utilizando bibliotecas como NumPy y SciPy, demostró ser efectiva para la simulación y validación de resultados.

Sin embargo, para aplicaciones a gran escala, se recomienda considerar optimizaciones adicionales o el uso de lenguajes de programación de más bajo nivel.

10.2 Recomendaciones

1. **Optimización de la Implementación**: A pesar de la efectividad de la implementación en Python, se sugiere explorar otras implementaciones en lenguajes como C++ o Fortran para mejorar el rendimiento computacional en problemas de gran escala.

2. **Estudio de Parámetros de Regularización**: Se aconseja realizar estudios más detallados sobre la influencia de los parámetros de regularización en la convergencia del algoritmo. Desarrollar métodos automáticos para la selección de estos parámetros podría ser una línea de investigación futura prometedora.

3. **Aplicaciones en Problemas Reales**: Se debe fomentar la aplicación del algoritmo en problemas reales del ámbito de la ingeniería, economía, y ciencias sociales. Colaborar con expertos en estos campos puede proporcionar nuevas perspectivas y retos para el algoritmo, enriqueciendo así su desarrollo y aplicación.

4. **Desarrollo de Herramientas de Software**: Crear herramientas de software de código abierto que implementen el algoritmo de punto proximal con interfaces amigables puede facilitar su uso por parte de investigadores y profesionales de diversas disciplinas.

5. **Educación y Capacitación**: Finalmente, se recomienda incorporar el estudio del algoritmo de punto proximal y su implementación en Python en los currículos de programas de matemática aplicada y ciencias de la computación. Esto ayudará a preparar a la próxima generación de científicos y profesionales en el uso de herramientas avanzadas de optimización.

Estas conclusiones y recomendaciones buscan no solo sintetizar los hallazgos de esta tesis, sino también proporcionar una hoja de ruta para futuras investigaciones y aplicaciones del algoritmo de punto proximal en espacios de Hilbert.

Bibliografía

Bauschke, H. and Combettes, P. (2019). Convex analysis and monotone operator theory in hilbert spaces, corrected printing.

Benazic, R. (2000). *Topología en Espacios Euclidianos*. XVIII Coloquio de la Sociedad Matemática Peruana (SMP), IMCA-Lima, Perú, 2 edition.

Cano Lengua, M. A. (2011). Método del punto proximal para minimizar funciones cuasiconvexas usando el subdiferencial de clarke. *Revista Universidad del Callao*.

Clarke, F. H. (1990). *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM.

Cruzado Acuña, S. (2018). Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa. *Universidad Nacional Mayor de San Marcos*.

Fan, L., Liu, S., and Gao, S. (2003). Generalized monotonicity and convexity of nondifferentiable functions. *Journal of Mathematical Analysis and Applications*, 279(1):276–289.

Lages Lima, E. (1981). *Curso de Análise*, volume 2. Projeto Euclides, Impa, Rio de Janeiro.

Lázaro Carrión, M. (2002). *Sucesiones y Series*. Moshera.

Martinet, B. (1970). Regularisation, d'inéquations variationelles par approximations succesives. *Revue Francaise d'informatique et de Recherche operationelle*.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic theory*, volume 1. Oxford university press New York.

Mordukhovich, B. S., Nam, N. M., and Yen, N. (2006). Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming. *Optimization*, 55(5-6):685–708.

Papa Quiroz, E. A. (2019). Minimización cuasiconvexa: teoría, metodos y aplicaciones. *Facultad de Ciencias Matematicas, Universidad Nacional Mayor de San Marcos*.

Peypouquet, J. (2015). *Convex Analysis and Subdifferential Calculus*. Springer.

Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. *SIAM journal on control and optimization*, 14(5):877–898.

Rojo, A. (1973). *Álgebra II*, volume 1 of *Álgebra Lineal*. Libreria: El Ateneo, 4 edition.

The Luc, D. (1989). *Theory of vector optimization*, volume 319 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, 1 edition.