

UNIVERSIDAD NACIONAL DEL SANTA
ESCUELA DE POSGRADO
PROGRAMA DE DOCTORADO EN MATEMÁTICA



UNS
ESCUELA DE
POSGRADO

**Un método Proximal para optimización Multi objetivo
Cuasi-convexa y su aplicación a la economía**

**Tesis para obtener el Grado de
Doctor en Matemática**

Autor:

Mg. Bermeo Carrasco, Osmar Arnaldo

Asesor:

Dr. Moore Flores, Teodoro
Código ORCID: 0000-0002-1755-3459

Línea de Investigación:
Problemas de optimización

Nuevo Chimbote - PERÚ
2023



UNS
ESCUELA DE
POSGRADO

CONSTANCIA DE ASESORAMIENTO DE TESIS

Yo, Dr. Teodoro Moore Flores, mediante la presente certifico mi asesoramiento de la Tesis Doctoral titulada: "Un Método Proximal Optimización Multiobjetivo Cuasi-Convexa y su aplicación a la Economía", elaborado por el magister Osmar Arnaldo Bermeo Carrasco para obtener el Grado Académico de Doctor en Matemática en la Escuela de Posgrado de la Universidad Nacional del Santa.

Nuevo Chimbote, febrero del 2023

Dr. Teodoro Moore Flores

ASESOR

CODIGO ORCID: 0000-0002-1755-3459

DNI N° 32763522



UNS
ESCUELA DE
POSGRADO

CONFORMIDAD DEL JURADO EVALUADOR

“UN MÉTODO PROXIMAL PARA OPTIMIZACIÓN MULTIOBJETIVO CUASI-
CONVEXA Y SU APLICACIÓN A LA ECONOMÍA”

TESIS PARA OPTAR EL GRADO DE DOCTOR EN MATEMÁTICA

Revisado y Aprobado por el Jurado Evaluador:

Dr. Milton Milciades Cortez Gutiérrez

PRESIDENTE

CODIGO ORCID: 0000-0003-4939-7734

DNI N° 18162818

Dr. Herón Juan Morales Marchena

SECRETARIO

CODIGO ORCID: 0000-0002-5394-0958

DNI N° 32837715

Dr. Teodoro Moore Flores

VOCAL

CODIGO ORCID: 0000-0002-1755-3459

DNI N° 32763522



UNS
ESCUELA DE
POSGRADO

ACTA DE EVALUACIÓN DE SUSTENTACIÓN DE TESIS

A los diecisiete días del mes de febrero del año 2023, siendo las 11:40 horas, en el aula multimedia N° 01 de la Escuela de Posgrado de la Universidad Nacional del Santa, se reunieron los miembros del Jurado Evaluador, designados mediante Resolución Directoral N° 626-2022-EPG-UNS de fecha 22 de noviembre de 2022, conformado por los docentes: Dr. Milton Milciades Cortez Gutiérrez (Presidente), Dr. Herón Juan Morales Marchena (Secretario) y Dr. Teodoro Moore Flores (Vocal), con la finalidad de evaluar la tesis titulada: **UN MÉTODO PROXIMAL OPTIMIZACIÓN MULTI OBJETIVO CUASI-CONVEXA Y SU APLICACIÓN A LA ECONOMÍA**; presentado por el tesista **Osmar Arnaldo Bermeo Carrasco**, egresado del programa de **Doctorado en Matemática**.

Sustentación autorizada mediante Resolución Directoral N° 035-2023-EPG-UNS de fecha 13 de febrero de 2023.

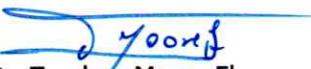
El presidente del jurado autorizó el inicio del acto académico; producido y concluido el acto de sustentación de tesis, los miembros del jurado procedieron a la evaluación respectiva, haciendo una serie de preguntas y recomendaciones al tesista, quien dio respuestas a las interrogantes y observaciones.

El jurado después de deliberar sobre aspectos relacionados con el trabajo, contenido y sustentación del mismo y con las sugerencias pertinentes, declara la sustentación como: APROBADO asignándole la calificación de: DIECINUEVE

Siendo las 12:30 horas del mismo día se da por finalizado el acto académico, firmando la presente acta en señal de conformidad.


Dr. Milton Milciades Cortez Gutiérrez
Presidente


Dr. Herón Juan Morales Marchena
Secretario


Dr. Teodoro Moore Flores
Vocal

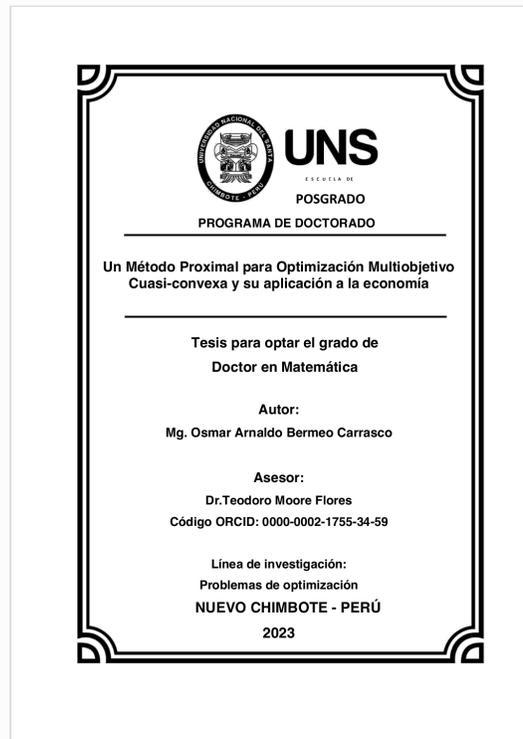


Recibo digital

Este recibo confirma que su trabajo ha sido recibido por **Turnitin**. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Osmar Bermejo
Título del ejercicio: Informe_tesis
Título de la entrega: Un Método Proximal para Optimización Multiobjetivo Cuasi-...
Nombre del archivo: TESIS_DOCTORAL_OSMAR_BERMEJO_UNS_2022_R.docx
Tamaño del archivo: 3.31M
Total páginas: 159
Total de palabras: 25,353
Total de caracteres: 159,383
Fecha de entrega: 14-sept.-2023 06:48p. m. (UTC-0500)
Identificador de la entrega... 2166400205



Un Método Proximal para Optimización Multiobjetivo Cuasi-convexa y su aplicación a la economía

INFORME DE ORIGINALIDAD

2%

INDICE DE SIMILITUD

2%

FUENTES DE INTERNET

1%

PUBLICACIONES

1%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	repositorio.unac.edu.pe Fuente de Internet	1%
2	www.ewh.ieee.org Fuente de Internet	1%
3	cybertesis.unmsm.edu.pe Fuente de Internet	<1%
4	repositorio.uns.edu.pe Fuente de Internet	<1%
5	Erik Papa Quiroz, Segundo Cruzado Acuña. "Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa en espacios Euclidianos", Pesquimat, 2019 Publicación	<1%
6	view.genial.ly Fuente de Internet	<1%
7	hdl.handle.net Fuente de Internet	<1%

Dedicatoria

Dedicado a mis padres
Esmilda y Samuel, a mi novia
Mariela y mis hermanos

Osmar Arnaldo Bermeo Carrasco

Agradecimiento

Agradeciendo a Dios por darme fortaleza, a mis padres por sus enseñanzas para enfrentar las dificultades de la vida, a mi novia Mariela por su apoyo incondicional donde es mi bazo derecho y socia de la vida, a mis hermanos por su apoyo invaluable, en el lidiar de la vida, a la UNS por la oportunidad que me brinda de seguir estudios posgrado y de manera especial a la Dr. Teodoro Moore Flores, por su asesoría y asistencia puntual, así también al Dr. Erik Papa Quiroz por sus notables sugerencias, a mi gran amigo Dr. Aldo Bazán y a Mg. Rolando Ichpas.

Osmar Arnaldo Bermeo Carrasco

Índice General

Conformidad del asesor	ii
Aprobación del Jurado Evaluador	iii
Acta de sustentación	iv
Recibo digital.....	v
Dedicatoria.....	vi
Índice general.....	vii
Resumen.....	ix
Abstracto.....	x
INTRODUCCION.....	1
1. PROBLEMA DE INVESTIGACIÓN	
1.1. Planteamiento y fundamentación del problema de investigación ...	4
1.2 Antecedentes de la investigación	9
1.3 Formulación del problema de investigación	12
1.3.1 Problema general	12
1.3.2 Problemas específicos	12
1.4 Delimitación de estudio	12
1.5 Justificación e importancia de la investigación	13
1.5.1 Justificación teórica	13
1.5.2 Justificación económica	13
1.5.3 Justificación tecnológica	13
1.6 Objetivos de la investigación	13
1.6.1 Objetivo general	13
1.6.2 Objetivos específicos	13
2. MARCO TEORICO	15
2.1 Fundamentos teóricos de la investigación	15
2.1.1 Orden de preferencia	15
2.1.2 Conos y relaciones de orden	18

2.1.3 Eficiencia, débil eficiencia y dominancia de solución	22
2.1.4 Conceptos básicos de análisis real	27
2.1.5 La norma en el espacio euclidiano R^n	28
2.1.6 Bolas en el espacio euclidiano R^n	28
2.1.7 Sucesiones y subsucesiones en el espacio euclidiano R^n	29
2.1.8 Series en el espacio euclidiano R^n	31
2.1.9 Conjuntos abiertos y cerrados en R^n	32
2.1.11 Puntos de acumulación en R^n	34
2.1.12 funciones continuas	35
2.2 Marco conceptual	36
2.2.1 Teoría básica de análisis convexo	36
2.2.2 Distancia proximal	47
2.2.3 Subdiferenciales	61
2.2.4 Subdiferencial de Frechet y en el límite	61
2.2.5 Subdiferencial de Clarke	78
2.2.7 Condiciones de optimalidad y condiciones de descenso	79
2.2.8 Convergencia de sucesiones	79

3.-	MÉTODO DE PUNTO PROXIMAL ESCALARIZADO INEXACTO CON DISTANCIA PROXIMAL GENERALIZADA	80
	3.1 El algoritmo	81
	3.2 Existencia de iterados	83
	3.3 Resultados de convergencia	92
4.	TASA DE CONVERGENCIA DEL ALGORITMO ISSPMR	93
	4.1 Análisis de la tasa de convergencia	98
5.	EXPERIMENTACION NUMERICA	107
6.	APLICACIÓN DEL ALGORITMO EN ECONOMIA	112
7.	MARCO METODOLOGICO	112
	7.1 Hipótesis central de la investigación	112
	7.1.1 Hipótesis general	112
	7.1.2 Hipótesis específicas	112
	7.2 Variables e indicadores de la investigación	113
	7.3 Métodos de investigación	113
	7.4 Diseño o esquema de investigación	113
	7.5 Población y muestra	113
	7.6 Actividades del proceso de investigación	114
	7.7 Técnicas e instrumentos de recolección de datos.....	115
	7.8 Técnicas de procesamiento y análisis de los resultados.....	114
8.	RESULTADOS Y DISCUSIÓN	116
	8.1 Resultados	116

8.2	Discusión	116
9.	CONCLUSIONES Y RECOMENDACIONES	118
9.1	Conclusiones	118
9.2	Recomendaciones	118
10	REFERENCIA BIBLIOGRAFICAS Y VIRTUAL.....	119
11	ANEXOS	125

INDICE DE LISTA DE FIGURAS

1.1	Conjunto factible y espacio factible	5
2.1	Ilustración de cono	17
2.2	Ilustración del cono convexo	17
2.3	Soluciones eficientes de Pareto	21
2.4	Soluciones débiles de Pareto	23
2.5	Puntos no dominados y puntos débilmente no dominados	26
2.6	Conjunto abierto	32
2.7	Conjunto de nivel inferior	36
2.8	Conjunto de nivel superior	37
2.9	Epigrafo de la función f	37
2.10	Función semicontinua inferior	38
2.11	Conjunto convexo	40
2.12	Función convexa	41
2.13	Función cuasi-convexa	43
2.14	Función conveva no diferenciable en todo \mathbb{R}	61
2.15	Función no convexa, no diferenciable en todo \mathbb{R}	63
2.16	Función convexa de dos variables no diferenciable en todo \mathbb{R}^2	66
2.17	Conjunto de subdiferenciales de f para los valores $a > 0, b = 0$	67
2.18	Conjunto de subdiferenciales de f para los valores $a = 0, b > 0$	68
2.19	Conjunto de subdiferenciales de f cuando $a = b$	69
2.20	Conjunto de subdiferenciales de f para los valores $b = ma, m > 0$...	69
2.21	Conjunto de subdiferenciales de f para los valores $b = ma, m < 0$..	70

5.1	Funciones componentes objetivo ejemplo 5.1	97
5.2	Curvas de nivel de las funciones objetivo ejemplo 5.1	98
5.3	Soluciones de Pareto ejemplo 5.1	98
5.4	Frontera de Pareto ejemplo 5.1, con distancia proximal Kullback-Leiber Bregman	100
5.5	Frontera de Pareto ejemplo 5.1, con distancia proximal homogénea de segundo orden	101
5.6	Conjunto de restricciones ejemplo 5.2	102
5.7	Funciones componentes objetivo ejemplo 5.2	102
5.8	Distancias de los puntos consecutivos ejemplo 5.2 con $\lambda_k = 0.2$	104
5.9	Distancias de los puntos consecutivos ejemplo 5.2 con $\lambda_k = 0.4$	105
6.1	Funciones componentes objetivo utilidad y costo	107
6.2	Soluciones de Pareto en el conjunto de restricciones C	107
6.3	Frontera de Pareto, funciones utilidad y costo	109
6.4	Distancias de puntos consecutivos	110

NOTACIONES Y SIMBOLOGIAS

NOTACION	DESCRIPCION
X	conjunto definido, como conjunto factible en el espacio de decisión $X \subseteq \mathbb{R}^n$
$Y = F(X)$	conjunto definido, como conjunto factible en el espacio objetivo $Y \subseteq \mathbb{R}^m$
\preceq	relación binaria en un conjunto $C \subseteq \mathbb{R}^n$
(C, \preceq)	conjunto preordenado
$\bar{x} \leq \bar{y}$	$x_k \leq y_k, k = 1, 2 \dots n$
$\bar{x} < \bar{y}$	$x_k < y_k, k = 1, 2 \dots n$
\mathbb{R}^n	espacio Euclidiano
$\mathbb{R}_{\geq}^n = \mathbb{R}_+^n$	octante no negativo en \mathbb{R}^n
$\mathbb{R}_{>}^n = \mathbb{R}_{++}^n$	octante positivo en \mathbb{R}^n
$\bar{\mathbb{R}}$	$\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ conjunto de los números reales extendidos
X_E	conjunto de soluciones eficientes (o soluciones de Pareto)
$Y_N = F(X_E)$	conjunto no dominado (o frontera de Pareto)
$x^k \rightarrow \hat{x}$	denota que $\lim_{k \rightarrow \infty} x^k = \hat{x}$
$\operatorname{argmin}\{f(x), x \in C \subseteq \mathbb{R}^n\}$	conjunto de minimizadores de $f, f: C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$
$d(\cdot, \bar{y})$	distancia proximal $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$
$H(\bar{x}, \bar{y})$	distancia proximal inducida a $d, H: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$
$D(X)$	Conjunto de distancias proximales
$\mathcal{F}(X)$	Conjunto de distancias proximales inducidas
$f'(\bar{x}, d)$	derivada direccional de f en \bar{x} en dirección de $d \in \mathbb{R}^n$
$\partial f(x)$	Subdiferencial de f en x
$\hat{\partial} f(x)$	Conjunto de subgradientes denominado subdiferencial de Frechet

$B(\bar{x}_0, \varepsilon)$	$\{\bar{x} \in \mathbb{R}^n: \ \bar{x} - \bar{x}_0\ < \varepsilon\}$, bola abierta o entorno de centro \bar{x}_0 y radio ε
$\delta_C(\cdot)$	define la función indicatriz en un conjunto no vacío $C \subseteq \mathbb{R}^n$
$N_C(\bar{x}_0)$	$\{v \in \mathbb{R}^n: \langle v, y - \bar{x}_0 \rangle \leq 0, \forall y \in C\}$, cono normal al punto \bar{x}_0 relacionado al conjunto $C \subseteq \mathbb{R}^n$
$epi(f)$	$\{(\bar{x}, \rho) \in \mathbb{R}^n \times \mathbb{R}: f(\bar{x}) \leq \rho\}$, conjunto epigrafo de una función f
$L_f(\rho)$	$\{\bar{x} \in \mathbb{R}^n: f(\bar{x}) \leq \rho\}$, conjunto de nivel de la función f
$int(C)$	interior del conjunto C
e^k	error de aproximación de convergencia
$\ e^k\ $	Norma del error de la aproximación
Ω_k	$\{x \in C: F(x) \leq F(x^k), \forall k \in \mathbb{N}\}$
$\sum_{k=0}^{+\infty} \eta_k < +\infty$	Serie de término k-ésimo converge a un número real
$\partial^{Lim} f(x)$	Conjunto de subgradientes denominado subdiferencial en el límite
$f^0(\bar{x}, v)$	derivada direccional generalizado de f en \bar{x} en dirección de $v \in \mathbb{R}^n$
$\partial^0 f(\bar{x})$	Conjunto de subgradientes denominado subdiferencial de Clarke

INDICE DE TABLAS

5.1	Resultados de la experimentación numérica con la distancia proximal Kullback-Leiber Bregman ejemplo 5.1	100
5.2	Resultados de la experimentación numérica con la distancia proximal homogénea de segundo orden ejemplo 5.1	101
5.3	Resultados de la experimentación numérica con la distancia proximal Kullback-Leiber Bregman ejemplo 5.2 con $\lambda_k = 0.2$	103
5.4	Resultados de la experimentación numérica con la distancia proximal Kullback-Leiber Bregman ejemplo 5.2 con $\lambda_k = 0.4$	105
6.1	Resultados de la experimentación numérica con la distancia proximal Kullback-Leiber Bregman de ejemplo aplicado a la economía	108

RESUMEN

En el presente trabajo de tesis presentamos un método proximal escalarizado inexacto para optimización multiobjetivo con distancia proximal generalizada, la finalidad del método es resolver problemas multiobjetivo cuasi-convejos con restricciones en el espacio Euclidiano, con las funciones objetivo que sean localmente Lipschitz. Considerando las hipótesis necesarias se prueba que la sucesión generada por el método propuesto está bien definida. Se presenta resultados de convergencia de la sucesión generada por el algoritmo, bajo ciertas condiciones de la función objetivo, donde se prueba que dicha sucesión converge hacia un crítico Pareto-Clarke, seguidamente se analiza la tasa de convergencia del algoritmo, considerando ciertas condiciones, resultando que la tasa de convergencia es lineal y superlineal. Finalmente se realiza la experimentación numérica del método propuesto a problemas biobjetivos y además se valida el método aplicándolo a un problema en economía, donde se presentan algunos experimentos computacionales.

Palabras claves. Método proximal, función cuasi-convexa, optimización multiobjetivo, soluciones de Pareto, punto crítico Pareto-Clarke, distancia proximal, tasa de convergencia.

ABSTRACT

In this thesis we present an inexact scalarized proximal method for multiobjective optimization with generalized proximal distance, the purpose of the method is to solve quasi-convex multiobjective problems with restrictions in Euclidean space, with the objective functions that are locally Lipschitz. Considering the necessary hypotheses, it is proved that the sequence generated by the proposed method is well defined. Convergence results of the sequence generated by the algorithm are presented, under certain conditions of the objective function, where it is proved that said sequence converges towards a critical Pareto-Clarke, then the rate of convergence of the algorithm is analyzed, considering certain conditions, resulting that the rate of convergence is linear and superlinear. Finally, the numerical experimentation of the proposed method to biobjective problems is carried out and the method is also validated by applying it to a problem in economics, where some computational experiments are presented.

Keywords. Proximal method, quasi-convex function, multiobjective optimization, Pareto solutions, Pareto-Clarke critical point, proximal distance, convergence rate.

INTRODUCCION

Gran parte de los problemas del mundo real implican la optimización simultánea de varios objetivos que generalmente presentan conflictos entre ellos; es decir, la mejora en uno conduce a empeorar el otro. La presencia de tales tipos de problemas es tan significativa, que consume gran parte de nuestro tiempo cotidiano en tomar una decisión. Se trata, por ejemplo, de escoger de manera óptima para llegar al trabajo, establecer el orden de nuestras tareas, elegir el restaurante para el almuerzo, hacer las compras en el supermercado, preparar la cena y la distribución de actividades en el tiempo de ocio restante.

También es el mismo tipo de problemas que enfrentan los ingenieros, economistas y técnicos a la hora de diseñar e implementar sistemas de todo tipo: existen múltiples objetivos a cumplir y se espera lograrlos todos en la medida de lo posible. Aunque la mayoría de los problemas de decisión involucran este tipo de situaciones, las propuestas computacionales de automatización que se han presentado para resolverlos habitualmente se limitan a convertir el problema de objetivos múltiples en uno en que existe un solo objetivo.

El primer tratado de optimización multiobjetivo fue dado inicialmente por el economista político F. Y. Edgeworth (1881), en su obra *Mathematical Psychics*, luego el economista Wilfredo Pareto (1896), en su obra manual de economía política, donde extendió la teoría y proporcionó condiciones necesarias para lo que hoy se denomina óptimo de Pareto. Sin embargo, el primer tratamiento matemático formal de un problema de optimización vectorial en espacios de dimensión finita se encuentra incluido en el célebre trabajo de programación no lineal de Kuhn y Tucker (1951), y el primero de los casos vectoriales fue dado por otro economista L. Hurwicz en (1958).

Dado la referencia anterior donde Kuhn y Tucker ya proporcionaron resultados sobre la programación multiobjetivo, también se destacó la aparición de los otros trabajos importantes como los de Carlos Romero (1991 y 1993), Rios-Insua (1990), Vincke (1992), Jahn (1991), Miettnien K.M (1999), Jahn J (2004) y

Ehrgott. M (2005). Desde los puntos de vista de los tipos de problemas a resolver hay que destacar que una gran parte de la biografía citada hace énfasis en los problemas multiobjetivos lineales.

En la actualidad las últimas investigaciones realizadas, se destaca que se han limitado a resolver problemas de optimización multiobjetivo no lineales considerando que presentan mayor complejidad, la cual ha motivado a los investigadores a desarrollar métodos que permitan resolver este tipo de problemas. Uno de estos métodos es el método de punto proximal para minimización multiobjetivo cuasi-convexa irrestricto en el espacio Euclidiano introducido por Fernández [28] en el 2014 el cual es un método iterativo que se caracteriza por resolver en cada iteración un problema más simple de optimización, de tal forma que paso a paso las iteraciones se aproximan a la solución buscada.

Además cabe mencionar el método proximal en optimización vectorial introducido por Chuan L y Chih J [26] donde establecen un algoritmo para resolver un problema de optimización vectorial irrestricto, donde consideraron que el espacio de decisión es el espacio de Hilber real X y el espacio objetivo es el espacio de Banach real Y y la función a optimizar sea convexa de X en $Y \cup \{+\infty\}$, y finalmente el método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa introducido por Cruzado Acuña [22] en el 2019 resolvió problemas irrestrictos definidos en el espacio Euclidiano, donde las funciones vectoriales dadas son localmente Lipschitz, cuasi-convexas, demuestra que la sucesión generada por el método está bien definida y converge a un punto crítico Pareto-Clarke.

Siguiendo la secuencia de la evolución de la optimización multiobjetivo y las aplicaciones al contexto real es una motivación para la presente investigación tratar con problemas multiobjetivo restrictos.

El método proximal para optimización multiobjetivo cuasi-convexa con restricciones desarrollado permitió mostrar que una sucesión generada por algoritmo converge a un punto crítico Pareto-Clarke del problema de optimización multiobjetivo.

Finalmente mencionar que este trabajo de investigación ha dado fruto al artículo científico titulado “Un algoritmo de punto proximal escalarizado inexacto para minimización multiobjetivo con restricciones y funciones objetivo cuasi-convexas,

que ha sido sometido a publicación a la revista internacional European Journal of Operational Research.

CAPÍTULO I

PROBLEMA DE INVESTIGACION

1.1. Planteamiento y fundamentación del problema de investigación

En el mundo real se busca optimizar los recursos de la naturaleza, en este sentido optimizar significa encontrar la mejor solución a diversos problemas en diferentes áreas como por ejemplo en economía, en ingeniería, en medicina. Este tipo de problemas es tan significativo, que toma gran parte de nuestro tiempo en la toma una decisión ya sea en el ámbito personal, como profesional.

Por ejemplo, comprar una casa, donde podríamos centrarnos en el precio y eso significaría elegir la casa de menor costo, indicándonos un solo objetivo, sin embargo al elegir una casa se considera otros objetivos que han llevado a tomar esa decisión como, por ejemplo, área del terreno, el impuesto, el material de construcción entre otros.

Si fuera un problema de objetivo único, bastaría con llamar a todas las inmobiliarias y hacer una lista específica de precios (sin tomar en consideración área del terreno, el impuesto, material de construcción, etc.) y escogería la casa con menor precio sin mayores complicaciones. El uso del modelo matemático de búsqueda no sería el más apropiado dado que si no se está considerando los otros objetivos se podría estar tomando una mala decisión.

En la actualidad existen modelos matemáticos con condiciones específicas, que se ajustan mejor a la naturaleza de estos problemas, como es la optimización de múltiples objetivos o optimización multiobjetivo. El modelo matemático se denota a continuación

$$\text{Optimizar } F(x) = \text{Opt}\{(F_1(x), \dots, F_m(x)) : x \in X \subset \mathbb{R}^n\} \quad (1.1)$$

donde $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $F(x) = (F_1(x), \dots, F_m(x)), m \geq 2$, denota el vector de funciones objetivo definido en el espacio variable o llamado **espacio de decisión** \mathbb{R}^n , con valores en el espacio objetivo o llamado espacio criterio \mathbb{R}^m , $X \subseteq \mathbb{R}^n$ representa el espacio factible de decisión o conjunto factible. $\bar{x} = (x_1, x_2, \dots, x_n)$, representa el vector de las n variables de decisión, $Y = F(X) \subset \mathbb{R}^m$, representa el conjunto factible definido en el **espacio objetivo**.

Por ejemplo, Consideremos la siguiente Figura 1.1

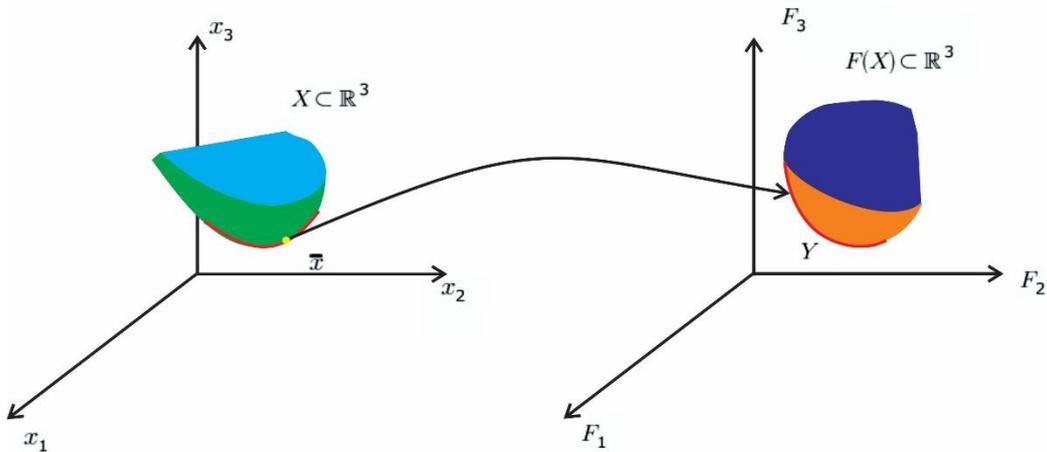


Figura 1.1 Conjunto factible y espacio factible

Se representa gráficamente, a la izquierda, el conjunto factible para las variables de decisión x_1, x_2 y x_3 y a la derecha, cuando son insertadas estas variables en las funciones F_1, F_2 y F_3 , generan un espacio factible tridimensional la cual son soluciones para las funciones objetivo.

Los modelos matemáticos (1.1) independientemente de la naturaleza de función objetivo F y del espacio factible de decisión X , es necesario conocer las características de la función a optimizar y el espacio factible. Por ejemplo

$$\begin{cases} \min F(x) \\ x \in X \end{cases} = \min\{(F_1(x), F_2(x)) : x \in X \subset \mathbb{R}^2\}$$

donde X es el primer cuadrante y la función $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, puede ser una función arbitraria

Las funciones convexas y cuasi-convexas juegan un papel importante en la optimización. En muchas aplicaciones a la teoría económica una generalización de los conceptos de función convexa se hace necesaria. Los conceptos de función cuasi-convexa dan más generalidad a las conclusiones de la teoría económica. Según [50] la cuasi-convexidad es una propiedad más débil que la convexidad de ahí que algunas propiedades de estas últimas no se cumplen para las primeras.

El crecimiento tecnológico constante que se ha producido ha implicado desarrollar nuevos problemas de optimización, lo que ha implicado el estudio de nuevas teorías matemáticas y nuevas soluciones numéricas que conduzca a la resolución de los problemas. En este trabajo de investigación abordaremos el problema de optimización multiobjetivo restringido donde las funciones objetivo son cuasi-convexas.

Este tipo de problemas tiene una alta complejidad y ha conllevado a muchos investigadores a desarrollar métodos numéricos que permitan solucionar este tipo de problemas. Uno de estos métodos es el método proximal en optimización vectorial Bonnel H., Iusem A.N. Svaiter B. F [16] en el 2019, trabajaron el método punto proximal para optimización convexa con valores escalares.

En el año 2016 Apolinario, Papa Quiroz y Oliveira [1] trabajaron con funciones vectoriales, considerando que estas son cuasi-convexas, localmente Lipschitz, y que están acotadas inferiormente, la cual presentaron como objeto resolver el problema de minimización (1.1) en su versión multiobjetivo.

La iteración que consideraron fue: Dado un punto inicial arbitrario $x^k \in \mathbb{R}^n$, se debe encontrar un $x^{k+1} \in \Omega_k$, $\Omega_k = \{x \in \mathbb{R}^n: F(x) \leq F(x^k)\}$ tal que

$$0 \in \partial^0 \left(\langle F(\cdot), Z_k \rangle + \left(\frac{\lambda_k}{2} \right) \| \cdot - x^k \|^2 \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1})$$

donde ∂^0 es el subdiferencial de Clarke, $\lambda_k > 0$, $\{Z_k\} \subset \mathbb{R}_+^n$, $\|Z_k\| = 1$ y $N_{\Omega_k}(x^{k+1})$, representa el cono normal de Ω_k en x^{k+1} . Los autores demostraron que la sucesión $\{x^k\}$ generada por el método está bien definido y converge a un punto crítico Pareto-Clarke del problema y cuando F es convexa, esta sucesión converge a una solución de Pareto débil.

Bonnel H., Iusem A.N. Svaiter B. F (2019) [16] presentaron el método proximal donde consideran funciones convexas, y no funciones cuasi-convexas, sin embargo, Papa Quiroz y Cruzado (2019) [56] presentaron un artículo, donde introdujeron un algoritmo de punto proximal inexacto para resolver problemas de minimización cuasi-convexas multiobjetivos sin restricciones donde ampliaron el algoritmo presentado por Apolinario, Papa Quiroz y Oliveira (2016) [1].

Papa Quiroz E., Borda D., Collantes F. (2020). [55], en el artículo un método proximal para minimización multiobjetivo cuasiconvexo en el ortante no negativo y su aplicación a la teoría de la demanda en microeconomía presentaron un algoritmo proximal utilizando distancias proximales, para esolver problemas de minimización multiobjetivo cuasi-convexos sujetos a restricciones no negativas y obtienen resultados parciales de convergencia.

Dadas las referencias mencionadas, la motivación de la investigación es estudiar la extensión del método proximal para optimización multiobjetivo con restricciones con funciones objetivos estrictamente o semi-estrictas cuasi-convexas y su aplicación a la economía.

1.2. Estado del arte del tema de investigación

La teoría de optimización multiobjetivo ha tenido un gran avance en los últimos años donde muchos investigadores han logrado establecer modelos matemáticos para diferentes áreas, por ejemplo, en campo de la economía han sido utilizados para optimizar las utilidades, los costos y los ingresos Luc,D.T(1989).

Maino M., Pittet J., Kodrich C., (1993) señalan que las técnicas de programación multiobjetivo, llamadas técnicas de optimización vectorial se enfrentan al problema de optimizar simultáneamente varios objetivos, además indican que como es imposible definir un óptimo cuando existen varios objetivos en conflicto, la programación multiobjetivo en vez de buscar una solución óptima, trata de encontrar un conjunto de soluciones eficientes no dominadas u óptimos de Pareto.

Carlos Romero (1996) indica que la programación multiobjetivo también llamada optimización vectorial constituye un enfoque multicriterio de gran potencialidad cuando el contexto decisional está definido por una serie de objetivos a optimizar que deben de satisfacer un determinado conjunto de restricciones. El propósito del enfoque multiobjetivo consiste en segregar del conjunto de soluciones posibles un subconjunto propio del mismo cuyos elementos gocen de la condición de optimalidad Paretiana.

Matthias Ehrgott. (2005) señala que la vida se trata de decisiones donde dichas decisiones, sin importar si las toma un grupo o un individuo, por lo general implican varios objetivos en conflicto, y además hace referencia que, desde sus primeras raíces, que fueron puestas por el economista Wilfredo Pareto a fines del siglo XIX, la disciplina ha prosperado y crecido, especialmente durante las últimas décadas. Hoy en día, muchos sistemas de soporte de decisiones incorporan métodos para tratar con objetivos contradictorios. La base para tales sistemas es una matemática teoría de la optimización bajo múltiples objetivos.

A continuación, describiremos las últimas investigaciones relacionadas al método proximal para la optimización multiobjetivo cuasi-convexa. Para resolver problemas con múltiples objetivos, usando el método proximal es necesario establecer propiedades tanto en el conjunto factible, como en el espacio factible, donde cabe resaltar que el conjunto factible representa el espacio de decisión y el espacio factible representa el espacio objetivo.

En el artículo un método proximal interior en optimización vectorial Villacorta, K.D.V., Oliveira, P.R (2011) estudiaron el problema de la optimización vectorial donde encontraron puntos eficientes para mapeos de \mathbb{R}^n a \mathbb{R}^m , con respecto al orden parcial inducida por un cono cerrado, convexo y puntiagudo $C \subset \mathbb{R}^m$, con interior no vacío, desarrollaron para este problema una extensión del método del punto proximal para un problema de optimización convexa con valores escalares con una convergencia modificada, utilizaron una función de Bregman con una condición de detección de convergencia para construir una aproximación generalizada, demostraron que la sucesión generada por el algoritmo converge a una solución débilmente eficiente en un problema de optimización vectorial en un conjunto no poliédrico.

En el método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa Acuña. C (2018) resolvió problemas irrestrictos de minimización multiobjetivo cuasiconvexa definidos en el espacio Euclidiano, donde las funciones vectoriales son localmente Lipschitz, donde demuestra que la sucesión generada por el método está bien definida y converge globalmente a un punto de acumulación, y que la sucesión generada por el algoritmo converge hacia

un punto crítico Pareto-Clarke del problema y además valida el método mediante algunos experimentos computacionales.

Así mismo otra investigación, H. Bonnel, A.N. Iusem, B.F. Svaiterc (2019) desarrollaron el método proximal en optimización vectorial, donde hicieron una extensión del método punto proximal para optimización convexa con valores escalares en esta extensión los subproblemas consisten en encontrar soluciones para regularizaciones adecuadas del problema original, presentaron una versión exacta como inexacta, en la que los subproblemas resuelven soluciones aproximadas es decir en ambos casos buscan la convergencia débil de la sucesión.

Finalmente, en el artículo un método Proximal para minimización multiobjetivo cuasi-convexa en el ortante no negativo y su aplicación a la teoría de la demanda en microeconomía, Papa E., Borda D., Collantes F (2020) los autores propusieron un método de punto proximal inexacto para resolver problemas de minimización multiobjetivo con funciones objetivas cuasiconvexas localmente lipschitziana en el ortante no negativo en el espacio Euclidiano. Consideraron que la función objetivo sea acotada, semicontinua inferior y utilizaron distancias proximales, donde demostraron la convergencia de la sucesión generada por el algoritmo, dieron algunos resultados computacionales y una aplicación a la teoría de la demanda.

1.3. Antecedentes de la investigación

Los antecedentes relacionados a la investigación son los siguientes

Ceng, L., Yao, J (2007). Métodos proximales aproximados en optimización de vectores [19]. Este artículo los autores estudian el problema de optimización vectorial para encontrar puntos débilmente eficientes para mapeos en un espacio de Banach Y , con respecto al orden parcial inducido por un cono cerrado, convexo y puntiagudo $C \subset Y$ con un interior no vacío.

El método proximal en la optimización vectorial extiende para desarrollar un método proximal aproximado, para este problema en virtud del método de punto proximal aproximado debe encontrar una raíz de un operador monótono máximo. En esta aproximación proximal, los subproblemas consisten en encontrar puntos

débilmente eficientes para regularizaciones adecuadas del mapeo original. Presentan una versión tanto absoluta como relativa, en la que los subproblemas se resuelven solo aproximadamente. La convergencia débil establece la generación de la sucesión convergente a un punto débil eficiente.

L. M. Graña Drummond Y Ellen H. Fukuda (2009). Convergencia del método de gradiente proyectado para optimización vectorial [37]. Tomando como referencia Graña Drummond y Iusem (2004) donde propusieron una extensión del método de gradiente proyectado para problemas de optimización de vectores restringidos. Los autores simplemente mostraron estacionariedad de todos los grupos puntos y, para otra versión del algoritmo, bajo algunas afirmaciones adicionales, proporcionan convergencia de soluciones débilmente eficientes. En este trabajo, primero corrigieron un ligero error en la prueba del resultado de continuidad del trabajo desarrollado por Graña Drummond y Iusem (2004), y luego ampliaron su análisis de convergencia.

Villacorta, K.D.V., Oliveira, P.R (2011). Un método proximal interior en optimización vectorial [72]. Este artículo los autores estudian el problema de la optimización vectorial donde encontraron un rendimiento débil de puntos eficientes para mapeos de R^n a R^m , con respecto al orden parcial inducida por un cono cerrado, convexo y puntiagudo $C \subset R^m$, con interior no vacío. Desarrollaron para este problema una extensión del método del punto proximal para un problema de optimización convexa con valores escalares con una convergencia modificada condición de detección que nos permite construir un método proximal interior para resolución de problemas de optimización vectorial en un conjunto no poliédrico.

Papa Quiroz, E.A., Mallma Ramirez, L., Oliveira, P.R (2015). Un método proximal inexacto para la minimización cuasiconvexa [54]. En este artículo proponen los autores un método de punto proximal inexacto para resolver problemas de minimización restringida con funciones objetivo cuasiconvexas localmente Lipschitz. Suponiendo que la función también está acotada desde abajo, semicontinuos inferiores y utilizando distancias proximales, muestran que la sucesión generada por el método converge a un punto estacionario del problema.

H. C Fernández Apolinario, E. A. Papa Quiroz y P.R. Oliveira (2016). "Un método de escalarización de punto proximal para minimización multiobjetivo cuasi-

convexa” [1]. En este artículo los autores propusieron un método de escalarización de punto proximal para resolver problemas de minimización multiobjetivo sin restricciones con Lipschitz local y con funciones vectoriales cuasi-convexo. Demostraron, bajo supuestos naturales, que la sucesión generada por el método está bien definida y converge globalmente a un punto crítico de Pareto-Clarke. El método propuesto por los autores considera como una extensión para el caso no convexo, del método proximal inexacto para problemas de minimización multiobjetivo convexa estudiados por Bonnel, H., Iusem, A.N., Svaiter, B.F. (2005). [26]

Segundo Cruzado Acuña (2018). “Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa” [22]. En esta investigación el autor presenta un método de punto proximal escalarizado inexacto para resolver problemas irrestrictos de minimización multiobjetivo cuasiconvexa definidos en el espacio Euclidiano, donde las funciones vectoriales son localmente Lipschitz. Prueba que la sucesión generada por el método está bien definida y converge globalmente a un punto de acumulación. Además, las sucesiones convergen hacia un punto crítico Pareto-Clarke del problema. Finalmente valida el método propuesto y los resultados encontrados, mediante algunos experimentos computacionales.

H. Bonnel, A.N. Iusem, B.F. Svaiter (2019). “El método proximal en optimización vectorial” [16]. En la cual los autores hacen una extensión del método punto proximal para optimización convexa con valores escalares, en esta extensión los subproblemas consisten en encontrar soluciones para regularizaciones adecuadas del problema original, presenta tanto una versión exacta como inexacta, en la que los subproblemas solo resuelven soluciones aproximadas es decir en ambos casos buscan la convergencia débil de la sucesión.

Kin Keung Lai¹, Shashi Kant Mishra, Geetanjali Panda, Md Abu Talhamainuddin Ansary, y Bhagwat Ram (2020). En el método de descenso más pronunciado para problemas optimización multiobjetivo sin restricciones [39]. En este artículo los autores hacen mención que el gradiente- q es la generalización del gradiente basado en la derivada- q . Donde la versión q construye el método de descenso más empinado para problemas de optimización multiobjetivo sin

restricciones y recuperado al clásico cuando q es igual a 1. En este método, el proceso de búsqueda se mueve paso a paso de forma global al principio luego particularmente en una vecindad. Este método no depende de un punto de partida. Finalmente verifican el algoritmo propuesto para encontrar puntos críticos con ejemplos numéricos.

Papa E., Borda D., Collantes F (2020). Un método Proximal para minimización multiobjetivo cuasi-convexo en el ortante no negativo y su aplicación a la teoría de la demanda en microeconomía [55]. En este artículo proponen los autores un método de punto proximal inexacto para resolver problemas de minimización multiobjetivo con funciones objetivas cuasiconvexas localmente lipschitziana en el ortante no negativo. Suponen que la función también está acotada desde abajo, semicontinuo inferior y utilizan distancias proximales, obtienen la convergencia de las sucesiones generadas por el algoritmo. Dieron algunos resultados computacionales y una aplicación a la teoría de la demanda.

1.4. Formulación del problema de investigación

1.4.1. Problema general

¿Será posible demostrar la extensión del Método Proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones y su aplicación a la economía?

1.4.2. Problemas específicos

1. ¿Es posible introducir un algoritmo de punto proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones en espacios euclidianos?
2. ¿Es posible la convergencia de la sucesión generada por el algoritmo a un punto crítico o a un mínimo?
- 3.- ¿Se puede analizar la velocidad de convergencia del algoritmo?
4. ¿Se puede aplicar el algoritmo a la solución de un problema en economía?

1.5. Delimitación del estudio

La siguiente investigación estará limitada a estudiar el método proximal para optimización multiobjetivo sobre funciones cuasi-convexas definidas sobre el espacio Euclidiano.

1.6. Justificación e importancia de la investigación

1.6.1. Justificación teórica.

Para la comunidad matemática se presentaría una demostración de un algoritmo extendiendo en el campo del conocimiento de matemática aplicada.

1.6.2. Justificación económica.

Se podría aplicar a problemas de equilibrios económicos en economía como por ejemplo a las competencias perfectas donde en [10] menciona que el modelo de competencia perfecta es uno de los modelos de mercado más importantes en microeconomía

1.6.3. Justificación tecnológica.

Considerando que un método es un conjunto de instrucciones que permite solucionar problemas, implica convertirse en un programa o algoritmo, por lo tanto, la finalidad es transformar el algoritmo en un software.

1.7. Objetivos de la investigación: General y específicos

1.7.1. Objetivo general

Extender el Método Proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones y su aplicación a la economía.

1.7.2. Objetivos específicos

- 1.- Introducir un algoritmo de punto proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones en espacios euclidianos
- 2.- Demostrar la convergencia de la sucesión generada por el algoritmo a un punto crítico o a un mínimo.

3.- Analizar y demostrar la velocidad de convergencia del algoritmo

4.- Aplicar el algoritmo a la solución de un problema en economía.

CAPÍTULO II

MARCO TEÓRICO

2.1. Fundamentos teóricos de la investigación

El desarrollo de esta investigación se basará en la teórica relacionada a los métodos proximales para optimización multiobjetivo cuasi-convexa y los procedimientos que permitirán la extensión del método. A continuación, se enunciará los conceptos básicos.

2.1.1 Orden de Preferencia

En un problema de optimización tradicional de la forma

$$\min_{x \in X} f(x) = \min\{f: \mathbb{R}^n \rightarrow \mathbb{R} : x \in X \subset \mathbb{R}^n\}$$

Si hacemos una evaluación de la función objetivo nos permite decidir entre dos puntos factibles cual es la mejor. Es decir, un punto x^{k+1} mejorara a otro x^k si se verifica la relación $f(x^{k+1}) \leq f(x^k)$, $\forall x^k \in X$, para determinar el valor mínimo de f , y si se busca determinar el valor máximo de f diremos que un punto x^{k+1} mejorara a otro x^k si se verifica la relación $f(x^k) \leq f(x^{k+1})$. Esto permite a muchos algoritmos encontrar direcciones de mejora de la función objetivo f que nos vayan conduciendo hasta encontrar su valor óptimo.

Sin embargo, en un problema con múltiples objetivos cada uno de los valores de las funciones objetivo en un punto x^k , $(F_1(x^k), \dots, F_m(x^k))$ no es claro como decidir si otra combinación dada en otro punto x^{k+1} de X , $(F_1(x^{k+1}), \dots, F_m(x^{k+1}))$ es o no mejor. Dado que en los problemas de optimización multiobjetivo (POM) suelen estar en conflicto, implicando que raramente se pueda dar la existencia de una única solución que optimice a todas funciones objetivo de manera simultánea, por lo tanto, existirá un conjunto de soluciones consideradas igualmente válidas para el problema.

Determinar la elección de al menos una de estas soluciones implicaría una toma de decisiones mediante un proceso de negociación de preferencias, entonces el método a utilizar dependerá esencialmente del criterio utilizado en el espacio objetivo para ordenar soluciones. Por tanto, para el problema de optimización multiobjetivo es necesario establecer un orden de preferencia en

un conjunto $Y \subset \mathbb{R}^n$, siendo fundamental, dado que nos permite ordenar los elementos del conjunto en función de un orden previamente establecido.

A continuación, introducimos la definición de relaciones binarias, sus propiedades, considerados como base de las más importantes clases de relaciones que se dan en optimización multiobjetivo.

Definición 2.1.1.1 -Una relación binaria \mathcal{R} en un conjunto Y es llamada

1. Reflexiva, si $(z, z) \in \mathcal{R}$ para todo $z \in Y$
2. Simétrica, si $(x, z) \in \mathcal{R} \Rightarrow (z, x) \in \mathcal{R}$, para cualquier $x, z \in Y$
3. Asimétrica, si $(x, z) \in \mathcal{R} \Rightarrow (z, x) \notin \mathcal{R}$, para cualquier $x, z \in Y$
4. Antisimétrica, si $(x, z) \in \mathcal{R}$ y $(z, x) \in \mathcal{R} \Rightarrow x = z$, para cualquier $x, z \in Y$
5. Transitiva, si $(x, z) \in \mathcal{R}$ y $(z, w) \in \mathcal{R} \Rightarrow (x, w) \in \mathcal{R}$, para cualquier $x, z, w \in Y$
6. Conectada, conexa o completa, si $(x, z) \in \mathcal{R}$ o $(z, x) \in \mathcal{R}$, para cualquier $x, z \in Y$

Por ejemplo, La relación “mayor o igual que” (\geq) en el conjunto de los enteros \mathbb{Z} es reflexiva antisimétrica y transitiva.

i) Reflexiva

Como $a \geq a$ para cada entero a ($a \in \mathbb{Z}$), luego \geq es reflexiva

ii) Antisimétrica

Si $a \geq b$ y $b \geq a$, entonces $a = b$ por lo tanto \geq es antisimétrica

iii) Transitiva

Si $a \geq b$ y $b \geq c$, entonces $a \geq c$, por lo tanto \geq es transitiva

Podemos notar fácilmente que no cumple con la simetría

Definición 2.1.1.2 -Una relación binaria R sobre un conjunto X se dice que es un preorden si es reflexiva y transitiva.

La relación binaria \preceq , que de aquí en adelante asumiremos para referirnos indirectamente a la relación R que “preferido a”. En el caso que \preceq sea un preorden entonces el par (X, \preceq) es llamado un conjunto preordenado

Definición 2.1.1.3- Una relación binaria \preceq en un conjunto Y se dice que es

1. un orden parcial, si es un preorden antisimétrico
2. un orden total, si se trata de un orden parcial conectado
3. un orden parcial estricto, si es asimétrica y transitiva.

Por ejemplo, sea el conjunto $A = \{2,3,4,5\}$ y $R: A \rightarrow A$. R es una relación “a divide exactamente a b” entonces R es de orden parcial.

En efecto los elementos de R estarán dados por

$$R = \{(2,2), (2,4), (3,3), (4,4), (5,5)\}, \text{ donde}$$

i) Reflexiva

$$(2,2), (3,3), (4,4), (5,5) \in R, \text{ por tanto, } R \text{ es reflexiva}$$

ii) Antisimetrica

$$(2,4) \in R \wedge (4,2) \notin R, \text{ por tanto, } R \text{ es antisimetrica}$$

iii) Transitiva

$$(2,2) \wedge (2,4) \in R \rightarrow (2,4) \in R$$

$$(2,4) \wedge (4,4) \in R \rightarrow (2,4) \in R, \text{ entonces } R \text{ es transitiva}$$

Por lo tanto, R es de orden parcial

Definición 2.1.1.4- (Orden débil por componentes)

Sean $\bar{x}, \bar{y} \in \mathbb{R}^n$, entonces $\bar{x} \preceq \bar{y}$ si y solo si $x_k \leq y_k, k = 1, 2, \dots, n$

Definición 2.1.1.5- (Orden por componentes)

Sean $\bar{x}, \bar{y} \in \mathbb{R}^n$, entonces $\bar{x} \preceq \bar{y}$ si y solo si $x_k \leq y_k, k = 1, 2, \dots, n, \bar{x} \neq \bar{y}$

Definición 2.1.1.6 - (Orden estricto por componentes)

Sean $\bar{x}, \bar{y} \in \mathbb{R}^n$, entonces $\bar{x} < \bar{y}$ si y solo si $x_k < y_k, k = 1, 2, \dots, n$

Considerando las definiciones de orden débil y estricto definiremos los subconjuntos sobre el espacio euclidiano \mathbb{R}^n .

$\mathbb{R}_{\geq}^n = \{y \in \mathbb{R}^n: y \geq 0\}$, representa el ortante no negativo de \mathbb{R}^n

$\mathbb{R}_{>}^n = \{y \in \mathbb{R}^n: y > 0\} = \mathbb{R}_{\geq}^n \setminus \{0\}$

$\mathbb{R}_{>}^n = \{y \in \mathbb{R}^n: y > 0\} = \text{int}(\mathbb{R}_{\geq}^n)$, representa el ortante positivo \mathbb{R}^n

2.1.2.- Conos y relaciones de orden

Definición 2.1.2.1 Sea C un subconjunto no vacío de \mathbb{R}^n . Diremos que

- (a) C es un cono, si dados $\bar{x} \in C$ y $\lambda \geq 0$, entonces $\lambda\bar{x} \in C$.
- (b) Un cono C es llamado puntiagudo, si dados $\bar{x} \in C$ y $-\bar{x} \in C$ entonces $\bar{x} = 0$.

Por ejemplo, sea el conjunto $C = \{d \in \mathbb{R}^2: d_k \geq 0, k = 1,2\} = \mathbb{R}_{\geq}^2$, donde se representa de forma gráfica en la Figura 1



Figura 2.1. Ilustración de cono

y además en \mathbb{R}^n estará dado por $C = \{d \in \mathbb{R}^n: d_k \geq 0, k = 1,2, \dots, p\} = \mathbb{R}_+^n$, que es un cono puntiagudo, el cual se usará en este trabajo.

Definición 2.1.2.2- Un cono C en R^p es llamado:

- i) No trivial o propio, si $C \neq \emptyset$ y $C \neq \mathbb{R}^p$
- ii) Convexo, si $\alpha\bar{d}_1 + (1 - \alpha)\bar{d}_2 \in C \forall \bar{d}_1, \bar{d}_2 \in C, y 0 \leq \alpha \leq 1$

Por ejemplo, en la figura 2.2 se muestra al cono $C = \mathbb{R}_{\geq}^3 = \{d \in \mathbb{R}^3: d_k \geq 0, k = 1,2,3\}$, de elementos no negativos, determinado por el orden débil de componentes (ver definición 2.1.1.4).

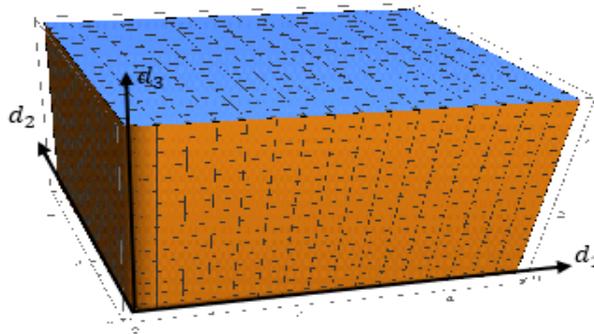


Figura 2.2. Ilustración de cono convexo

Definición 2.1.2.3 -Sea \mathcal{R} una relación binaria sobre \mathbb{R}^n , donde definimos:

$$C_{\mathcal{R}} = \{\bar{y} - \bar{x} : \bar{x} \mathcal{R} \bar{y}\}$$

como el conjunto de elementos no negativos de \mathbb{R}^n de acuerdo con la relación \mathcal{R} .

Definición 2.1.2.4-Diremos que la relación binaria \mathcal{R} es

1.-Compatible con la multiplicación escalar, si para cualquier $\bar{x}, \bar{y} \in \mathcal{R}$ y para todo $\alpha \in \mathbb{R}_>$ se cumple que $(\alpha\bar{x}, \alpha\bar{y}) \in \mathcal{R}$.

2.- Compatible con la adición, si para cualquier $\bar{x}, \bar{y} \in \mathcal{R}$ y para todo $\bar{z} \in \mathbb{R}^n$ se cumple que $(\bar{x} + \bar{z}) \mathcal{R} (\bar{y} + \bar{z})$.

Definición 2.1.2.5. Sea C un cono y la relación de orden \mathcal{R}_C definida por

$$\bar{x} \mathcal{R}_C \bar{y} \Leftrightarrow \bar{y} - \bar{x} \in C$$

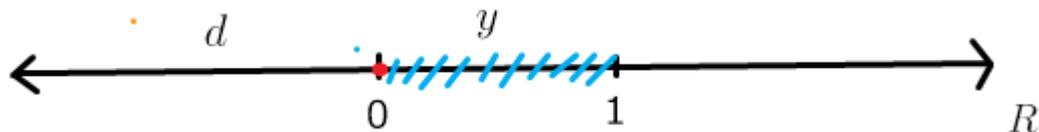
Proposición 2.1.2.6. Sea C un cono entonces la relación \mathcal{R}_C es compatible con la multiplicación escalar y la adición en \mathbb{R}^n

Definición 2.1.2.7. Sea $S \subseteq \mathbb{R}^n$ un cono convexo y $x \in S$. Diremos que d es un vector normal a S en x si se cumple $\langle d, y - x \rangle \leq 0 \quad \forall y \in S$.

Definición 2.1.2.8. Sea $C \subset \mathbb{R}^n$ y $\bar{x} \in C$. El cono normal al punto \bar{x} relacionado con el conjunto C este dado por $N_C(\bar{x}) = \{v \in \mathbb{R}^n : \langle v, y - \bar{x} \rangle \leq 0, \quad \forall y \in C\}$

Ejemplo 2.1.2.9. Sea el conjunto convexo $C = \{x \in \mathbb{R} / 0 \leq x \leq 1\}$, determinaremos el cono normal para cada elemento del conjunto C .

i. Para $\bar{x} = 0$

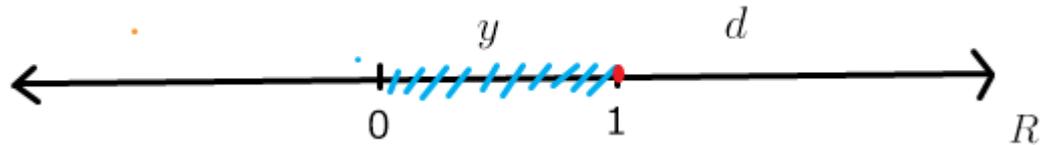


$$N_C(0) = \{d \in \mathbb{R} : \langle d, y - 0 \rangle \leq 0, \quad \forall y \in C\}$$

Si $0 \leq y \leq 1$, y $d \leq 0 \rightarrow dy \leq 0$, por lo tanto

$$N_C(0) = \langle -\infty, 0 \rangle$$

ii. Para $\bar{x} = 1$

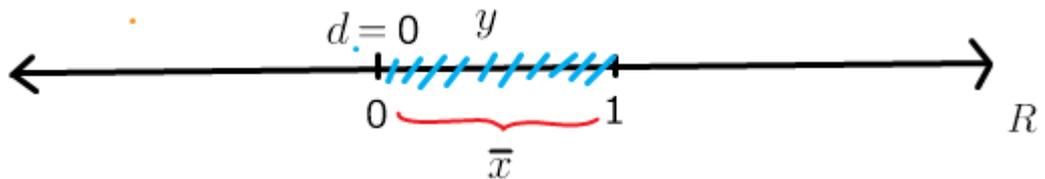


$$N_C(1) = \{d \in \mathbb{R} : \langle d, y - 1 \rangle \leq 0, \quad \forall y \in C\}$$

Si $0 \leq y \leq 1 \rightarrow -1 \leq y - 1 \leq 0$ y $d \geq 0 \rightarrow d(y - 1) \leq 0$, por lo tanto

$$N_C(1) = [0, +\infty)$$

iii. Para $\bar{x} \in \langle 0, 1 \rangle$



$$N_C(\bar{x}) = \{d \in \mathbb{R} : \langle d, y - \bar{x} \rangle \leq 0, \quad \forall y \in C\}$$

Si $0 < \bar{x} < 1 \rightarrow -1 < -\bar{x} < 0$ por otro lado $0 \leq y \leq 1$, entonces $y - 1 < y - \bar{x} < y$, luego $-1 \leq y - \bar{x} \leq 1$, ahora considerando los valores de d , tenemos lo siguiente:

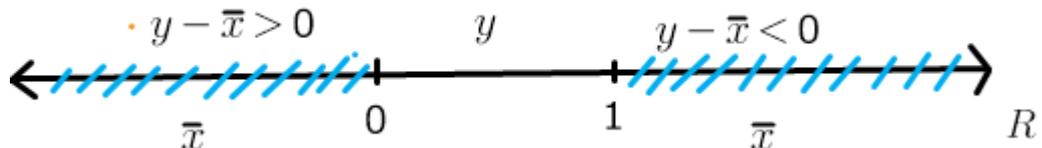
1. Si $d = 0$ se tiene que $d(y - \bar{x}) \leq 0 \rightarrow N_C(\bar{x}) = 0$.
2. Si $d < 0$ se tiene que $d \leq d(y - \bar{x}) \leq -d \rightarrow N_C(\bar{x}) = \emptyset$
3. Si $d > 0$ se tiene que $-d \leq d(y - \bar{x}) \leq d \rightarrow d(y - \bar{x}) + d \geq 0$

$d((y - \bar{x}) + 1) \geq 0 \rightarrow y \geq \bar{x} - 1 \rightarrow y \geq \bar{x} - 1 > -1 \rightarrow y > -1$, lo cual es una contradicción dado que $0 \leq y \leq 1 \rightarrow -1 \leq -y \leq 0$

por lo tanto $N_C(\bar{x}) = \emptyset$, finalmente

$$N_C(\bar{x}) = 0, \quad \bar{x} \in \langle 0, 1 \rangle$$

iv. Para $\bar{x} \notin [0, 1]$ se tiene $y \neq \bar{x} \rightarrow y - \bar{x} \neq 0 \rightarrow [(y - \bar{x}) > 0 \vee (y - \bar{x}) < 0]$



Ahora considerando los valores de d , y $(y - \bar{x}) > 0 \vee (y - \bar{x}) < 0$ tenemos lo siguiente:

1. Si $d = 0$ se tiene que $0 > 0 \vee 0 < 0$ lo cual es una contradicción por lo tanto $N_C(\bar{x}) = \emptyset$
2. Si $d < 0$ se tiene que $d(y - \bar{x}) < 0 \vee d(y - \bar{x}) > 0 \rightarrow N_C(\bar{x}) = \emptyset$
3. Si $d > 0$ se tiene que $d(y - \bar{x}) > 0 \vee d(y - \bar{x}) < 0 \rightarrow N_C(\bar{x}) = \emptyset$

Proposición 2.1.2.9. Sea C un conjunto convexo en \mathbb{R}^n , si $\bar{x} \in \text{int}(C)$ entonces

$$N_C(\bar{x}) = \{0\}$$

Demostración. Sea $\bar{x} \in C$ y $\delta > 0$ talque $B(\bar{x}, \delta) \subset C$, consideremos que

$$v \in N_C(\bar{x})$$

por definición $\langle v, y - \bar{x} \rangle \leq 0$, para todo $y \in C$.

Ahora para un $t > 0$ suficientemente pequeño $\bar{x} + tv \in B(\bar{x}, \delta) \subset C$, por lo tanto $\langle v, \bar{x} + tv - \bar{x} \rangle \leq 0 \rightarrow \langle v, tv \rangle \leq 0 \rightarrow t\langle v, v \rangle \leq 0, \rightarrow t\|v\|^2 \leq 0 \rightarrow \|v\|^2 \leq 0$, implicando que $v = 0$.

Proposición 2.1.2.10. Sea $C \subset \mathbb{R}^n$ un conjunto convexo y sea $\bar{x} \in C$ entonces $N_C(\bar{x})$ es un cono convexo.

Demostración. Mostraremos que $N_C(\bar{x})$ es un cono y luego que es convexo.

i.- $N_C(\bar{x})$ es un cono.

Consideremos $v \in N_C(\bar{x})$ y $\lambda \geq 0$. Por la definición de cono normal, témenos que $\langle v, y - \bar{x} \rangle \leq 0$ para cualquier $y \in C$.

Si $\lambda \geq 0$ entonces $\lambda\langle v, y - \bar{x} \rangle = \langle \lambda v, y - \bar{x} \rangle \leq 0$ para cualquier $x \in C$.

Así tenemos que $\lambda v \in N_C(\bar{x})$ por lo tanto, el resultado.

ii.- $N_C(\bar{x})$ es un cono normal convexo. Sea $v_1, v_2 \in N_C(\bar{x})$ y $0 \leq \lambda \leq 1$. Entonces tenemos que $\langle v_1, y - \bar{x} \rangle \leq 0$ y $\langle v_2, y - \bar{x} \rangle \leq 0$ para cualquier $x \in C$. Luego se tiene

$$\langle \lambda v_1 + (1 - \lambda)v_2, y - \bar{x} \rangle = \lambda\langle v_1, y - \bar{x} \rangle + (1 - \lambda)\langle v_2, y - \bar{x} \rangle \leq 0, \quad x \in C$$

Así se concluye que $\lambda v_1 + (1 - \lambda)v_2 \in N_C(\bar{x})$ y por lo tanto el cono normal es convexo.

2.1.3.- Eficiencia, débil eficiencia y no dominancia de soluciones

Siguiendo el modelo matemático definido anteriormente, es decir

$$\text{Optimizar } F(x) = \text{Opt}\{(F_1(x), \dots, F_m(x)) : x \in X \subset \mathbb{R}^n\} \quad (1.1)$$

donde $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $F(x) = (F_1(x), \dots, F_m(x))$, $m \geq 2$, denota el vector de funciones objetivo definido en el espacio denominado variable o llamado espacio de decisión \mathbb{R}^n , con valores en el espacio objetivo o llamado espacio criterio \mathbb{R}^m , $X \subseteq \mathbb{R}^n$ representa el **espacio factible** de decisión o **conjunto factible**, $\bar{x} = (x_1, x_2, \dots, x_n)$ representa el vector de las n variables de decisión, $Y = F(X) \subset \mathbb{R}^m$, representa el conjunto factible definido en el **espacio objetivo**, se presentara las siguientes definiciones.

Definición 2.1.3.1. (Miettinen [51], Definición 2.2.1). Un punto $\hat{x} \in \mathbb{R}^n$ es una solución de Pareto (también conocida como solución eficiente o óptimo de Pareto) para el problema (1.1), si no existe otro $x \in \mathbb{R}^n$ tal que $F_i(x) \leq F_i(\hat{x})$, para todo $i \in \{1, 2, \dots, m\}$ y $F_i(x) < F_i(\hat{x})$, para algún $i \in \{1, 2, \dots, m\}$.

Por ejemplo, consideremos la función objetivo $F: \mathbb{R} \rightarrow \mathbb{R}^2$, definida mediante $F(x) = (x^2, (x-1)^2)$, donde $F: \mathbb{R} \rightarrow \mathbb{R}^2$ representa el vector de funciones objetivo a optimizar y \mathbb{R} el espacio de decisión. Por lo tanto, el problema multiobjetivo con restricciones estará dado por

$$\min_{x \in X} F(x) = \min\{(x^2, (x-1)^2) : x \in X \subset \mathbb{R}^n\}$$

En la siguiente grafica de la Figura 2.3 se observa el grafico de las funciones objetivo

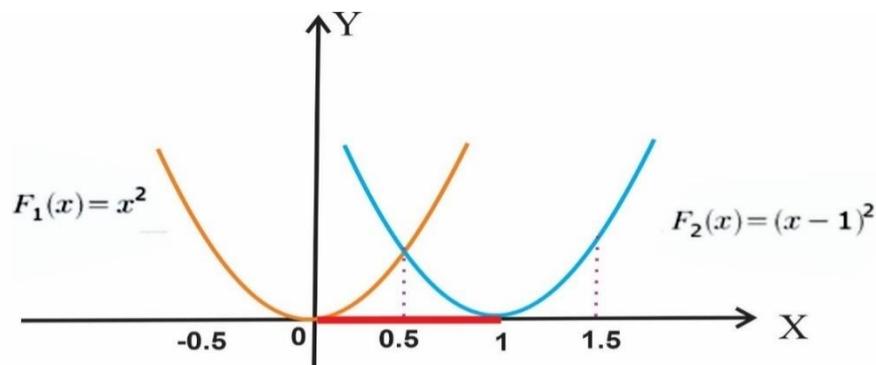


Figura 2.3. Soluciones eficientes de Pareto

En la cual el punto $x = 0$ es una solución de Pareto (o óptimo de Pareto) dado que no existe otro elemento x distinto de $x = 0$, talque $F_1(x) \leq F_1(0) = 0$ y $F_2(x) \leq F_2(0) = 1$, de forma similar para el punto $x = 0.5$ es una solución de Pareto (o óptimo de Pareto) dado que no existe otro elemento x distinto de $x = 0.5$, talque $F_1(x) \leq F_1(0.5) = 0.25$ y $F_2(x) \leq F_2(0.5) = 0.25$.

Si queremos que $F_1(x) \leq 0.25$ entonces x pertenecería al intervalo $[-0.5, 0.5]$ y para que $F_2(x) \leq 0.25$, x pertenecería al intervalo $[0.5, 1.5]$, por lo tanto, las soluciones de Pareto (o óptimo de Pareto) será el conjunto $X = \{x \in \mathbb{R}: 0 \leq x \leq 1\}$, descrito geoméricamente por la línea roja de la figura.

Observación 2.1.3.2

De acuerdo con la definición anterior y en el caso de minimización del problema (1.1)

se considera lo siguiente.

- Si \hat{x} es eficiente, entonces $F(\hat{x})$ es llamada solución no dominada
- Si $x^t, x^k \in X$ y $F(x^t) \leq F(x^k)$, entonces x^t domina a x^k
(denotado por $x^t \preceq x^k$) y $F(x^t)$ domina a $F(x^k)$
- El conjunto de todas las soluciones eficientes (o soluciones de Pareto), denotado por X_E , es denominado **conjunto eficiente**.
- El conjunto de los puntos no dominados $\hat{y} = F(\hat{x}) \in Y$ donde $\hat{x} \in X_E$ denotado por $Y_N = F(X_E)$, es denominado **conjunto no dominado** o **frontera de Pareto**.

Definición 2.1.3.3. (Miettinen [51], Definición 2.5.1) Un punto $\hat{x} \in R^n$ es una solución débil de Pareto (también conocida como solución débilmente eficiente o óptimo de Pareto) para el problema (1.1), para minimización de F , un punto $\hat{x} \in R^n$ es una solución débil de Pareto, si no existe otro $x \in R^n$ tal que $F_i(x) < F_i(\hat{x})$, para todo $i \in \{1, 2, \dots, m\}$ y para maximización de F , un punto $\hat{x} \in R^n$ es una solución débil de Pareto, si no existe otro $x \in R^n$ tal que $F_i(\hat{x}) < F_i(x)$, para todo $i \in \{1, 2, \dots, m\}$. Además, si \hat{x} es una solución débilmente eficiente, entonces el punto $F(\hat{x}) = \hat{y}$ es denominado **débilmente no dominado**, por otro lado, la **débil eficiencia y no dominancia** de conjuntos, donde se denota por X_{wE} y Y_{wN} .

Por ejemplo, consideremos el problema de optimización vectorial definido mediante

$$\max F(x, y) = (x, y)$$

$$\text{S.a } R = \{(x, y) \in \mathbb{R}^2 / 0 \leq x \leq 1, 0 \leq y \leq 1\}$$

Donde dado que la función F es la función identidad, el conjunto imagen $Y = F(R)$ es equivalente al conjunto R , asumimos que $C = \mathbb{R}_{\geq}^2$, además describiremos el conjunto eficiente X_E y conjunto no dominado o frontera de Pareto $Y_N = F(X_E)$ en la siguiente figura 4

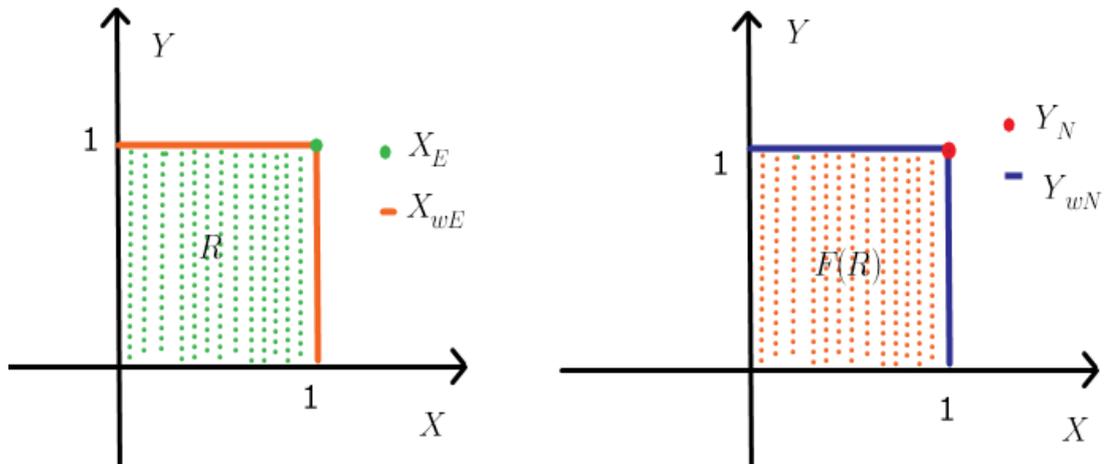


Figura 2.4. Soluciones débiles de Pareto

del grafico observamos que el punto $\hat{x} = (1,1)$ es la única solución eficiente del problema es decir que no existe $(x, y) \neq (1,1)$, con $(x, y) \in R$ tal que $F(\hat{x}) \preceq F(\bar{x})$ implicando que $F_1(1) \leq F_1(x)$, $F_2(1) \leq F_2(x)$. En efecto si $\hat{x} = (1,1)$, $\bar{x} = (x, 1)$

$$(1,1) \preceq (x, 1) \rightarrow \bar{0} \preceq (x, 1) - (1,1)$$

$$\bar{0} \preceq (x - 1, 0) \Leftrightarrow 0 \leq (x - 1) \wedge 0 \leq 0$$

Concluyendo que es una contradicción dado que $0 \leq x \leq 1$ por lo tanto $\hat{x} = (1,1)$ es la única solución eficiente del problema $Y_N = \{(1,1)\}$.

Por otro lado, el conjunto de soluciones débilmente eficiente será

$X_{wE} = \{(x, y) \in R: x = 1 \text{ o } y = 1\}$, veamos que el punto de la forma $(x_0, 1)$ cumple con la definición de solución débilmente eficiente. En efecto, sea $\hat{x} = (x_0, 1)$ una solución débilmente eficiente por lo tanto no existe un $\bar{x} \in R$ tal que $F(\hat{x}) < F(\bar{x}), \forall \bar{x} \in R$.

Supongamos que lo hubiera, es decir $F(x_0, 1) < F(x, 1) \rightarrow (x_0, 1) < (x, 1)$, implicando $\bar{0} < (x - x_0, 1)$, luego $(0,0) < (x - x_0, 1) \Leftrightarrow 0 \neq 0 \wedge 0 < x - x_0$, por

lo tanto, las soluciones de la forma $(x_0, 1)$ son débilmente eficientes, de manera análoga para las soluciones de la forma $(y_0, 1)$ cumple con la condición.

Finalmente

$$Y_{wN} = \{(x, y) \in F(R): x = 1 \text{ o } y = 1\}$$

Observación 2.1.3.4

Denotaremos al conjunto de soluciones de Pareto por: $\arg \min\{F(x): x \in R^n\}$ y al conjunto de soluciones débiles de Pareto por: $\arg \min_w\{F(x): x \in R^n\}$, en el caso de minimización del problema (1.1) y al valor óptimo de optimizar el problema (1.1) lo denotaremos F^* .

Por ejemplo, consideremos la función $f(x) = \text{sen}(x)$, donde el conjunto de soluciones de Pareto estará por

$$\arg \min\{f(x): x \in R\} = \left\{x \in R / x_n = \frac{3\pi}{2} + 2n\pi, n \in \mathbb{Z} / f(x_n) = \inf(f) = -1\right\}$$

Definición 2.1.3.5. Una función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ es una representación escalar estricta de la aplicación $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ cuando dados $x, \bar{x} \in \mathbb{R}^n$ tenemos que:

$$F(x) \preceq F(\bar{x}) \Rightarrow f(x) \leq f(\bar{x}) \text{ y } F(x) \prec F(\bar{x}) \Rightarrow f(x) < f(\bar{x})$$

Además, diremos que f es una representación escalar débil de F si

$$F(x) \prec F(\bar{x}) \Rightarrow f(x) < f(\bar{x})$$

Observar que toda representación escalar f estricta de F es también una representación escalar débil de F .

Proposición 2.1.3.6. La función $f: \mathbb{R}^n \rightarrow \mathbb{R}$ es una representación escalar estricta de $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ si y solo si existe una función estrictamente creciente $g: \mathbb{R}^m \rightarrow \mathbb{R}$ tal que $f = g \circ F$.

Demostración. Véase [38], (Proposición 2.3)

Ejemplo 2.1.3.7. Sea la aplicación $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$. La función $f: \mathbb{R}^n \rightarrow \mathbb{R}$, definida por $f(x) = \langle F(x), u \rangle$, con $u \in \mathbb{R}_+^m \setminus \{0\}$, fijo será una representación escalar estricta de F .

En efecto, dados $x, \bar{x} \in \mathbb{R}^n$, si $F(x) \leq F(\bar{x})$, entonces $F(\bar{x}) - F(x) \in \mathbb{R}_+^m$ implicando que $F_i(\bar{x}) - F_i(x) \geq 0$ para todo $i \in \{1, \dots, m\}$. Luego para todo $u \in \mathbb{R}_+^n \setminus \{0\}$, tenemos

$$[F_i(\bar{x}) - F_i(x)]u_i \geq 0, \forall i \in \{1, \dots, m\}, \text{ lo que implica que}$$

$$\langle F(\bar{x}) - F(x), u \rangle \geq 0, \forall u \in \mathbb{R}_+^n \setminus \{0\}, \text{ equivalente a } \langle F(x), u \rangle \leq \langle F(\bar{x}), u \rangle.$$

Así mismo si $F(x) < F(\bar{x})$, entonces $F(\bar{x}) - F(x) \in \mathbb{R}_{++}^m$ implicando que $F_i(\bar{x}) - F_i(x) > 0$ para todo $i \in \{1, \dots, m\}$. Luego para todo $u \in \mathbb{R}_+^n \setminus \{0\}$, tenemos

$$[F_i(\bar{x}) - F_i(x)]u_i > 0, \forall i \in \{1, \dots, m\}, \text{ lo que implica que}$$

$$\langle F(\bar{x}) - F(x), u \rangle > 0, \forall u \in \mathbb{R}_+^n \setminus \{0\}, \text{ equivalente a } \langle F(x), u \rangle < \langle F(\bar{x}), u \rangle.$$

Proposición 2.1.3.8. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una representación escalar débil de la aplicación $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\arg \min\{f(x): x \in \mathbb{R}^n\}$ conjunto de minimizadores de f . Entonces tenemos $\arg \min\{f(x): x \in \mathbb{R}^n\} \subseteq \arg \min_w\{f(x): x \in \mathbb{R}^n\}$.

Demostración

Sea $\bar{x} \in \arg \min\{f(x): x \in \mathbb{R}^n\}$, donde f una representación escalar débil arbitraria de F . Supongamos que $\bar{x} \notin \arg \min_w\{F(x): x \in \mathbb{R}^n\}$, entonces existe $\hat{x} \in \mathbb{R}^n$ talque $F(\hat{x}) < F(\bar{x})$.

Luego como f es una representación débil de F , tendremos que $f(\hat{x}) < f(\bar{x})$, por lo que nos lleva a una contradicción dado que $\bar{x} \in \arg \min\{f(x): x \in \mathbb{R}^n\}$.

Definición 2.1.3.9 - Dominancia de Pareto (ver Sawaragi [67]).

Sean dos soluciones $x^t, x^k \in \mathbb{R}^m$ se dice que x^t domina a x^k (denotado por $x^t \leq x^k$), sí y solo si x^t es parcialmente menor que x^k , es decir $F_i(x^t) < F_i(x^k)$, para todo $i = 1, 2, \dots, m$, y existe al menos un $j \in \{1, 2, \dots, m\}$.

Por ejemplo, sea el problema

$$\min_{x \in X} F(x) = \min\{(F_1(x), F_2(x)) : x \in X \subset \mathbb{R}^2\} \text{ y consideremos que a}$$

$F(x^t) = (F_1(x^t), F_2(x^t))$, se le aplica el concepto de Pareto y de acuerdo con la gráfica de la figura 5.

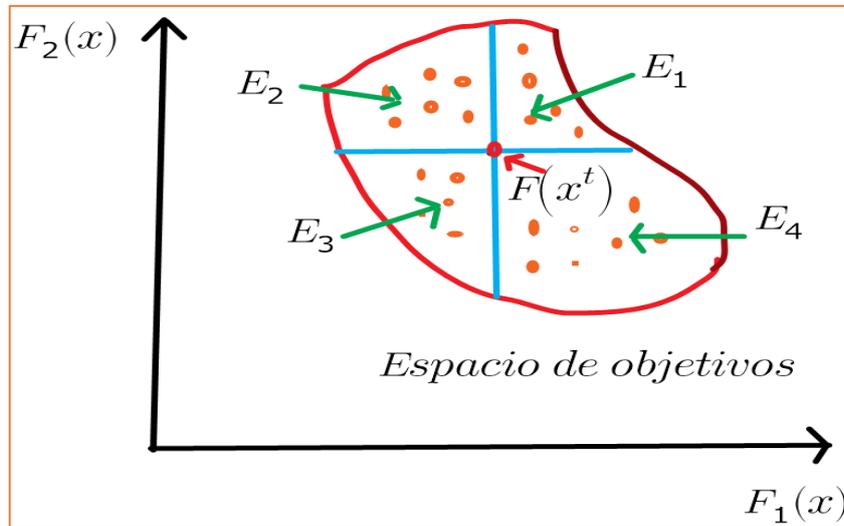


Figura 2.5. Puntos no dominados y puntos débilmente no dominados

Donde tomando la solución $F(x^t)$ definida en el espacio de los objetivos como referencia, se define cuatro regiones E_1, E_2, E_3, E_4 , la cual las posibles relaciones de dominancia de Pareto entre soluciones serán de la siguiente forma:

- $F(x^t)$, domina las soluciones contenidas en la región E_1
- $F(x^t)$, es incomparable con las soluciones contenidas en la región E_2
- $F(x^t)$, es dominada por las soluciones contenidas en la región E_3
- $F(x^t)$, es incomparable con las soluciones contenidas en la región E_4

2.1.4.-Conceptos básicos de análisis real

2.1.4.-Producto interno en el espacio Euclidiano \mathbb{R}^n

Definición 2.1.4.1.-Sea $\bar{x}, \bar{y} \in \mathbb{R}^n$, el producto interno de \bar{x} e \bar{y} denotado por $\langle \bar{x}, \bar{y} \rangle$ es el número real $\langle \bar{x}, \bar{y} \rangle = \sum_{i=1}^n x_i y_i$.

Teorema 2.1.4.2.-Sean $\bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^n$, y $\alpha \in \mathbb{R}$, se cumple que.

1. $\langle \bar{x}, \bar{x} \rangle \geq 0, \quad \forall \bar{x} \in \mathbb{R}^n$
2. $\langle \bar{x}, \bar{x} \rangle = 0, \Leftrightarrow \bar{x} = \bar{0}$
3. $\langle \bar{x}, \bar{y} \rangle = \langle \bar{y}, \bar{x} \rangle, \quad \forall \bar{x}, \bar{y} \in \mathbb{R}^n$
4. $\langle \bar{x} + \bar{z}, \bar{y} \rangle = \langle \bar{x}, \bar{y} \rangle + \langle \bar{z}, \bar{y} \rangle, \quad \forall \bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^n$
5. $\langle \alpha \bar{x}, \bar{y} \rangle = \alpha \langle \bar{x}, \bar{y} \rangle, \quad \forall \alpha \in \mathbb{R} \text{ y } \forall \bar{x}, \bar{y} \in \mathbb{R}^n$

Demostración. Ver [7], página 5.

2.1.5 La norma en el espacio Euclidiano \mathbb{R}^n

Definición 2.1.5.1.- Sea $\bar{x} \in \mathbb{R}^n$, la norma de \bar{x} , denotado por $\|\bar{x}\|$, es definido mediante $\|\bar{x}\| = \sqrt{\langle \bar{x}, \bar{x} \rangle}$

Teorema 2.1.5.2.- (Desigualdad de Cauchy-Schwarz)

1.- $|\langle \bar{x}, \bar{y} \rangle| \leq \|\bar{x}\| \|\bar{y}\|$, para todo $\bar{x}, \bar{y} \in \mathbb{R}^n$

2.- $|\langle \bar{x}, \bar{y} \rangle| = \|\bar{x}\| \|\bar{y}\|$, si solo si existe $\alpha \in \mathbb{R}$ talque $\bar{x} = \alpha \bar{y}$

Demostración. Véase [7], página 8.

Teorema 2.1.5.3.- Sean $\bar{x}, \bar{y} \in \mathbb{R}^n$, y $\alpha \in \mathbb{R}$, se cumple las siguientes propiedades.

1. $\|\bar{x}\| \geq 0$, para todo $\bar{x} \in \mathbb{R}^n$

2. $\|\bar{x}\| = 0$, si solo si $\bar{x} = \bar{0}$

3. $\|\bar{x} + \bar{y}\| \leq \|\bar{x}\| + \|\bar{y}\|$, para todo $\bar{x}, \bar{y} \in \mathbb{R}^n$ (Desigualdad triangular)

4. $\|\alpha \bar{x}\| = |\alpha| \|\bar{x}\|$, para todo $\alpha \in \mathbb{R}$ y para todo $\bar{x} \in \mathbb{R}^n$

Demostración. Véase [7], página 9.

Proposición 2.1.5.4.- Sean $\bar{x}, \bar{y} \in \mathbb{R}^n$, se cumple las siguientes propiedades.

1.- $|\|\bar{x}\| - \|\bar{y}\|| \leq \|\bar{x} - \bar{y}\|$, para todo $\bar{x}, \bar{y} \in \mathbb{R}^n$

2.- $\|\bar{x} + \bar{y}\|^2 + \|\bar{x} - \bar{y}\|^2 = 2(\|\bar{x}\|^2 + \|\bar{y}\|^2)$, para todo $\bar{x}, \bar{y} \in \mathbb{R}^n$ (Identidad del paralelogramo)

3.- $\|\bar{x} + \bar{y}\|^2 - \|\bar{x} - \bar{y}\|^2 = 4\langle \bar{x}, \bar{y} \rangle$, para todo $\bar{x}, \bar{y} \in \mathbb{R}^n$ (Identidad de polarización)

Demostración. Véase [7], página 10.

2.1.6.-Bolas en el espacio Euclidiano \mathbb{R}^n

Definición 2.1.6.1.-Sea $\bar{x}_0 \in \mathbb{R}^n$, y $\delta > 0$.

1.-Una bola abierta de centro \bar{x}_0 y radio δ es el conjunto definido por:

$$B(\bar{x}_0, r) = B_\delta(\bar{x}_0) = \{\bar{x} \in \mathbb{R}^n: \|\bar{x} - \bar{x}_0\| < \delta\}$$

2.- Una bola cerrada de centro \bar{x}_0 y radio δ es el conjunto definido por:

$$B[\bar{x}_0, r] = B_r[\bar{x}_0] = \{\bar{x} \in \mathbb{R}^n : \|\bar{x} - \bar{x}_0\| \leq r\}$$

2.1.7.-Sucesiones y subsucesiones en el espacio euclidiano \mathbb{R}^n

Definición 2.1.7.1.-Se denomina sucesión en \mathbb{R}^n a la función $x: \mathbb{N} \rightarrow \mathbb{R}^n$ que a cada $k \in \mathbb{N}$ le asocia el vector $x(k) = x_k \in \mathbb{R}^n$ denominado k -ésimo término de la sucesión y además se denota por $\{x_k\}_{k \in \mathbb{N}}$, o de forma equivalente por $\{x_k\} \subseteq \mathbb{R}^n$.

Definición 2.1.7.2.-Sea $\{x_k\} \subseteq \mathbb{R}$

1.-La sucesión $\{x_k\}$ es creciente si:

$$x_k \leq x_{k+1} \quad \forall k \geq 1$$

2.-La sucesión $\{x_k\}$ es decreciente si:

$$x_k \geq x_{k+1} \quad \forall k \geq 1$$

Por ejemplo, sea la sucesión $x: \mathbb{N} \rightarrow \mathbb{R}$ definida mediante $x_k = \left\{\frac{1}{2^k}\right\}$, $k \in \mathbb{N}$ es decreciente es decir $x_k \geq x_{k+1} \quad \forall k \geq 1$

Definición 2.1.7.3.-Sea $\{x_k\} \subseteq \mathbb{R}^n$, diremos que $\bar{x} \in \mathbb{R}^n$ es el límite de la $\{x_k\}$ cuando $k \rightarrow \infty$, y se expresa por $\lim_{k \rightarrow \infty} x_k = \bar{x}$ si para $\varepsilon > 0$, existe un $k_0 \in \mathbb{N}$ tal que si $k \geq k_0$, entonces $\|x_k - \bar{x}\| < \varepsilon$.

De forma simbólicamente se expresa

$$\lim_{k \rightarrow \infty} x_k = \bar{x} \Leftrightarrow \forall \varepsilon > 0, \exists k_0 \in \mathbb{N}; k \geq k_0 \Rightarrow \|x_k - \bar{x}\| < \varepsilon$$

Por ejemplo, la sucesión $x: \mathbb{N} \rightarrow \mathbb{R}$ definida mediante $x_k = \left\{\frac{1}{2^k}\right\}$, $k \in \mathbb{N}$ tiene como límite a 0, dado que $\lim_{k \rightarrow \infty} x_k = 0$.

Definición 2.1.7.4.-Sea $\{x_k\} \subseteq \mathbb{R}^n$, diremos que:

1.- $\{x_k\}$ es una sucesión convergente, si solo si existe $\bar{x} \in \mathbb{R}^n$ tal que $\lim_{k \rightarrow \infty} x_k = \bar{x}$

2.- $\{x_k\}$ es una sucesión divergente si solo si no es convergente.

Por ejemplo, la sucesión $x: \mathbb{N} \rightarrow \mathbb{R}$ definida mediante $x_k = \left\{\frac{1}{2^k}\right\}$, $k \in \mathbb{N}$ es convergente, dado que $\lim_{k \rightarrow \infty} x_k = 0$

Corolario 2.1.7.5.- Si $\{x_k\}, \{z_k\} \subseteq \mathbb{R}^n$ y $\{\lambda_k\} \subseteq \mathbb{R}$ son sucesiones convergentes, entonces $\{x_k \pm z_k\}, \{\lambda_k x_k\} \subseteq \mathbb{R}^n, \{\langle x_k, z_k \rangle\}, \{\|x_k\|\}, \{d(x_k, z_k)\} \subseteq \mathbb{R}^n$ son también sucesiones convergentes y además se cumple que:

$$1.- \lim_{k \rightarrow \infty} (x_k \pm z_k) = \lim_{k \rightarrow \infty} x_k \pm \lim_{k \rightarrow \infty} z_k$$

$$2.- \lim_{k \rightarrow \infty} (\lambda_k x_k) = \left(\lim_{k \rightarrow \infty} \lambda_k \right) \left(\lim_{k \rightarrow \infty} x_k \right)$$

$$3.- \lim_{k \rightarrow \infty} \langle x_k, z_k \rangle = \langle \lim_{k \rightarrow \infty} x_k, \lim_{k \rightarrow \infty} z_k \rangle$$

$$4.- \lim_{k \rightarrow \infty} \|x_k\| = \left\| \lim_{k \rightarrow \infty} x_k \right\|$$

$$5.- \lim_{k \rightarrow \infty} d(x_k, z_k) = d\left(\lim_{k \rightarrow \infty} x_k, \lim_{k \rightarrow \infty} z_k\right)$$

Demostración. Véase [7], página 21.

Definición 2.1.7.6.- Sea $\{x_k\} \subseteq \mathbb{R}^n$, decimos que $\{x_k\}$ es una sucesión acotada si existe $M > 0$ talque $\|x_k\| \leq M, \forall k \in \mathbb{N}$.

Proposición 2.1.7.7.- Si $\{x_k\} \subseteq \mathbb{R}^n$ es una sucesión convergente, entonces $\{x_k\}$ es acotada.

Demostración. Véase [7], página 21.

Definición 2.1.7.8.- Sea $\{x_k\} \subseteq \mathbb{R}^n$ y $j : \mathbb{N} \rightarrow \mathbb{N}$ una función creciente (es decir $j_1 < j_2 < \dots < j_k < \dots$). La función compuesta $x \circ j : \mathbb{N} \rightarrow \mathbb{R}^n$ que a cada $k \in \mathbb{N}$ le asocia el vector $x \circ j(k) = x(j(k)) = x_{j_k}$ la cual es denominada subsucesión de $\{x_k\}$.

La notación $\{x_{j_k}\} \subseteq \{x_k\}$ significa que $\{x_{j_k}\}$ es una subsucesión $\{x_k\}$.

Proposición 2.1.7.9.- Sea $\{z_k\} \subseteq \mathbb{R}^n$. Si $\lim_{k \rightarrow \infty} z_k = \bar{z}$ entonces toda subsucesión $\{z_{j_k}\}$ de $\{z_k\}$ converge hacia \bar{z}

Demostración. Véase [7], página 22.

Teorema 2.1.7.10.- Si la sucesión $\{z_k\} \subseteq \mathbb{R}^n$ es acotada en \mathbb{R}^n entonces posee una subsucesión convergente.

Demostración. Véase [7], página 22.

Definición 2.1.7.11.-El punto $\bar{z} \in \mathbb{R}^n$ es el valor adherente a la sucesión $\{z_k\} \subseteq \mathbb{R}^n$ si y solo si existe $\{z_{j_k}\} \subset \{z_k\}$ tal que $\lim_{k \rightarrow \infty} z_{j_k} = \bar{z}$

Teorema 2.1.7.12.- Sea una sucesión $\{z_k\} \subseteq \mathbb{R}^n$. Se tiene las siguientes equivalencias:

- 1.- $\{z_k\}$ es convergente.
- 2.- $\{z_k\}$ es acotada y posee un único valor adherente.

Demostración. Véase [7], página 23.

2.1.8 Series en el Espacio Euclidiano \mathbb{R}^n

En la presente sección se presentará los conceptos básicos necesarios para la finalidad de esta tesis.

Definición 2.1.8.1. Sea $\{z_k\}_{k \geq 1}$ una sucesión de números reales. La serie infinita de números esta denotada por:

$$\sum_{k=1}^{\infty} z_k = z_1 + z_2 + \dots + z_n + \dots$$

donde z_k es denominado el k -ésimo término de la serie.

Además de la serie de números reales $\sum_{k=1}^{\infty} z_k$, formaremos una sucesión $\{s_k\}_{k \geq 1}$ definida mediante la siguiente forma:

$s_1 = z_1, s_2 = z_1 + z_2, \dots, s_k = z_1 + z_2 + \dots + z_n$, donde a la sucesión $\{z_k\}_{k \geq 1}$ se le denomina sucesión de sumas parciales de la serie infinita, siendo s_k la k -ésima suma parcial de la serie.

Definición 2.1.8.2.-Sea la serie infinita $\sum_{k=1}^{\infty} z_k$ y la sucesión de sumas parciales $\{s_k\}_{k \geq 1}$.

1.-Si $\lim_{k \rightarrow +\infty} s_k = l$ existe, entonces diremos que la serie infinita $\sum_{k=1}^{\infty} z_k$ es convergente y converge hacia l siendo l un numero finito.

2.- Si la serie infinita $\sum_{k=1}^{\infty} z_k$ es convergente, la suma de la serie se puede escribir de la forma

$$\sum_{k=1}^{\infty} z_k = \lim_{k \rightarrow +\infty} s_k = l,$$

Ejemplo 2.1.8.3.- Hallar la suma de la serie infinita $\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)}$ en caso que sea convergente.

Solución. En efecto el n -ésimo término de la serie es $z_n = \frac{1}{(n+1)(n+2)}$ donde descomponiendo en fracciones parciales se tiene que:

$$z_n = \frac{1}{(n+1)(n+2)} = \frac{A}{(n+1)} + \frac{B}{(n+2)} \text{ y operando tenemos que } A = 1, B = -1, \text{ luego}$$

$$z_n = \frac{1}{(n+1)} - \frac{1}{(n+2)}, \text{ de la forma:}$$

$$z_1 = \frac{1}{2} - \frac{1}{3}$$

$$z_2 = \frac{1}{3} - \frac{1}{4}$$

.

.

$$z_n = \frac{1}{(n+1)} - \frac{1}{(n+2)}$$

$$\text{Donde } S_n = z_1 + z_2 + \dots + z_n = \frac{1}{2} - \frac{1}{(n+2)}$$

Por lo tanto $\lim_{n \rightarrow +\infty} S_n = \frac{1}{2}$ existe, entonces la serie infinita $\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)}$ es

convergente y su suma es igual a $\frac{1}{2}$, es decir $\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)} = \frac{1}{2}$.

Proposición 2.1.8.4.- Si $\sum_{k=1}^{\infty} z_k$ es convergente, entonces $\lim_{k \rightarrow +\infty} z_k = 0$.

Demostración. Vease [22], página 11

2.1.9. Conjuntos abiertos y cerrados en \mathbb{R}^n

Definición 2.1.9.1.-Sea $C \subseteq \mathbb{R}^n$:

- 1.-Diremos que $\bar{z} \in \mathbb{R}^n$ es un punto interior del conjunto C si existe $\delta > 0$ tal que $B(\bar{z}, \delta) \subseteq C$.
2. El conjunto de todos los puntos interiores de C es llamado interior de C y es denotado por $int(C)$.
- 3.- Diremos que $V \subseteq \mathbb{R}^n$ es un entorno o vecindad de \bar{z} si $\bar{z} \in int(V)$
- 4.-Diremos que B es un conjunto abierto si $B = int(B)$

Ejemplo 2.1.9.2. Sea $A = \{(x, y) \in \mathbb{R}^2: 1 < x < 2, 1 < y < 2\}$, es un conjunto abierto, y se ilustra en la figura 2.6

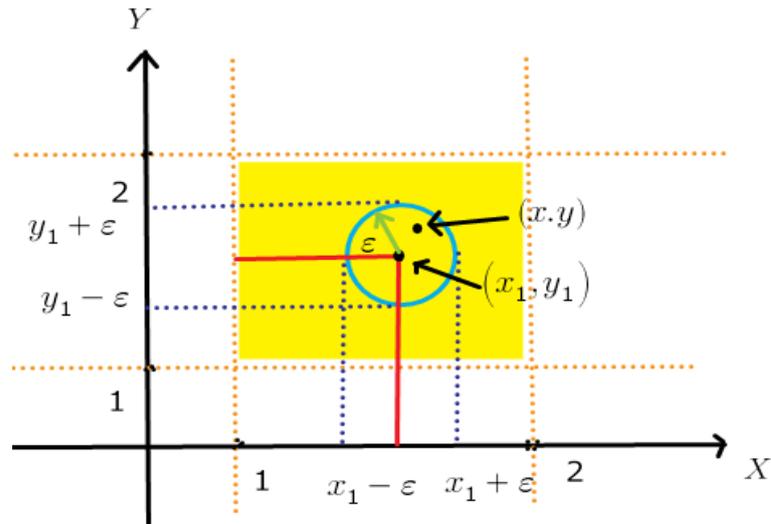


Figura 2.6. Conjunto abierto

En efecto del grafico sea $\bar{x} = (x_1, y_1) \in A$ entonces $1 < x_1 < 2$ y $1 < y_1 < 2$. Ahora definimos $\varepsilon = \min\{x_1 - 1, 2 - x_1, y_1 - 1, 2 - y_1\}$ por lo tanto si $(x, y) \in B(\bar{x}, \varepsilon)$ entonces $1 < x_1 - \varepsilon < x < x_1 + \varepsilon < 2$ y $1 < y_1 - \varepsilon < y < y_1 + \varepsilon < 2$, por lo tanto $(x, y) \in A$, implicando $B(\bar{x}, \varepsilon) \subset A$ y en consecuencia el conjunto A es abierto.

Teorema 2.1.9.3.-Sean \emptyset y \mathbb{R}^n donde se cumple que:

- 1.- \emptyset y \mathbb{R}^n son conjuntos abiertos
- 2.- Si B_1 y B_2 son conjuntos abiertos, entonces $B_1 \cap B_2$ es un conjunto abierto.
- 3.- Si $\{X_i\}_{i \in I}$ es una familia arbitraria de conjuntos abiertos, entonces $\bigcup_{i \in I} X_i$ es un conjunto abierto.

Demostración. Véase [7], página 30.

Definición 2.1.9.4. Sea $Y \subseteq \mathbb{R}^n$:

- 1.- Diremos que $\bar{y} \in \mathbb{R}^n$ es un punto adherente de Y si existe $\{y_n\} \subseteq \mathbb{R}^n$ tal que $\lim_{n \rightarrow \infty} y_n = \bar{y}$.
- 2.- El conjunto de todos los puntos adherentes de Y es llamado clausura o cerradura de Y y se denota por \bar{Y} .
- 3.- Diremos que Y es un conjunto cerrado si $Y = \bar{Y}$.

Teorema 2.1.9.5.-Se cumplen las siguientes propiedades:

- 1.- $Y \subseteq \bar{Y}$ para todo $Y \subseteq \mathbb{R}^n$.

2.- Si $Y \subseteq Z$, entonces $\bar{Y} \subseteq \bar{Z}$.

Demostración. Véase [7], página 35.

Teorema 2.1.9.6.- Sea $Y \subseteq \mathbb{R}^n$ y $\bar{y} \in \mathbb{R}^n$. Se dan las equivalentes:

1.- $\bar{y} \in \bar{Y}$.

2.- $B(\bar{y}, \delta) \cap Y \neq \emptyset, \forall \delta > 0$.

Demostración. Véase [7], página 36.

Teorema 2.1.9.7. Sea $Y \subseteq \mathbb{R}^n$, el conjunto Y es cerrado si y solo si su complemento $Y^c = \mathbb{R}^n - Y$ es abierto.

Demostración. Véase [7], página 37.

Teorema 2.1.9.8. Sean \emptyset y \mathbb{R}^n donde se cumple que:

1.- \emptyset y \mathbb{R}^n son conjuntos cerrados

2.- Si B_1 y B_2 son conjuntos cerrados, entonces $B_1 \cup B_2$ es un conjunto cerrado.

3.- Si $\{X_i\}_{i \in I}$ es una familia arbitraria de conjuntos cerrados, entonces $\bigcap_{i \in I} X_i$ es un conjunto cerrado.

2.1.10. Puntos de acumulación en \mathbb{R}^n

Definición 2.1. 10.1 Sea $Y \subseteq \mathbb{R}^n$:

1. Diremos que $\bar{y} \in \mathbb{R}^n$ es un punto de acumulación o punto límite del conjunto Y si y solo si $B(\bar{y}, \delta) \cap (Y - \{\bar{y}\}) \neq \emptyset$ para todo $\delta > 0$.

2. El conjunto de todos los puntos de acumulación de Y es llamado conjunto derivado de Y y es denotado por Y' .

3. Si $\bar{y} \in Y$ no es un punto de acumulación de Y diremos que \bar{y} es un punto aislado de Y , esto significa que existe $\delta > 0$, tal que \bar{y} es el único punto de Y en $B(\bar{y}, \delta)$.

4. Diremos que Y es un conjunto discreto si y solo si todos sus puntos son aislados.

Proposición 2.1. 10.2 Sea $X \subseteq \mathbb{R}^n$. $\bar{x} \in X'$, si y solo si $x \in \overline{X - \{\bar{x}\}}$

Demostración. Véase [7], página 40.

Teorema 2.1.10.3. Sea $Y \subseteq \mathbb{R}^n$ y $\bar{y} \in \mathbb{R}^n$. Se tiene las afirmaciones equivalentes.

1. $\bar{y} \in Y'$
2. Existe una sucesión $\{y_n\} \subseteq \overline{Y - \{y\}}$ tal que $\lim_{n \rightarrow +\infty} y_n = \bar{y}$

Demostración. Véase [7], página 41.

Teorema 2.1.10.4. (Bolzano Weierstras) Todo conjunto infinito y acotado $Y \subseteq \mathbb{R}^n$ tiene al menos un punto de acumulación.

Demostración. Véase [7], pág. 41.

2.1.11. Funciones continuas

Definición 2.1.11.1. Sea $X \subseteq \mathbb{R}^m$ y $f: X \rightarrow \mathbb{R}^n$.

1. Diremos que f es continua en $\bar{x}_0 \in X$, si para todo $\varepsilon > 0$, existe un $\delta = \delta(\varepsilon) > 0$, tal que si $\bar{x} \in X$ y $\|\bar{x} - \bar{x}_0\| < \delta$, entonces $\|f(\bar{x}) - f(\bar{x}_0)\| < \varepsilon$
2. Diremos que g es discontinua en $\bar{z}_0 \in Z$, sí g no es continua en \bar{z}_0 .
3. Diremos que g es discontinua en $X \subseteq Z$, si y solo si g no es continua en \bar{x}_0 , para $\bar{x}_0 \in X$.

Teorema 2.1.11.2. Sean $G: Y \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$ y $\bar{y} \in Y$. Son equivalentes:

1. G es continua en \bar{y} .
2. $\{y_n\} \subseteq Y$ tal que $\lim_{k \rightarrow +\infty} y_k = \bar{y}$, entonces $\lim_{k \rightarrow +\infty} G(x_k) = G(\bar{x})$

Demostración. Véase [8], página. 38.

Corolario 2.1.11.3. Sean $f, g: X \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$ y $\rho: X \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ funciones continuas en $\bar{x} \in X$ entonces $f \pm g$, ρf , $\|f\|$, $\langle f, g \rangle$, $d(f, g)$ y $\frac{1}{\rho}$ ($\rho(x) \neq 0$) son continuas en \bar{x} .

Demostración. Véase [8], página 39.

Teorema 2.1.11.4. Sea $g: Z \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^n$ una función vectorial, donde $g = (g_1, \dots, g_n)$ y $\bar{z} \in Z$. Son equivalentes.

1. g es continua en \bar{z} .
2. g_i son continuas en \bar{z} , $\forall i \in \{1, \dots, n\}$.

2.2. Marco conceptual

2.2.1. Teoría básica de Análisis convexo

En el análisis convexo se permite los elementos $-\infty, +\infty$ para extender la recta real, es decir que la línea recta extendida es denotado por

$[-\infty, +\infty] = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\} = \overline{\mathbb{R}}$. Las reglas aritéticas son extendidas a los elementos $\{-\infty\}$ y $\{+\infty\}$ de la manera usual, dejando las expresiones $+\infty + (-\infty)$ como indefinidas.

Además en el espacio euclidano \mathbb{R}^n la extensión estará dado por

$$\mathbb{R}^n \cup \{-\infty\} \cup \{+\infty\} = \overline{\mathbb{R}^n}$$

Considerando la función $\varphi: Y \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, donde puede ser prolongada a una función de valores extendidos $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por

$$g(x) = \begin{cases} \varphi(y), & \text{si } y \in Y \\ +\infty, & \text{si } y \notin Y \end{cases}$$

Definición 2.2.1.1. Sea la función $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ de valores extendidos.

1. El dominio de f denotado por $dom(f)$, es llamado dominio efectivo si es definido por

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

2. La función f es propia si $dom(f) \neq \emptyset$

3. Diremos que f es coerciva si para cada sucesión $\{x^k\}_{k \in \mathbb{N}}$ en \mathbb{R}^n se tiene que

$$\lim_{k \rightarrow +\infty} \sup f(x^k) = +\infty$$

o equivalentemente si

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

Ejemplo 2.2.1.2. Sea $z^k \in \mathbb{R}^n$ un punto fijo y la función $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, definida por $g(z) = \|z - z^k\|^2$, entonces g es coerciva.

En efecto, por la Proposición 2.1.6.4, tendremos que

$$\| \|z\| - \|z^k\| \|^2 \leq \|z - z^k\|^2 = g(z), \quad \forall z, z^k \in \mathbb{R}^n$$

Ahora dado que $\|z\| \rightarrow +\infty$ y $\|z^k\|$ representa un número fijo, luego al reemplazar dichos valores en la expresión anterior, obtendremos que $g(z) \rightarrow +\infty$.

Definición 2.2.1.3. Sea $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función de valores extendidos y $\rho \in \mathbb{R}$

1. El conjunto de nivel inferior de la función g está definido por

$$L_g(\rho) = \{x \in \mathbb{R}^n : g(x) \leq \rho\}$$

2. El conjunto de nivel superior de la función g está definido por

$$L_g(\rho) = \{x \in \mathbb{R}^n : g(x) \geq \rho\}$$

3. El epígrafo de la función g está definido por

$$epi(f) = \{(x, \rho) \in \mathbb{R}^n \times \mathbb{R} : g(x) \leq \rho\}$$

Ejemplo 2.2.1.4. Sea la función $f(x, y) = x^2 + y^2$.

1. El conjunto de nivel inferior estará dado por la siguiente ecuación y gráfica de la figura 7.

$$L_f(\alpha) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \alpha, \text{ con } \alpha \geq 0\}$$

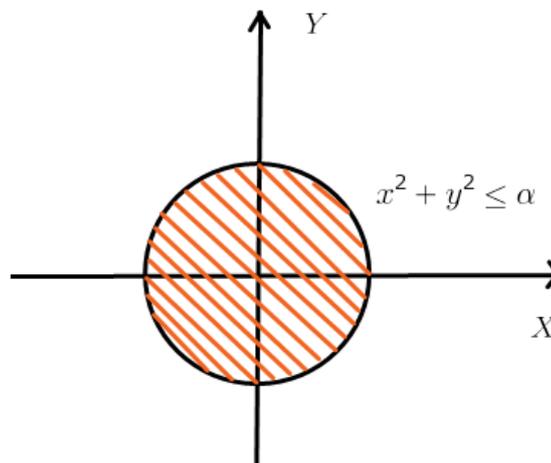


Figura 2.7. Conjunto de nivel inferior de f

2. El conjunto de nivel superior estará dado por la siguiente ecuación y gráfica de la figura 2.8

$$L_f(\alpha) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \geq \alpha, \text{ con } \alpha \geq 0\}$$

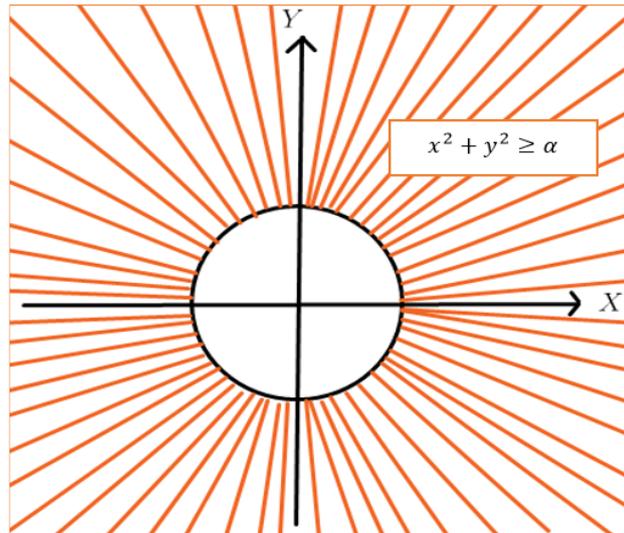


Figura 2.8. Conjunto de nivel superior

3. El epígrafo de f está definido por la siguiente ecuación y grafica

$$\text{epi}(f) = \{(x, y, \alpha) \in \mathbb{R}^2 \times \mathbb{R} : x^2 + y^2 \leq \alpha\}$$

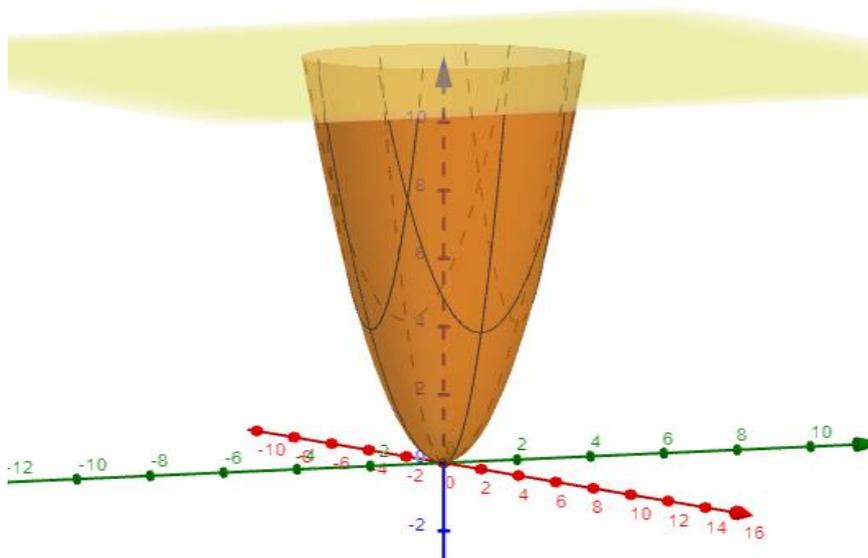


Figura 2.9. Epígrafo de la función f

Definición 2.2.1.5. Sea $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia. Diremos que g es localmente Lipschitz en $\bar{x} \in \text{dom}(g)$, si existe algún $\varepsilon_x > 0$ y $L_x > 0$ tal que:

$$|g(\bar{z}) - g(\bar{y})| \leq L_x \|\bar{z} - \bar{y}\| \quad \forall \bar{z}, \bar{y} \in B(\bar{x}, \varepsilon_x), \text{ donde}$$

$B(\bar{x}, \varepsilon_x) = \{\bar{y} \in \mathbb{R}^n: \|\bar{y} - \bar{x}\| < \varepsilon_x\}$ y $L_x > 0$ es la constante de Lipschitz.

Por otro lado, diremos que g es localmente Lipschitz en \mathbb{R}^n , si g es localmente Lipschitz en cada $\bar{x} \in \mathbb{R}^n$.

Definición 2.2.1.6. Sea la función $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ de valores extendidos, diremos que:

1. f es semicontinua inferior (sci) en un punto $\bar{x} \in \text{dom}(f) \subset \mathbb{R}^n$, si para cualquier sucesión $\{x^k\}_{k \in \mathbb{N}}$ en \mathbb{R}^n con x^k convergiendo a \bar{x} , se tiene que

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

Además, f es semicontinua inferior en \mathbb{R}^n , si es semicontinua inferior en todos los puntos de \mathbb{R}^n .

2. f es semicontinua superior (scs) en un punto $\bar{x} \in \text{dom}(f) \subset \mathbb{R}^n$, si para cualquier sucesión $\{x^k\}_{k \in \mathbb{N}}$ en \mathbb{R}^n con x^k convergiendo \bar{x} , se tiene que

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

Además, f es semicontinua superior en \mathbb{R}^n , si es semicontinua superior en todos los puntos de \mathbb{R}^n .

Ejemplo 2.2.1.7. Sea la función $g: [1,2] \rightarrow \mathbb{R}$ definida por

$$g(x) = \begin{cases} x^3 + 1, & \text{si } x \in]1,2] \\ 1, & \text{si } x = 1 \end{cases}$$

Se observa en gráfica de la figura 2.10 que la función g no es continua en $\bar{x} = 1$, sin embargo, la función es semicontinua inferior en dicho punto. En efecto, sea $\{x^k\} \subseteq [1,2]$ tal que $x^k \rightarrow 1$, como f es creciente para $n \in \mathbb{N}$ tenemos que $\inf_{k \geq n} g(x^k) = 2$, por lo tanto $\sup_{n \in \mathbb{N}} \inf_{k \geq n} g(x^k) = 2 > g(\bar{x}) = 1$.

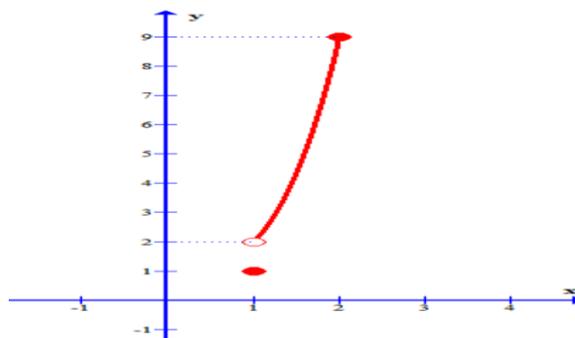


Figura 2.10. Función semicontinua inferior

Definición 2.2.1.8. Sea $C \subset \mathbb{R}^n, C \neq \emptyset$, la función indicatriz de C está definido por

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

Proposición 2.2.1.9. Sea $A \subset \mathbb{R}^n, A \neq \emptyset$. El conjunto A es cerrado si y solo si $\delta_A(\cdot)$ es semicontinua inferior. véase [22] página 17.

Teorema 2.2.1.10. $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función de valores extendidos. Entonces, f es semicontinua inferior si y solamente si el $\text{epi}(f)$ es un conjunto cerrado.

Demostración. Véase [13], página 16.

Teorema 2.2.1.11. Si $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, es una función coersiva para todo $\alpha \in \mathbb{R}$ entonces, el conjunto $L_f(\alpha)$ es limitado.

Demostración. Véase [13], página 17.

Teorema 2.2.1.12. Sea una función $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, coersiva y semicontinua inferiormente entonces el conjunto $L_f(\alpha)$ es compacto para todo $\alpha \in \mathbb{R}$.

Demostración. Véase [13], página 17.

Definición 2.2.1.13. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, y $\alpha \in \mathbb{R}$ talque $L_f(\alpha) \neq \emptyset$. Si f es coersiva, semicontinua inferior y limitadamente inferiormente. Entonces

$$\arg \min \{f(\bar{x}): \bar{x} \in L_f(\alpha)\} = \arg \min \{f(\bar{x}): \bar{x} \in \mathbb{R}^n\}$$

Demostración. Véase [13], página 17.

Proposición 2.2.1.14. Si la función $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es propia, y además semicontinua inferior y coerciva, entonces se tiene que el valor óptimo f^* es finito y el conjunto $\arg \min \{f(\bar{x}): \bar{x} \in \mathbb{R}^n\}$ es no vacío y compacto.

Demostración. Véase [65], Teorema 1.9.

Definición 2.2.1.15. Un conjunto $T \subset \mathbb{R}^n$ es llamado convexo si para cada $x, y \in T$, se tiene

$$x\lambda + (1 - \lambda)y \in T \quad \forall \lambda \in [0,1]$$

Ejemplo 2.2.1.16.-El conjunto a seguir

$$\mathcal{M} = \{(x, y) \in \mathbb{R}^2: x^2 + y^2 \leq \alpha, \alpha \in \mathbb{R}_0^+\}$$

es convexo

En efecto

Sean $P_0, P_1 \in \mathcal{M}, \lambda \in [0,1]$.

Probaremos que $P_1\lambda + (1 - \lambda)P_0 \in \mathcal{M} \quad \forall \lambda \in [0,1]$

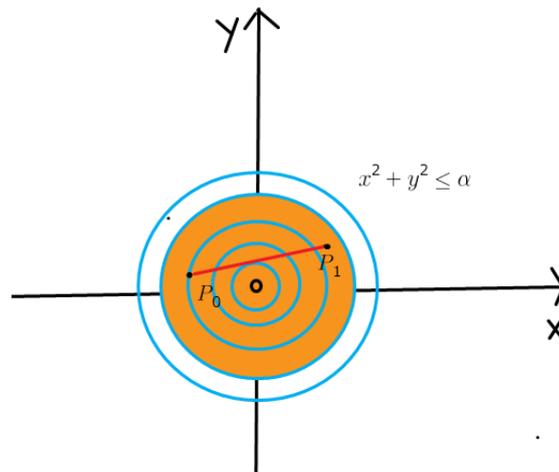


Figura 2.11. Conjunto convexo

En efecto considerando $P_0 = (x_1, y_1)$ y $P_1 = (x_2, y_2)$ y usando la definición

$$(x_1, y_1)\lambda + (1 - \lambda)(x_2, y_2) = (x_1\lambda + (1 - \lambda)x_2, y_1\lambda + (1 - \lambda)y_2)$$

$$(x_1\lambda + (1 - \lambda)x_2)^2 + (y_1\lambda + (1 - \lambda)y_2)^2 =$$

$$= \lambda^2 x_1^2 + 2\lambda(1 - \lambda)x_1 x_2 + (1 - \lambda)^2 x_2^2 + \lambda^2 y_1^2 + 2\lambda(1 - \lambda)y_1 y_2 + (1 - \lambda)^2 y_2^2$$

$$= \lambda^2 x_1^2 + \lambda^2 y_1^2 + (1 - \lambda)^2 x_2^2 + (1 - \lambda)^2 y_2^2 + 2\lambda(1 - \lambda)x_1 x_2 + 2\lambda(1 - \lambda)y_1 y_2$$

$$= \lambda^2(x_1^2 + y_1^2) + (1 - \lambda)^2(x_2^2 + y_2^2) + 2\lambda(1 - \lambda)x_1 x_2 + 2\lambda(1 - \lambda)y_1 y_2 \dots (\beta)$$

Por la teoría de números reales tenemos

$$(x_1 - x_2)^2 \geq 0 \rightarrow x_1^2 + x_2^2 \geq 2x_1 x_2 \rightarrow 2x_1 x_2 \leq x_1^2 + x_2^2$$

$$(y_1 - y_2)^2 \geq 0 \rightarrow y_1^2 + y_2^2 \geq 2y_1 y_2 \rightarrow 2y_1 y_2 \leq y_1^2 + y_2^2 \dots (\theta)$$

Ahora (θ) en (β)

$$\leq \lambda^2(x_1^2 + y_1^2) + (1 - \lambda)^2(x_2^2 + y_2^2) + \lambda(1 - \lambda)(x_1^2 + x_2^2) + \lambda(1 - \lambda)(y_1^2 + y_2^2)$$

$$\begin{aligned} &\leq \lambda^2(x_1^2 + y_1^2) + (1 - \lambda)^2(x_2^2 + y_2^2) + \lambda(1 - \lambda)(x_1^2 + y_1^2) + \lambda(1 - \lambda)(x_2^2 + y_2^2) \\ &\quad (x_1\lambda + (1 - \lambda)x_2)^2 + (y_1\lambda + (1 - \lambda)y_2)^2 \\ &\quad \leq \lambda^2\alpha + (1 - \lambda)^2\alpha + \lambda(1 - \lambda)\alpha + \lambda(1 - \lambda)\alpha \\ &\quad (x_1\lambda + (1 - \lambda)x_2)^2 + (y_1\lambda + (1 - \lambda)y_2)^2 \leq \alpha \end{aligned}$$

Por lo tanto $P_1\lambda + (1 - \lambda)P_0 \in \mathcal{M} \quad \forall \lambda \in [0,1]$, como se muestra en la gráfica de la figura 2.11.

Concluimos que el conjunto \mathcal{M} es convexo

Definición 2.2.1.17. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia. Entonces f es convexa en \mathbb{R}^n si para todo $x, y \in \mathbb{R}^n$ y para todo $\alpha \in [0,1]$ se cumple que:

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

Ejemplo 2.2.1.18. Demostrar que la siguiente función es convexa

$$f(x, y) = x^2 + y^2$$

Solución.

Por demostrar que si para todo $P_0, P_1 \in \mathbb{R}^2$ y para todo $\alpha \in [0,1]$ se cumple que:

$$f(P_0\alpha + (1 - \alpha)P_1) \leq \alpha f(P_0) + (1 - \alpha)f(P_1)$$

En efecto

Considerando $P_0 = (x_1, y_1)$, $P_1 = (x_2, y_2)$ y $(x_1, y_1)\alpha \in \mathbb{R}^2$, $(1 - \alpha)(x_2, y_2) \in \mathbb{R}^2$

Entonces $P_0\alpha + (1 - \alpha)P_1 \in \mathbb{R}^2$

Además $(x_1, y_1)\alpha + (1 - \alpha)(x_2, y_2) = (x_1\alpha + (1 - \alpha)x_2, y_1\alpha + (1 - \alpha)y_2) \dots (*)$

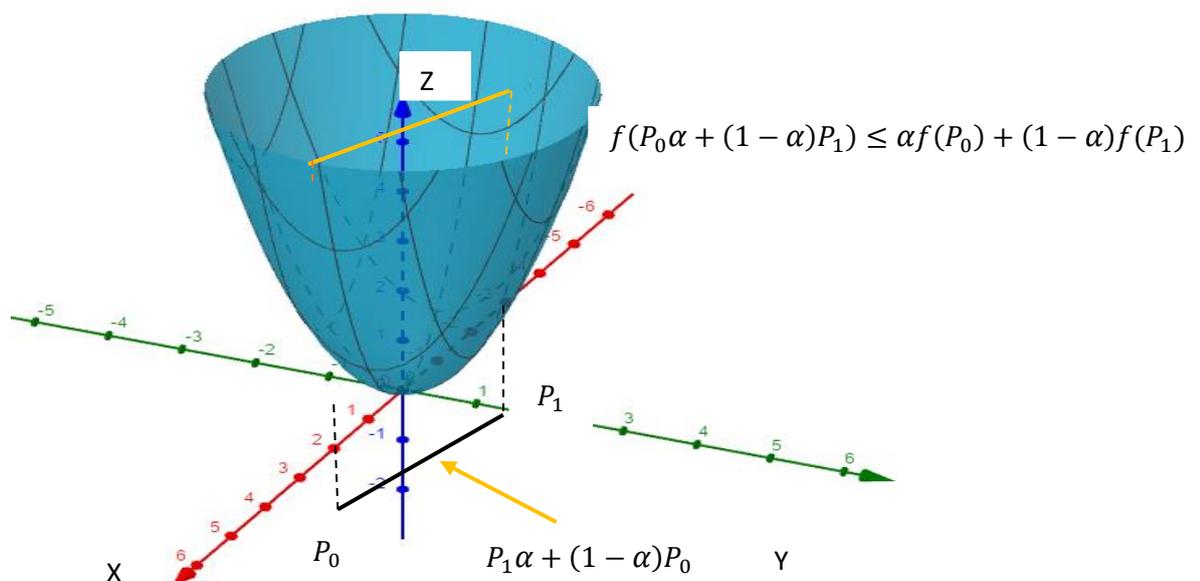


Figura 2.12. Función convexa

Reemplazando (*) en la función tenemos

$$\begin{aligned}
 f((x_1, y_1)\alpha + (1 - \alpha)(x_2, y_2)) &= f(x_1\alpha + (1 - \alpha)x_2, y_1\alpha + (1 - \alpha)y_2) \\
 f(x_1\alpha + (1 - \alpha)x_2, y_1\alpha + (1 - \alpha)y_2) &= (x_1\alpha + (1 - \alpha)x_2)^2 + (y_1\alpha + (1 - \alpha)y_2)^2 \\
 &= \alpha^2 x_1^2 + \alpha^2 y_1^2 + (1 - \alpha)^2 x_2^2 + (1 - \alpha)^2 y_2^2 + 2\alpha(1 - \alpha)x_1 x_2 + 2\alpha(1 - \alpha)y_1 y_2 \\
 &= \alpha^2(x_1^2 + y_1^2) + (1 - \alpha)^2(x_2^2 + y_2^2) + 2\alpha(1 - \alpha)x_1 x_2 + 2\alpha(1 - \alpha)y_1 y_2 \dots (**)
 \end{aligned}$$

Por la teoría de números reales tenemos

$$\begin{aligned}
 (x_1 - x_2)^2 \geq 0 &\rightarrow x_1^2 + x_2^2 \geq 2x_1 x_2 \rightarrow 2x_1 x_2 \leq x_1^2 + x_2^2 \\
 (y_1 - y_2)^2 \geq 0 &\rightarrow y_1^2 + y_2^2 \geq 2y_1 y_2 \rightarrow 2y_1 y_2 \leq y_1^2 + y_2^2 \dots (*)
 \end{aligned}$$

Ahora (*) en (**)

$$\begin{aligned}
 &f((x_1, y_1)\alpha + (1 - \alpha)(x_2, y_2)) \leq \\
 &\leq \alpha^2(x_1^2 + y_1^2) + (1 - \alpha)^2(x_2^2 + y_2^2) + \alpha(1 - \alpha)(x_1^2 + x_2^2) + \alpha(1 - \alpha)(y_1^2 + y_2^2) \\
 &\leq \alpha^2(x_1^2 + y_1^2) + (1 - \alpha)^2(x_2^2 + y_2^2) + (1 - \alpha)(x_1^2 + y_1^2) + \alpha(1 - \alpha)(x_2^2 + y_2^2) \\
 &\leq (\alpha^2 + \alpha - \alpha^2)(x_1^2 + y_1^2) + ((1 - \alpha)^2 + \alpha(1 - \alpha))(x_2^2 + y_2^2) \\
 &\leq (\alpha^2 + \alpha - \alpha^2)(x_1^2 + y_1^2) + (1 - \alpha)((1 - \alpha^2 + \alpha - \alpha^2) + \alpha)(x_2^2 + y_2^2) \\
 &\leq \alpha(x_1^2 + y_1^2) + (1 - \alpha)(x_2^2 + y_2^2)
 \end{aligned}$$

Por lo tanto

$$f(P_0\alpha + (1 - \alpha)P_1) \leq \alpha f(P_0) + (1 - \alpha)f(P_1)$$

Concluimos con la prueba

Definición 2.2.1.19. Sea la función propia $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. Entonces f es cuasi-convexa en \mathbb{R}^n si para todo $x, y \in \mathbb{R}^n$ y para todo $\lambda \in [0, 1]$ se cumple que:

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

Por otro lado, cuando la desigualdad es estricta es decir $f(x\gamma + (1 - \gamma)y) < \max\{f(x), f(y)\}$ y $\gamma \in [0, 1]$, para $x \neq y$ la función f es estrictamente cuasi-convexa.

Ejemplo 2.2.1.20

Sea la función definida mediante

$$f(x) = \begin{cases} (x-1)^2 & , 1 \leq x < 2 \\ 1 & , 2 \leq x < 4 \end{cases}$$

Determinar si la función f es cuasi-convexa.

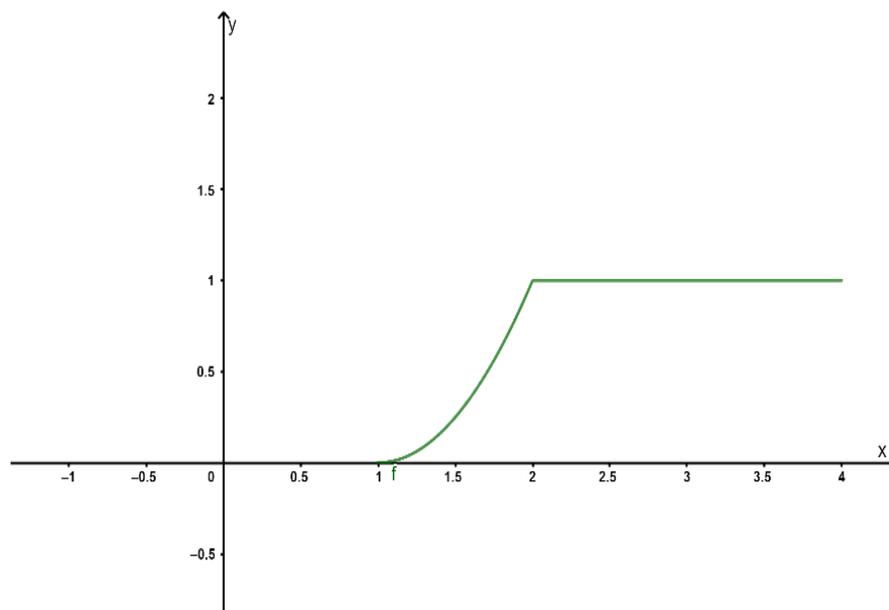


Figura 2.13. Función cuasi-convexa

Por probar que:

$$f(x\gamma + (1-\gamma)y) \leq \max\{(x-1)^2, 1\} \quad \forall x \in [1, 2), \forall y \in [2, 4) \text{ y } \gamma \in [0, 1]$$

En efecto dado que $x \leq y \Rightarrow x \leq x\gamma + (1-\gamma)y \leq y$ con $0 \leq \gamma \leq 1$ y como la función f es creciente (ver figura 2.13) entonces

$$(x-1)^2 = f(x) \leq f(x\gamma + (1-\gamma)y) \leq f(y) = 1 = \max\{(x-1)^2, 1\}$$

Por lo tanto, la función f es cuasi-convexa.

Teorema 2.2.1.21. Sea la función $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ propia. Entonces, f es cuasi-convexa si, y solo si, el conjunto $\{x \in \mathbb{R}^n: f(x) \leq C\}$ es convexo, para cada $C \in \mathbb{R}$.

Demostración. Ver ([6], Teorema 3.5.2)

Ejemplo 2.2.1.22.

Sea la función f definida mediante

$$f(x, y) = \sqrt[3]{2x^2 + 3y^2}$$

Se determinará si la función f es cuasi-convexa

En efecto por el teorema 3.4.10 demostraremos que el conjunto

$\{(x, y) \in \mathbb{R}^2: \sqrt[3]{2x^2 + 3y^2} \leq C\}$, es convexo o equivalentemente que el conjunto

$\{(x, y) \in \mathbb{R}^2: 2x^2 + 3y^2 \leq C^3\}$, sea convexo.

Por lo hecho en el ejemplo 2.2.1.16 el conjunto $\{(x, y) \in \mathbb{R}^2: 2x^2 + 3y^2 \leq C^3\}$ es convexo por lo tanto concluimos que la función f es cuasiconvexa.

Teorema 2.2.1.23. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase $C^1(\mathbb{R}^n)$. Entonces f es cuasi-convexa si, y solo si $\forall \bar{x}, \bar{y} \in \mathbb{R}^n$ si $f(\bar{x}) \leq f(\bar{y})$, se tiene que

$$\langle \nabla f(\bar{y}); \bar{x} - \bar{y} \rangle \leq 0$$

Demostración. Ver Mangazarian ([46], Teorema 4)

Ejemplo 2.2.1.24

Sea la función

$$f(x, y) = -e^{-x^2 - y^2}$$

definida en el conjunto

$$D = \{(x, y) \in \mathbb{R}^2: 1 \leq x \leq 4, 1 \leq y \leq 4\}$$

En efecto la función f es diferenciable en todo \mathbb{R}^2 , en particular en el conjunto D

Considerando

$$\bar{x} = (x, y), \bar{y} = (x_1, y_1), y \ x \leq x_1, y \leq y_1 \Rightarrow x - x_1 \leq 0, y - y_1 \leq 0 \dots (*)$$

$$\text{Además, si } 1 \leq x_1 \leq 4, 1 \leq y_1 \leq 4 \Rightarrow 2x_1 e^{-x_1^2 - y_1^2} \geq 0, 2y_1 e^{-x_1^2 - y_1^2} \geq 0 \dots (**)$$

Ahora de (*) y (**) tenemos

$$2x_1 e^{-x_1^2 - y_1^2} (x - x_1) \leq 0, \quad y \quad 2y_1 e^{-x_1^2 - y_1^2} (y - y_1) \leq 0$$

Ahora sumando se tiene $2x_1 e^{-x_1^2 - y_1^2} (x - x_1) + 2y_1 e^{-x_1^2 - y_1^2} (y - y_1) \leq 0$

$$\Rightarrow (2x_1 e^{-x_1^2 - y_1^2}, 2y_1 e^{-x_1^2 - y_1^2})(x - x_1, y - y_1) \leq 0$$

$$\langle \nabla f(\bar{y}); \bar{x} - \bar{y} \rangle \leq 0$$

Por lo tanto, la función f es cuasi-convexa.

Teorema 2.2.1.25. Sea $z = g(x, y)$ una función que tiene derivadas parciales continuas de primer y segundo orden, definida sobre un conjunto abierto y convexo C del plano. Entonces

$$\text{a) } g \text{ es convexa sí y solo si } \frac{\partial^2 g}{\partial x^2} \geq 0, \frac{\partial^2 g}{\partial y^2} \geq 0, \text{ y } \begin{vmatrix} \frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial y \partial x} \\ \frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial y^2} \end{vmatrix} \geq 0$$

$$\text{b) } g \text{ es cóncava sí y solo si } \frac{\partial^2 g}{\partial x^2} \leq 0, \frac{\partial^2 g}{\partial y^2} \leq 0, \text{ y } \begin{vmatrix} \frac{\partial^2 g}{\partial x^2} & \frac{\partial^2 g}{\partial y \partial x} \\ \frac{\partial^2 g}{\partial x \partial y} & \frac{\partial^2 g}{\partial y^2} \end{vmatrix} \geq 0$$

donde todas las desigualdades se deben verificar en todo el conjunto C .

Ejemplo 2.2.1.26

Sea la función f definida por

$$f(x, y) = 2x^2 + 3y^2 - 8x + 12y + 3xy$$

Se puede verificar que la función f es convexa. En efecto determinando las derivadas parciales de primer orden y segundo orden

$$\frac{\partial f}{\partial x} = 4x - 8 + 3y, \quad \frac{\partial f}{\partial y} = 6y - 12 + 3x$$

$$\frac{\partial^2 f}{\partial x^2} = 4, \quad \frac{\partial^2 f}{\partial y^2} = 6, \quad \frac{\partial^2 f}{\partial y \partial x} = 3, \quad \frac{\partial^2 f}{\partial x \partial y} = 3$$

Luego podemos observar que las derivadas parciales de primer orden y segundo son continuas en el plano cartesiano \mathbb{R}^2 y además

$$\frac{\partial^2 f}{\partial x^2} = 4 \geq 0, \quad \frac{\partial^2 f}{\partial y^2} = 6 \geq 0, \quad \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} = \begin{vmatrix} 4 & 3 \\ 3 & 6 \end{vmatrix} = 15 \geq 0$$

Por lo tanto, por el Teorema 2.2.1.25. se concluye que la función es convexa.

Definición 2.2.1.27- Sea la función $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, entonces F es \mathbb{R}_+^m -convexa si solamente si todas las funciones componentes de F , $F_i: \mathbb{R}^n \rightarrow \mathbb{R}$ son convexas.

Definición 2.2.1.28- Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función, entonces F es \mathbb{R}_+ -cuasiconvexa si todas las funciones componentes de F , $F_i: \mathbb{R}^n \rightarrow \mathbb{R}$ son cuasiconvexas.

Definición 2.2.1.29- Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función, entonces F es localmente Lipschitz en \mathbb{R}^n si todas las funciones componentes de F , $F_i: \mathbb{R}^n \rightarrow \mathbb{R}$ son localmente Lipschitz en \mathbb{R}^n .

Definición 2.2.1.30- Sea $f: D \subset \mathbb{R} \rightarrow \mathbb{R}$ y \bar{x} un punto acumulación en el conjunto D , entonces

i) El límite superior de la función f en \bar{x} está definido por

$$\limsup_{x \rightarrow \bar{x}} f(x) = \inf_{\delta > 0} \left\{ \sup_{x \in B(\bar{x}, \delta) \cap D} f(x) \right\}.$$

ii) El límite inferior de la función f en \bar{x} está definido por

$$\liminf_{x \rightarrow \bar{x}} f(x) = \sup_{\delta > 0} \left\{ \inf_{x \in B(\bar{x}, \delta) \cap D} f(x) \right\}.$$

Definición 2.2.1.31- Sea $f: C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función definida en el conjunto convexo C . La función f se denomina estrictamente cuasi-convexa si

$$f(\lambda x_1 + (1 - \lambda)x_2) < \max\{f(x_1), f(x_2)\}$$

para todo $x_1, x_2 \in C, x_1 \neq x_2$ y $\lambda \in (0,1)$.

Definición 2.2.1.32- Sea $f: C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función definida en el conjunto convexo C . La función se denomina semi-estricta cuasi-convexa si

$$f(x_1) < f(x_2) \Rightarrow f(\lambda x_1 + (1 - \lambda)x_2) < f(x_2)$$

donde $x_1, x_2 \in C$ y $0 < \lambda < 1$.

2.2.2. Distancia proximal

En esta sección consideraremos la distancia proximal y la distancia proximal inducida la cual fue propuesta inicialmente por Auslender e Teboulle en (1999), luego usada en diferentes investigaciones como por ejemplo Mallma Ramirez. L et al. [45] entre otros.

Definición 2.2.2.1 Una función $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ es llamada distancia proximal con respecto a un conjunto convexo abierto no vacío $X \subset \mathbb{R}^n$ si para cada $y \in X$, se cumple que:

(i) $d(\cdot, y)$ es propia, semicontinua inferior, estrictamente convexo y continuamente diferenciable en X .

(ii) $\text{dom } d(\cdot, y) \subset \bar{X}$ y $\text{dom } (\partial_1 d(\cdot, y)) = X$, donde $\partial_1 d(\cdot, y)$ denota el mapeo sub-diferencial clásico de la función $d(\cdot, y)$ con respecto a la primera variable.

(iii) $d(\cdot, y)$ es coercitivo sobre \mathbb{R}^n (es decir, $\lim_{\|u\| \rightarrow \infty} d(u, y) = +\infty$).

(iv) $d(y, y) = 0$

Denotamos por $D(X)$ la familia de funciones que satisfacen la definición anterior.

Definición 2.2.2.2 Dado $d \in D(X)$, una función $H: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ se llama la distancia proximal inducido a d si H es una función valor finito en $X \times X$ y para cada $a, b \in X$ tenemos:

(li) $H(a, a) = 0$.

(lii) $\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a)$, $\forall c \in X$ con $\gamma \in \langle 0, 1 \rangle$.

Denotemos $(d, H) \in \mathcal{F}(X)$, como la distancia proximal que satisface las condiciones de la definición 2.1.4.2 También denotamos $(d, H) \in \mathcal{F}(\bar{X})$ si existe H y cumple que

(liii) H es valor finito en $\bar{X} \times X$ y satisface (i) y (ii), para cada $c \in \bar{X}$.

(liv) Para cada $c \in \bar{X}$, $H(c, \cdot)$ tiene conjuntos de nivel acotados en X .

Finalmente, denotamos $(d, H) \in \mathcal{F}_+(\bar{X})$ si:

(lv) $(d, H) \in \mathcal{F}(\bar{X})$

(lvi) $\forall y \in \bar{X}$ y $\forall \{y^k\} \subset X$ acotado con $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$, luego

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

(lvii) $\forall y \in \bar{X}$ y $\forall \{y^k\} \subset X$ tal que $\lim_{k \rightarrow +\infty} y^k = y$, luego $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$.

Por ejemplo, considerando que la distancia proximal $d(x, y) = \frac{(x-y)^2}{2}$ cumple con la definición 2.2.2.1 es decir que $d \in D(\mathbb{R})$ (conjunto de distancias proximales) y además que la función distancia proximal inducida sea

exactamente dada por $H(x, y) = \frac{(x-y)^2}{2}$, cumple con las condiciones de la definición.

En efecto $\nabla_1 d(x, y) = x - y \Rightarrow \nabla_1 d(b, a) = (b - a)$ y

$$(z - x)^2 = ((z - y) + (y - x))^2 = (z - y)^2 + (y - x)^2 + 2(z - y)(y - x) \dots (\alpha)$$

Ahora reemplazando $x = a$, $y = b$, $z = c$ en (α) tenemos

$$2(c - b)(b - a) = (c - a)^2 - (c - b)^2 - (b - a)^2$$

$$(c - b)(b - a) = \frac{(c - a)^2}{2} - \frac{(c - b)^2}{2} - \frac{(b - a)^2}{2} \dots (\beta)$$

Por otro lado, si $a, b \in \mathbb{R}$ entonces $(b - a)^2 > 0$, y dado que $0 < \gamma \leq 1$

$$\Rightarrow 0 < \gamma(b - a)^2 \leq (b - a)^2$$

$$\Rightarrow -(b - a)^2 \leq -\gamma(b - a)^2 < 0 \dots (\theta)$$

Luego (θ) en (β) , $(c - b)(b - a) \leq \frac{(c-a)^2}{2} - \frac{(c-b)^2}{2} - \gamma \frac{(b-a)^2}{2}$

Por lo tanto, el resultado

$$\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a), \forall c \in \mathbb{R} \text{ con } \gamma \in \langle 0, 1 \rangle.$$

En consecuencia $(d, H) \in \mathcal{F}(\mathbb{R})$, es la distancia proximal inducida que satisface las condiciones de la definición 2.2.2.2. y además cumple con (liii) y (liv)

En efecto

1.-La función H tiene un valor finito en $\mathbb{R} \times \mathbb{R}$ si para cada $c \in \mathbb{R}$ satisface (li) y (lii), es decir

$$(c - b)(b - a) \leq \frac{(c - a)^2}{2} - \frac{(c - b)^2}{2} - \gamma \frac{(b - a)^2}{2}, \quad c \in \mathbb{R}$$

2.-Si $x = c \Rightarrow H(c, y) = \frac{(c-y)^2}{2}$ tiene conjuntos de nivel acotados en \mathbb{R}

Es decir, para cada $x = c$ tenemos que $\frac{(c-y)^2}{2} = k$, con $k \geq 0$.

Finalmente $(d, H) \in \mathcal{F}_+(\mathbb{R})$ si cumple (lv), (lvi) y (lvii).

En efecto

1.-Por (liii) y (liv) $(d, H) \in \mathcal{F}(\mathbb{R})$

2.- $\forall y \in \mathbb{R}$ y $\forall \{y^k\} \subset \mathbb{R}$ acotada con $\lim_{k \rightarrow +\infty} \frac{(y - y^k)^2}{2} = 0$

$$\Rightarrow \lim_{k \rightarrow +\infty} (y - y^k)^2 = 0, \text{ por lo tanto } \lim_{k \rightarrow +\infty} y^k = y.$$

3.- $\forall y \in \mathbb{R}$ y $\forall \{y^k\} \subset \mathbb{R}$ tal que $\lim_{k \rightarrow +\infty} y^k = y$

$$\Rightarrow \lim_{k \rightarrow +\infty} (y^k - y) = 0, \text{ por lo tanto } \Rightarrow \lim_{k \rightarrow +\infty} \frac{(y^k - y)^2}{2} = 0$$

Definición 2.2.2.3 (función φ -divergencia)

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

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

ii) φ es estrictamente convexa en su dominio.

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

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

Denotamos por Φ las clases de funciones que satisfacen (i) – (iv)

Por ejemplo, sea la función φ definida mediante

$$\varphi(t) = -\log t + t - 1, \text{ con } \text{dom} \varphi = \langle 0, +\infty \rangle$$

Verificaremos que $\varphi \in \Phi$.

En efecto $\varphi'(t) = 1 - \frac{1}{t}$, $\varphi''(t) = \frac{1}{t^2}$, con $t \in \langle 0, +\infty \rangle$, así se tiene que:

i.- La función φ es dos veces continuamente diferenciable en $\langle 0, +\infty \rangle$

ii.- $\varphi''(t) > 0$, $\forall t \in \text{dom} \varphi$, y además φ es estrictamente convexa

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

iv) $\varphi(1) = \varphi'(1) = 0$ y $\varphi''(1) > 0$

Por lo tanto $\varphi \in \Phi$.

Proposición 2.2.2.4. Sea la función φ que pertenece a la clase de funciones Φ , entonces

1.- $\varphi(t) \geq 0$ y $\varphi(t) = 0$ si y solo si $t = 1$

2.- $\varphi(t)$ es decreciente en $\langle 0, 1 \rangle$ y creciente en $\langle 1, +\infty \rangle$

3.- $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$

Demostración

1.- Sea $\varphi \in \Phi$ y por la definición de φ -divergencia parte iv) se tiene que $\varphi(1) = \varphi'(1) = 0$ y $\varphi''(1) > 0^+$, luego en $t = 1$ es un mínimo local de φ y además por la condición ii) el valor $\varphi(1) = 1$ es un mínimo global, en consecuencia $\varphi(t) \geq 0, \forall t \in \text{dom}\varphi$ y $\varphi(t) = 0$ si y solo si $t = 1$.

2.-Por el apartado anterior la función φ tiene un mínimo en $t = 1$, en consecuencia φ es decreciente en el intervalo en $\langle 0,1 \rangle$ y creciente en el intervalo en $\langle 1, +\infty \rangle$.

3.-Dado que la función φ es estrictamente convexa y el apartado 2) se tiene que

$\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$ con $\text{dom } \varphi \subseteq [0, +\infty)$.

Definición 2.2.2.5. Distancia proximal φ -divergencia (Auslender y Teboulle (1999))

Si $\varphi \in \Phi$, entonces la aplicación $d: \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$, definida por

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

es una distancia proximal φ -divergencia.

A continuación, presentaremos algunas distancias proximales φ -divergencia

Sea $\varphi: \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia, convexa y semicontinua inferior tal que $\text{dom } \varphi \subset \mathbb{R}_+$ $\text{dom } \varphi' = \mathbb{R}_{++}$. Supongamos que φ es $C^2(\mathbb{R})$, estrictamente convexa y no negativa en \mathbb{R}_{++} , con $\varphi(1) = \varphi'(1) = 0$. Denotemos por Φ_1 como una subclase de funciones Φ satisfaciendo la condición

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

y por Φ_2 como una subclase de funciones Φ satisfaciendo la condición

$$\varphi_2 = \left\{ \varphi \in \Phi: \varphi''(1) \left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t - 1) \quad \forall t > 0 \right\}$$

Además, dado que $\ln t \leq t - 1$ para cualquier $t > 0$ y $\varphi''(1) > 0$, se tiene

que $\Phi_1 \subseteq \Phi_2 \subseteq \Phi$.

Por ejemplo, la función definida mediante $\varphi_1(t) = t - \ln t - 1$, con $\text{dom } \varphi = \langle 0, +\infty \rangle$ es una subclase Φ_1 y la función $\varphi_2(t) = t \ln t - t + 1$, $\text{dom } \varphi = [0, +\infty)$ es una subclase Φ_2 . Véase [18], página 35.

Ejemplo 2.2.2.6.

Considerando que la función φ -divergencia sea definida mediante $\varphi(t) = t - \ln t - 1$, $\text{dom } \varphi = \langle 0, +\infty \rangle$ la cual es una subclase de funciones $\varphi \in \Phi_1$, luego haciendo $t = \frac{x_j}{y_j}$, la función distancia φ -divergencia queda definida por

$$\bar{d}_\varphi^1(x, y) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right)$$

Analizaremos si \bar{d}_φ^1 , es una distancia proximal en \mathbb{R}_{++}^n , en efecto

i.-a) Si $\bar{d}_\varphi^1(x, y) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right)$, por lo tanto $\text{dom}(\bar{d}_\varphi^2) \neq \emptyset$, con $C =$

\mathbb{R}_{++}^n , entonces $\bar{d}_\varphi^2(\cdot, y)$ es función propia.

b) Si $\varphi \left(\frac{x_j}{y_j} \right) = \frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1, j = 1, \dots, n$ y $f(y) = y_j, j = 1, \dots, n$ son funciones continuamente diferenciables, por lo tanto, \bar{d}_φ^1 es una función continua para $x, y \in \mathbb{R}_{++}^n$, entonces $\bar{d}_\varphi^1(\cdot, y)$ es semicontinua inferior.

c) Si $\varphi'' \left(\frac{x_j}{y_j} \right) = \frac{1}{x_j^2} > 0$, y $f(y) = y_j$, con $j = 1, \dots, n, x_j, y_j > 0$, entonces $\bar{d}_\varphi^1(\cdot, y)$ es estrictamente convexa en \mathbb{R}_{++}^n .

d) Si $\varphi \left(\frac{x_j}{y_j} \right) = \frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1, j = 1, \dots, n$, y $f(y) = y_j, j = 1, \dots, n$ son funciones continuamente diferenciables en \mathbb{R}_{++}^n , entonces $\bar{d}_\varphi^1(\cdot, y)$ es continuamente diferenciable en \mathbb{R}_{++}^n .

ii.- Si $\bar{d}_\varphi^1(x, y) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right)$, el dominio $\text{dom } \bar{d}_\varphi^2(\cdot, y) \subset \mathbb{R}_+^n$ y

$$\text{dom} \left(\nabla_1 \bar{d}_\varphi^1(\cdot, y) \right) = \mathbb{R}_{++}^n.$$

iii.-Si $\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right)$, se tiene que $\lim_{\|x\| \rightarrow \infty} y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right) = \lim_{\|x\| \rightarrow \infty} x_j - y_j \ln(x_j) + y_j \ln(y_j) - y_j = +\infty$, entonces $\bar{d}_\varphi^1(\cdot, y)$ es coerciva sobre \mathbb{R}^n

iv.-Si \bar{d}_φ^1 está definida mediante

$$\bar{d}_\varphi^1(x, y) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right), \text{ donde si hacemos que } x = y \text{ tenemos}$$

$$\bar{d}_\varphi^1(x, x) = \sum_{j=1}^n y_j \left(\frac{x_j}{y_j} - \ln \left(\frac{x_j}{y_j} \right) - 1 \right) = 0$$

Por lo tanto, concluimos que $\bar{d}_\varphi^1(\cdot, y)$ es una distancia proximal

Observación 2.2.2.7

Sea $\varphi \in \Phi_1$ definimos la distancia proximal φ -divergencia

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

donde para cualquier $\varphi \in \Phi$ el $\operatorname{argmin}\{\varphi(t): t \in \mathbb{R}\} = \{1\}$, dado que en $t = 1$ hay un mínimo global, además φ es coerciva, por lo tanto $d_\varphi \in D(C)$, con $C = \mathbb{R}_{++}^n$.

Proposición 2.2.2.8 Sea $C \subset \mathbb{R}^n$ un conjunto convexo y las funciones convexas $f_i: C \rightarrow \mathbb{R}$, para cada $i = 1, \dots, p$. Entonces para cualquier $t_i \in \mathbb{R}_+, i = 1, \dots, m$, la función $f: C \rightarrow \mathbb{R}$ definida por

$$f(x) = \sum_{i=1}^m t_i f_i(x)$$

es convexa en C .

Demostración vease [11] página 35.

Ejemplo 2.2.2.9 Distancia proximal homogénea de segundo orden

Sea $\varphi: \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia, convexa y semicontinua inferior tal que $\operatorname{dom} \varphi \subset \mathbb{R}_+$, $\operatorname{dom} \partial\varphi = \mathbb{R}_{++}$. Supongamos que φ es $C^2(\mathbb{R})$, estrictamente

convexa y no negativa en \mathbb{R}_{++} , con $\varphi(1) = \varphi'(1) = 0$, y $\varphi \in \Phi_2$, satisfaciendo la condición, esto es,

$$\varphi''(1) \left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t - 1) \quad \forall t > 0.$$

Se tiene la función $\bar{d}_\varphi^2: \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ definida mediante

$$\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \varphi\left(\frac{x_j}{y_j}\right) + \frac{\sigma}{2} \|x - y\|^2$$

donde σ es una constante positiva.

l) Analizaremos si \bar{d}_φ^2 , es una distancia proximal en \mathbb{R}_{++}^n , en efecto

i.-a) $\bar{d}_\varphi^2(\cdot, y)$ es propia dado que $\text{dom}(\bar{d}_\varphi^2) \neq \emptyset$, $C = \mathbb{R}_{++}^n$,

b) $\bar{d}_\varphi^2(\cdot, y)$ es semicontinua inferior dado que \bar{d}_φ^2 es una función continua para $x, y \in \mathbb{R}_{++}^n$,

c) $\bar{d}_\varphi^2(\cdot, y)$ es estrictamente convexa en \mathbb{R}_{++}^n dado que $\varphi''\left(\frac{x_j}{y_j}\right) > 0$ con $j = 1, \dots, n$, $y_j > 0$ y si $h(x, y) = \frac{\sigma}{2} \|x - y\|^2 = \frac{\sigma}{2} ((x_1 - y_1)^2 + \dots + (x_n - y_n)^2)$, $h''_1(x, y) > 0$,

d) $\bar{d}_\varphi^2(\cdot, y)$ es continuamente diferenciable en \mathbb{R}_{++}^n , dado que φ es $C^2(\mathbb{R})$ y $\frac{\sigma}{2} \|x - y\|^2$ continuamente diferenciable en \mathbb{R}_{++}^n

ii.- $\text{dom} \bar{d}_\varphi^2(\cdot, y) \subset \mathbb{R}_+^n$ y $\text{dom}(\partial_1 \bar{d}_\varphi^2(\cdot, y)) = \mathbb{R}_{++}^n$

iii.- $\bar{d}_\varphi^2(\cdot, y)$ es coerciva sobre \mathbb{R}^n dado que $\lim_{\|x\| \rightarrow \infty} \bar{d}_\varphi^2(\cdot, y) = +\infty$, es decir

$$\lim_{\|x\| \rightarrow \infty} \sum_{j=1}^n y_j^2 \varphi\left(\frac{y_j}{y_j}\right) + \frac{\sigma}{2} \|y - y\|^2 = \lim_{\|x\| \rightarrow \infty} \sum_{j=1}^n y_j^2 \varphi\left(\frac{y_j}{y_j}\right) + \lim_{\|x\| \rightarrow \infty} \frac{\sigma}{2} \|y - y\|^2 = +\infty$$

iv.- $\bar{d}_\varphi^2(y, y) = 0$, si $x = y$ es decir

$$\bar{d}_\varphi^2(y, y) = \sum_{j=1}^n y_j^2 \varphi\left(\frac{y_j}{y_j}\right) + \frac{\sigma}{2} \|y - y\|^2 = 0$$

Concluimos que $\bar{d}_\varphi^2(\cdot, y)$ es una distancia proximal

II) Probaremos que la función $H: \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ es una distancia proximal inducida a d , en efecto.

Si $\varphi''(1) \left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t - 1) \forall t > 0$ y haciendo $t = \frac{b_j}{a_j}$

$$\varphi'\left(\frac{b_j}{a_j}\right) \leq \varphi''(1)\left(\frac{b_j}{a_j} - 1\right) \rightarrow a_j \varphi'\left(\frac{b_j}{a_j}\right) \leq \varphi''(1)(b_j - a_j)$$

multiplicando por $c_j \geq 0$ en la desigualdad anterior se tiene

$$c_j a_j \varphi'\left(\frac{b_j}{a_j}\right) \leq \varphi''(1) c_j (b_j - a_j) \dots (\alpha)$$

Por otro lado $-\varphi'(t) \leq -\varphi''(1)\left(1 - \frac{1}{t}\right)$ y haciendo $t = \frac{b_j}{a_j}$

$$-\varphi'\left(\frac{b_j}{a_j}\right) \leq -\varphi''(1)\left(1 - \frac{a_j}{b_j}\right) \rightarrow -b_j \varphi'\left(\frac{b_j}{a_j}\right) \leq -\varphi''(1)(b_j - a_j)$$

multiplicando por $a_j \geq 0$ en la desigualdad anterior se tiene

$$-a_j b_j \varphi'\left(\frac{b_j}{a_j}\right) \leq -\varphi''(1) a_j (b_j - a_j) \dots (\beta)$$

Ahora sumando (α) y (β) se tiene

$$c_j a_j \varphi'\left(\frac{b_j}{a_j}\right) - a_j b_j \varphi'\left(\frac{b_j}{a_j}\right) \leq \varphi''(1) c_j (b_j - a_j) + \varphi''(1) a_j (a_j - b_j)$$

$$(c_j - b_j) a_j \varphi'\left(\frac{b_j}{a_j}\right) \leq \varphi''(1) c_j (b_j - a_j) + \varphi''(1) a_j (a_j - b_j)$$

Luego sumando $\sigma(c_j - b_j)(b_j - a_j)$ y aplicando sumatoria para $j = 1, \dots, n$

$$\begin{aligned} & \sum_{j=1}^n \left[(c_j - b_j) a_j \varphi'\left(\frac{b_j}{a_j}\right) + \sigma(c_j - b_j)(b_j - a_j) \right] \\ & \leq \varphi''(1) \sum_{j=1}^n [c_j (b_j - a_j) + a_j (a_j - b_j) + \sigma(c_j - b_j)(b_j - a_j)] \end{aligned}$$

Obtenemos

$$\langle c - b, \nabla_1 d(b, a) \rangle \leq \varphi''(1) \langle b - a, c - a \rangle + \sigma \langle b - a, c - b \rangle$$

con $\nabla_1 d(b, a) = (a_1 \varphi'\left(\frac{b_1}{a_1}\right) + \sigma(b_1 - a_1), \dots, a_n \varphi'\left(\frac{b_n}{a_n}\right) + \sigma(b_n - a_n))$.

Por otro lado, usando las identidades y reemplazando en la desigualdad anterior tendremos

$$\begin{aligned}\langle b - a, c - a \rangle &= \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 + \|b - a\|^2) \\ \langle b - a, c - b \rangle &= \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 - \|b - a\|^2) \\ \langle c - b, \nabla_1 d(b, a) \rangle &\leq \\ &\leq \varphi''(1) \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 + \|b - a\|^2) \\ &\quad + \sigma \frac{1}{2}(\|c - a\|^2 - \|c - b\|^2 - \|b - a\|^2)\end{aligned}$$

Finalmente, para cualquier $\varphi \in \Phi_2$ se cumple

$$\langle c - b, \nabla_1 d_\varphi(b, a) \rangle \leq \bar{\eta} \|c - a\|^2 - \eta \|c - \bar{b}\|^2 - \gamma \|b - a\|^2 \quad \forall a, b \in \mathbb{R}_{++}^n, \forall c \in \mathbb{R}_+^n$$

donde $\bar{\eta} = \frac{(\sigma + \varphi''(1))}{2}$.

Por lo tanto $H(x, y) = \bar{\eta} \|x - y\|^2$, $\gamma = \sigma - \bar{\eta}$

En consecuencia, la distancia proximal inducida H cumple con las condiciones

(li) $H(a, a) = 0$

(lii) $\langle c - b, \nabla_1 d_\varphi(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a) \quad \forall a, b \in \mathbb{R}_{++}^n, \forall c \in \mathbb{R}_+^n$

$$(d, H) \in \mathcal{F}(\mathbb{R}_{++}^n)$$

(liii) H es valor finito en $\mathbb{R}_+^n \times \mathbb{R}_{++}^n$ y satisface (li) y (lii), para cada $c \in \mathbb{R}_+^n$.

(liv) Para cada $c \in \mathbb{R}_+^n$, $H(c, \cdot)$ tiene conjuntos de nivel acotados en \mathbb{R}_{++}^n

es decir $H(c, y) = \bar{\eta} \|c - y\|^2 = \bar{\eta}((c_1 - y_1)^2 + \dots + (c_n - y_n)^2)$, $c \in \mathbb{R}_+^n$

$\bar{\eta}((c_1 - y_1)^2 + \dots + (c_n - y_n)^2) = k$, luego $(d, H) \in \mathcal{F}(\mathbb{R}_+^n)$.

Finalmente, $(d, H) \in \mathcal{F}_+(\mathbb{R}_+^n)$ si:

(lv) $(d, H) \in \mathcal{F}(\mathbb{R}_+^n)$

(lvi) $\forall y \in \mathbb{R}_+^n$ $\exists \{y^k\} \subset \mathbb{R}_{++}^n$ acotado con $\lim_{k \rightarrow +\infty} \bar{\eta} \|y - y^k\|^2 = 0 \rightarrow$

$\lim_{k \rightarrow +\infty} \|y - y^k\|^2 = 0$, luego $\lim_{k \rightarrow +\infty} y^k = y$.

(lvii) $\forall y \in \mathbb{R}_+^n$ y $\forall \{y^k\} \subset \mathbb{R}_+^n$ tal que $\lim_{k \rightarrow +\infty} y^k = y \rightarrow \lim_{k \rightarrow +\infty} \|y^k - y\| = 0 \rightarrow$
 $\lim_{k \rightarrow +\infty} \bar{\eta} \|y - y^k\|^2 = 0, \bar{\eta} > 0$, luego $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$.

III) Probaremos que $\|x - y\|^2 \leq \theta H(x, y)$ con θ positivo

En efecto por la **Proposición 2.2.2.4**, $\varphi\left(\frac{x_j}{y_j}\right) \geq 0 \forall x_j, y_j, j = 1, 2, \dots, n$,

$(y_1, y_2, \dots, y_n) \in \mathbb{R}_{++}^n$ y σ constante positiva

$$\frac{\sigma}{2} \|x - y\|^2 \leq \left(\sum_{j=1}^n y_j^2 \varphi\left(\frac{x_j}{y_j}\right) + \frac{\sigma}{2} \|x - y\|^2 \right)$$

$$\|x - y\|^2 \leq \frac{2}{\sigma} \left(\sum_{j=1}^n y_j^2 \varphi\left(\frac{x_j}{y_j}\right) + \frac{\sigma}{2} \|x - y\|^2 \right)$$

Ahora haciendo $\theta = \frac{2}{\sigma}$, tenemos

$$\|x - y\|^2 \leq \theta H(x, y)$$

Ejemplo 2.2.2.10

Considerando que la función φ -divergencia sea definida mediante $\varphi(t) = t - \ln t - 1$, $\text{dom } \varphi = \langle 0, +\infty \rangle$ que cumple la condición $\varphi''(1) \left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1) \ln t$, $\varphi''(1) = 1, \forall t > 0$, es decir $\varphi \in \Phi_1$, por lo tanto, haciendo $t = \frac{x_j}{y_j}$, la función distancia queda definida por

$$\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2$$

donde σ es una constante positiva.

I) Analizaremos si \bar{d}_φ^2 , es una distancia proximal en \mathbb{R}_{++}^n , en efecto

i.-a) Si $\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2$, por lo tanto $\text{dom}(\bar{d}_\varphi^2) \neq$

\emptyset , con $C = \mathbb{R}_{++}^n$, entonces $\bar{d}_\varphi^2(\cdot, y)$ es función propia.

b) Si $\varphi\left(\frac{x_j}{y_j}\right) = \frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1, j = 1, \dots, n$ y $f(y) = y_j^2, j = 1, \dots, n$ son funciones continuamente diferenciables, por lo tanto, \bar{d}_φ^2 es una función continua para $x, y \in \mathbb{R}_{++}^n$, entonces $\bar{d}_\varphi^2(\cdot, y)$ es semicontinua inferior.

c) Si $\varphi''\left(\frac{x_j}{y_j}\right) = \frac{1}{x_j^2} > 0$, $f''(y) = 2 > 0$, con $j = 1, \dots, n$, $x_j, y_j > 0$ y si $h(x, y) = \frac{\sigma}{2} \|x - y\|^2 = \frac{\sigma}{2} ((x_1 - y_1)^2 + \dots + (x_n - y_n)^2)$, con $h''_i(x, y) = 2 > 0$, entonces $\bar{d}_\varphi^2(\cdot, y)$ es estrictamente convexa en \mathbb{R}_{++}^n .

d) Si $\varphi\left(\frac{x_j}{y_j}\right) = \frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1$, $j = 1, \dots, n$, $f(y) = y_j^2$, $j = 1, \dots, n$, y $h(x, y) = \frac{\sigma}{2} \|x - y\|^2$ son funciones continuamente diferenciables en \mathbb{R}_{++}^n , entonces $\bar{d}_\varphi^2(\cdot, y)$ es continuamente diferenciable en \mathbb{R}_{++}^n .

ii.- Si $\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1\right) + \frac{\sigma}{2} \|x - y\|^2$, $\text{dom} \bar{d}_\varphi^2(\cdot, y) \subset \mathbb{R}_+^n$ y $\text{dom}(\nabla_1 \bar{d}_\varphi^2(\cdot, y)) = \mathbb{R}_{++}^n$.

iii.- Si $\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1\right) + \frac{\sigma}{2} \|x - y\|^2$, se tiene que

$\lim_{\|x\| \rightarrow \infty} \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1\right) + \frac{\sigma}{2} \|x - y\|^2 = +\infty$, entonces $\bar{d}_\varphi^2(\cdot, y)$ es coerciva sobre \mathbb{R}^n

iv.- Si \bar{d}_φ^2 está definida mediante

$\bar{d}_\varphi^2(x, y) = \sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1\right) + \frac{\sigma}{2} \|x - y\|^2$, donde si hacemos que $x = y$ tenemos

$$\bar{d}_\varphi^2(x, x) = \sum_{j=1}^n x_j^2 \left(\frac{x_j}{x_j} \ln\left(\frac{x_j}{x_j}\right) - \frac{x_j}{x_j} + 1\right) + \frac{\sigma}{2} \|x - x\|^2 = 0$$

Por lo tanto, concluimos que $\bar{d}_\varphi^2(\cdot, y)$ es una distancia proximal

II) Probaremos que la función $H: \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ es una distancia proximal inducida a d , en efecto.

Si $\varphi \in \Phi_1$ se cumple que $\varphi''(1) \left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1) \ln t \quad \forall t > 0$, $\varphi''(1) = 1$ y haciendo $t = \frac{b_j}{a_j}$, $j = 1, \dots, n$ en $\varphi'(t) \leq \varphi''(1) \ln t$, se tiene

$$\frac{b_j - a_j}{a_j b_j} \leq \ln\left(\frac{b_j}{a_j}\right)$$

multiplicando por $a_j \geq 0$, $j = 1, \dots, n$ y sumando el termino $\sigma(b_j - a_j)$ en la desigualdad anterior se tiene

$$\left[\frac{a_j(b_j - a_j)}{a_j b_j} + \sigma(b_j - a_j)\right] \leq \left[a_j \ln\left(\frac{b_j}{a_j}\right) + \sigma(b_j - a_j)\right]$$

Sea $c_j > b_j, j = 1, \dots, n$ donde $(c_j - b_j) > 0$, y multiplicando a la desigualdad anterior tenemos

$$(c_j - b_j) \left[\frac{a_j(b_j - a_j)}{a_j b_j} + \sigma(b_j - a_j) \right] \leq (c_j - b_j) \left[a_j \ln \left(\frac{b_j}{a_j} \right) + \sigma(b_j - a_j) \right]$$

Considerando la siguiente ecuación

$$\left[(c_j - b_j) a_j \ln \left(\frac{b_j}{a_j} \right) \right] = \left[a_j c_j \ln \left(\frac{c_j}{a_j} \right) + (a_j - c_j) \right] - \left[a_j c_j \ln \left(\frac{c_j}{b_j} \right) + (b_j - c_j) \right] - \left[a_j b_j \ln \left(\frac{b_j}{a_j} \right) + (a_j - b_j) \right]$$

y reemplazando en la expresión anterior y aplicando sumatoria para $j = 1, \dots, n$ se tiene

$$\begin{aligned} \sum_{j=1}^n (c_j - b_j) \left[\frac{a_j(b_j - a_j)}{a_j b_j} + \sigma(b_j - a_j) \right] \\ \leq \sum_{j=1}^n \left[\left(a_j c_j \ln \left(\frac{c_j}{a_j} \right) + (a_j - c_j) \right) - \left(a_j c_j \ln \left(\frac{c_j}{b_j} \right) + (b_j - c_j) \right) \right. \\ \left. - \left(a_j b_j \ln \left(\frac{b_j}{a_j} \right) + (a_j - b_j) \right) + \sigma(b_j - a_j)(c_j - b_j) \right] \end{aligned}$$

donde obtenemos

$$\langle c - b, \nabla_1 d(b, a) \rangle \leq \sum_{j=1}^n \left[\left(a_j c_j \ln \left(\frac{c_j}{a_j} \right) + (a_j - c_j) \right) - \left(a_j c_j \ln \left(\frac{c_j}{b_j} \right) + (b_j - c_j) \right) - \left(a_j b_j \ln \left(\frac{b_j}{a_j} \right) + (a_j - b_j) \right) \right] + \sigma \langle b - a, c - b \rangle$$

$$\text{con } \nabla_1 \bar{d}_\varphi^2(b, a) = \left(\frac{a_1(b_1 - a_1)}{a_1 b_1} + \sigma(b_1 - a_1), \dots, \frac{a_n(b_n - a_n)}{a_n b_n} + \sigma(b_n - a_n) \right)$$

Por otro lado, usando la identidad y reemplazando en la desigualdad anterior tendremos

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

$$\begin{aligned} \langle c - b, \nabla_1 \bar{d}_\varphi^2(b, a) \rangle \leq \\ \leq \sum_{j=1}^n \left[\left(a_j c_j \ln \left(\frac{c_j}{a_j} \right) + (a_j - c_j) \right) - \left(a_j c_j \ln \left(\frac{c_j}{b_j} \right) + (b_j - c_j) \right) \right. \\ \left. - \left(a_j b_j \ln \left(\frac{b_j}{a_j} \right) + (a_j - b_j) \right) \right] + \frac{\sigma}{2} (\|c - a\|^2 - \|c - b\|^2 - \|b - a\|^2) \end{aligned}$$

Finalmente se cumple

$$\begin{aligned} \langle c - b, \nabla_1 \bar{d}_\varphi^2(b, a) \rangle &\leq \sum_{j=1}^n \left(a_j c_j \ln \left(\frac{c_j}{a_j} \right) + (a_j - c_j) \right) + \frac{\sigma}{2} \\ &\quad \| c - a \|^2 \\ &\quad - \sum_{j=1}^n \left(a_j c_j \ln \left(\frac{c_j}{b_j} \right) + (b_j - c_j) \right) - \frac{\sigma}{2} \| c - b \|^2 - \gamma \sum_{j=1}^n \left(a_j b_j \ln \left(\frac{b_j}{a_j} \right) \right. \\ &\quad \left. + (a_j - b) \right) - \frac{\sigma}{2} \| b - a \|^2 \quad \forall a, b \in \mathbb{R}_{++}^n, \forall c \in \mathbb{R}_+^n \end{aligned}$$

Por lo tanto $H(x, y) = \sum_{j=1}^n \left(x_j y_j \ln \left(\frac{x_j}{y_j} \right) + (y_j - x_j) \right) + \frac{\sigma}{2} \| x - y \|^2$, $\gamma = 1$

En consecuencia, la distancia proximal inducida H cumple con las condiciones

(li) $H(a, a) = 0$

(lii) $\langle c - b, \nabla_1 d_\varphi(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a) \quad \forall a, b \in \mathbb{R}_{++}^n, \forall c \in \mathbb{R}_+^n$

$$(d, H) \in \mathcal{F}(\mathbb{R}_{++}^n)$$

(liii) H es valor finito en $\mathbb{R}_+^n \times \mathbb{R}_{++}^n$ y satisface (li) y (lii), para cada $c \in \mathbb{R}_+^n$.

(liv) Para cada $c \in \mathbb{R}_+^n$, $H(c, \cdot)$ tiene conjuntos de nivel acotados en \mathbb{R}_{++}^n

es decir $H(c, y) = \sum_{j=1}^n \left(c_j y_j \ln \left(\frac{c_j}{y_j} \right) + (y_j - c_j) \right) + \frac{\sigma}{2} ((c_1 - y_1)^2 + \dots + (c_n - y_n)^2)$, $c \in \mathbb{R}_+^n$

$\rightarrow \sum_{j=1}^n \left(c_j y_j \ln \left(\frac{c_j}{y_j} \right) + (y_j - c_j) \right) + \frac{\sigma}{2} ((c_1 - y_1)^2 + \dots + (c_n - y_n)^2) = k$ luego

$(d, H) \in \mathcal{F}(\mathbb{R}_+^n)$

Finalmente, $(d, H) \in \mathcal{F}_+(\mathbb{R}_+^n)$ si:

(lv) $(d, H) \in \mathcal{F}(\mathbb{R}_+^n)$

(lvi) $\forall y \in \mathbb{R}_+^n$ $y \forall \{y^k\} \subset \mathbb{R}_{++}^n$ acotado con $\lim_{k \rightarrow +\infty} \sum_{j=1}^n \left(y y_j^k \ln \left(\frac{y}{y_j^k} \right) + \right.$

$\left. (y_j^k - y) \right) + \frac{\sigma}{2} \| y - y^k \|^2 = \lim_{k \rightarrow +\infty} \sum_{j=1}^n \left(y y_j^k \ln \left(\frac{y}{y_j^k} \right) + (y_j^k - y) \right) +$

$\lim_{k \rightarrow +\infty} \frac{\sigma}{2} \| y - y^k \|^2 = 0$, luego $\lim_{k \rightarrow +\infty} y^k = y$.

(lvii) $\forall y \in \mathbb{R}_+^n$ $y \forall \{y^k\} \subset \mathbb{R}_{++}^n$ tal que $\lim_{k \rightarrow +\infty} y^k = y$, luego $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$.

III) Probaremos que $\|x - y\|^2 \leq \theta H(x, y)$ con θ positivo

En efecto, $\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1 \geq 0$, $y_j \geq 0 \forall x_j, y_j j = 1, 2, \dots, n$, $(y_1, y_2, \dots, y_n) \in \mathbb{R}_{++}^n$ y σ constante positiva

$$\frac{\sigma}{2} \|x - y\|^2 \leq \left(\sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2 \right)$$

$$\|x - y\|^2 \leq \frac{2}{\sigma} \left(\sum_{j=1}^n y_j^2 \left(\frac{x_j}{y_j} - \ln\left(\frac{x_j}{y_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2 \right)$$

Ahora haciendo $\theta = \frac{2}{\sigma}$, tenemos

$$\|x - y\|^2 \leq \theta H(x, y)$$

2.2.3. Subdiferenciales

En esta sección se presenta las definiciones y ejemplos de subdiferenciales que se usan en la presente investigación.

2.2.4. Subdiferencial de Fréchet y en el límite

Definición 2.2.4.1. - Sea la función $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es llamada función propia si

a) $\text{dom} g \neq \emptyset$

b) $\forall x \in \text{dom} g$, tenemos que $g(x) > -\infty$

Definición 2.2.4.2. - Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia.

1. Para $x \in \text{dom}(f)$, donde $f(x)$ es finito, el conjunto de subgradientes regulares (también denominado **subdiferencial de Fréchet**) de f en x , denotado por $\hat{\partial}f(x)$ está formado por el conjunto de vectores $d \in \mathbb{R}^n$, tales que

$$f(y) \geq f(x) + \langle d, y - x \rangle + \theta(\|y - x\|), \text{ donde } \lim_{\substack{y \rightarrow x \\ y \neq x}} \frac{\theta(\|y - x\|)}{\|y - x\|} = 0$$

O equivalentemente

$$\hat{\partial}f(x) = \left\{ d \in \mathbb{R}^n : \liminf_{\substack{y \neq x, y \rightarrow x}} \frac{f(y) - f(x) - \langle d, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

Si $x \notin \text{dom}(f)$, entonces $\hat{\partial}f(x) = \emptyset$

2. El conjunto de subgradientes generales (también denominados **subdiferencial en el límite**) de f en x , denotado por $\partial^{Lim}f(x)$ está definido por

$$\partial^{Lim}f(x) = \{d \in \mathbb{R}^n : \exists x^t \rightarrow x, f(x^t) \rightarrow f(x), d^t \in \hat{\partial}f(x^t) \text{ y } d^t \rightarrow d\}$$

Ejemplo 2.2.4.3 Sea la función

$$f(x) = |x|$$

Se analizará el subdiferencial de f en todo su dominio

Por definición se sabe que:

$$|x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$

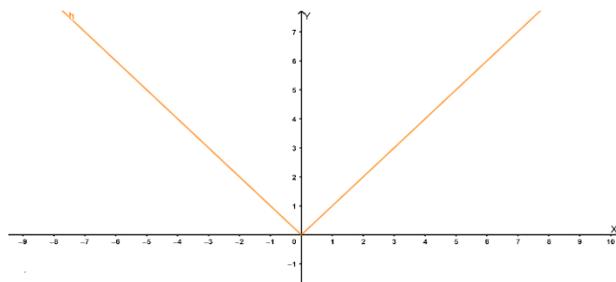


Figura 2.14. Función convexa no diferenciable en todo \mathbb{R}

a) Si $x = 0$, tenemos

$$\hat{\partial}f(0) = \liminf_{\substack{y \rightarrow 0 \\ y \neq 0}} \frac{f(y) - f(0) - \langle d, y - 0 \rangle}{|y - 0|} \geq 0$$

$$= \liminf_{\substack{y \rightarrow 0 \\ y \neq 0}} \frac{f(y) - f(0) - \langle d, y - 0 \rangle}{|y - 0|}$$

$$= \lim_{\substack{y \rightarrow 0 \\ y \neq 0}} \inf \frac{|y| - 0 - \langle d, y \rangle}{|y|}$$

$$= \lim_{\substack{y \rightarrow 0 \\ y \neq 0}} \inf \frac{|y| - d \cdot y}{|y|}$$

Ahora considerando el límite cuando $y \rightarrow 0$ por la derecha e izquierda

$$\hat{\partial} f_-(0) = \lim_{\substack{y \rightarrow 0^- \\ y \neq 0}} \inf \frac{-y - d \cdot y}{-y}$$

$$\hat{\partial} f_-(0) = \lim_{\substack{y \rightarrow 0^- \\ y \neq 0}} \inf (1 + d) \geq 0 \Rightarrow (1 + d) \geq 0 \Rightarrow d \geq -1 \dots (*)$$

$$\hat{\partial} f_+(0) = \lim_{\substack{y \rightarrow 0^+ \\ y \neq 0}} \inf \frac{y(1 - d)}{y}$$

$$\hat{\partial} f_+(0) = \lim_{\substack{y \rightarrow 0^+ \\ y \neq 0}} \inf (1 - d) \geq 0 \Rightarrow (1 - d) \geq 0 \Rightarrow d \leq 1 \dots (**)$$

Ahora de (*) y (**) se concluye que el subdiferencial de f en $x = 0$ es

$$d \in [-1, 1]$$

b) Si $x > 0$, tenemos

$$\hat{\partial} f(x) = \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{f(y) - f(x) - \langle d, y - x \rangle}{|y - x|} \geq 0$$

$$= \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{y - x - \langle d, y - x \rangle}{|y - x|}$$

$$\hat{\partial} f(x) = \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{y - x - d(y - x)}{|y - x|}$$

Por definición se sabe que:

$$|y - x| = \begin{cases} y - x, & y \geq x \\ -y + x, & y < x \end{cases}$$

Entonces

$$\hat{\partial} f(x) = \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{(y - x) - d(y - x)}{-(y - x)} \Rightarrow (d - 1) \geq 0 \Rightarrow d \geq 1 \dots (*)$$

$$\hat{\partial} f(x) = \lim_{\substack{y \rightarrow x^+ \\ y \neq x}} \inf \frac{(y - x) - d(y - x)}{(y - x)} \Rightarrow (1 - d) \geq 0 \Rightarrow d \leq 1 \dots (**)$$

Ahora de (*) y (**) se concluye que el subdiferencial de f para $x > 0$ es

$$\hat{\partial} f(x) = 1, \forall x > 0$$

De manera análoga para $x < 0$

$$\hat{\partial}f(x) = -1, \quad \forall x < 0$$

Finalmente, el sub-diferencial de la función f en todo su dominio estará dado por:

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

Ejemplo 2.2.4.4

Sea la función definida mediante

$$f(x) = \begin{cases} -(x-1)^2 + 1, & x \geq 1 \\ x, & x < 1 \end{cases}$$

Donde el gráfico de f se muestra en la siguiente figura.

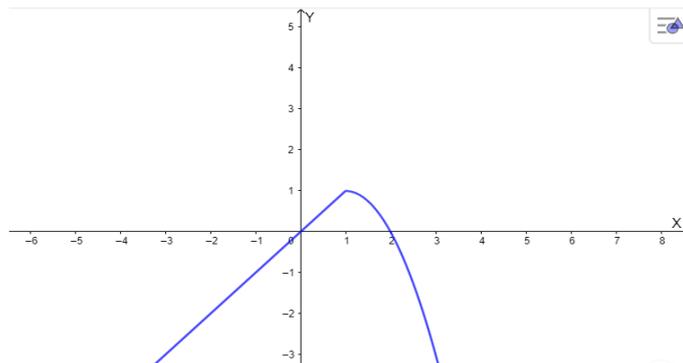


Figura 2.15. Función no convexa no diferenciable en todo \mathbb{R}

A continuación, analizaremos el subdiferencial de f en todo su dominio

i) Para $x < 1$ tenemos

$$\begin{aligned} \hat{\partial}f(x) &= \liminf_{\substack{y \rightarrow x \\ y \neq x}} \frac{f(y) - f(x) - \langle d, y - x \rangle}{|y - x|} \geq 0 \\ &= \liminf_{\substack{y \rightarrow x \\ y \neq x}} \frac{y - x - \langle d, y - x \rangle}{|y - x|} \\ \hat{\partial}f(x) &= \liminf_{\substack{y \rightarrow x \\ y \neq x}} \frac{y - x - d(y - x)}{|y - x|} \end{aligned}$$

Por definición se sabe que:

$$|y - x| = \begin{cases} y - x, & y \geq x \\ -y + x, & y < x \end{cases}$$

Entonces

$$\hat{\partial}f(x) = \liminf_{\substack{y \rightarrow x \\ y \neq x}} \frac{(y - x) - d(y - x)}{-(y - x)} \Rightarrow (d - 1) \geq 0 \Rightarrow d \geq 1 \dots (*)$$

$$\hat{\partial}f(x) = \liminf_{\substack{y \rightarrow x^+ \\ y \neq x}} \frac{(y - x) - d(y - x)}{(y - x)} \Rightarrow (1 - d) \geq 0 \Rightarrow d \leq 1 \dots (**)$$

Ahora de (*) y (**) se concluye que el subdiferencial de f para $x < 1$ es

$$\hat{\partial}f(x) = 1, \quad \forall x < 1$$

ii) Para $x = 1$ tenemos

$$\hat{\partial}f(1) = \liminf_{\substack{y \rightarrow 1 \\ y \neq 1}} \frac{f(y) - f(1) - \langle d, y - 1 \rangle}{|y - 1|} \geq 0$$

$$= \liminf_{\substack{y \rightarrow 1 \\ y \neq 1}} \frac{f(y) - f(1) - \langle d, y - 1 \rangle}{|y - 1|}$$

Ahora considerando el límite cuando $y \rightarrow 1$ por derecha y la izquierda

$$\hat{\partial}f_+(1) = \liminf_{\substack{y \rightarrow 1^+ \\ y \neq 1}} \frac{-(y - 1)^2 - d(y - 1)}{(y - 1)}$$

$$\hat{\partial}f_+(1) = \liminf_{\substack{y \rightarrow 1^+ \\ y \neq 1}} \frac{-(y - 1)(y - 1) - d(y - 1)}{(y - 1)}$$

$$\hat{\partial}f_+(1) = \liminf_{\substack{y \rightarrow 1^+ \\ y \neq 1}} (1 - y - d) \geq 0$$

$$\hat{\partial}f_+(1) = \sup_{\delta > 0} \left\{ \inf_{y \in B(1, \delta) \cap \{1\}} (1 - y - d) \right\} \geq 0$$

$$\hat{\partial}f_+(1) = \sup_{\delta > 0} \{-d\} \geq 0, \quad \Rightarrow d \leq 0 \dots (*)$$

$$\hat{\partial}f_-(1) = \liminf_{\substack{y \rightarrow 1^- \\ y \neq 1}} \frac{y - 1 - d(y - 1)}{-(y - 1)}$$

$$\hat{\partial}f_-(1) = \liminf_{\substack{y \rightarrow 1^- \\ y \neq 1}} \frac{(1 - d)}{-1}$$

$$\hat{\partial}f_-(1) = \lim_{\substack{y \rightarrow 1^- \\ y \neq 1}} \inf (d - 1) \geq 0 \Rightarrow d \geq 1 \dots (*)$$

Ahora de (*) y (**) se concluye que el subdiferencial de f en $x = 1$ es

$$\hat{\partial}f(1) = \emptyset$$

iii) Para $x > 1$ tenemos

$$\begin{aligned} \hat{\partial}f(x) &= \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{f(y) - f(x) - \langle d, y - x \rangle}{|y - x|} \geq 0 \\ &= \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{-(y - 1)^2 + 1 - (-(x - 1)^2 + 1) - \langle d, y - x \rangle}{|y - x|} \\ \hat{\partial}f(x) &= \lim_{\substack{y \rightarrow x \\ y \neq x}} \inf \frac{-(y - 1)^2 + (x - 1)^2 - \langle d, y - x \rangle}{|y - x|} \end{aligned}$$

Por definición se sabe que:

$$|y - x| = \begin{cases} y - x, & y \geq x \\ -y + x, & y < x \end{cases}$$

Entonces

$$\begin{aligned} \hat{\partial}f(x) &= \lim_{\substack{y \rightarrow x^- \\ y \neq x}} \inf \frac{-(y - 1)^2 + (x - 1)^2 - \langle d, y - x \rangle}{x - y} \\ \hat{\partial}f(x) &= \lim_{\substack{y \rightarrow x^- \\ y \neq x}} \inf \frac{(x - y)(x + y) - 2(x - y) + d(x - y)}{(x - y)} \\ \hat{\partial}f(x) &= \lim_{\substack{y \rightarrow x^- \\ y \neq x}} \inf (x + y - 2 + d) \geq 0 \\ \hat{\partial}f(x) &= \sup_{\delta > 0} \left\{ \inf_{y \in B(x^-, \delta) \cap \langle 1, +\infty \rangle} (x + y - 2 + d) \right\} \\ \hat{\partial}f(x) &= \sup_{\delta > 0} \{-\delta + 2(x - 1) + d\} \geq 0 \\ \hat{\partial}f(x) &= \sup_{\delta > 0} \{-\delta + 2(x - 1) + d\} \geq 0 \Rightarrow d \geq -2(x - 1) \dots (*) \end{aligned}$$

$$\hat{\partial}f(x) = \lim_{\substack{y \rightarrow x^+ \\ y \neq x}} \inf \frac{-(y - 1)^2 + (x - 1)^2 - \langle d, y - x \rangle}{x - y}$$

$$\hat{\partial}f(x) = \lim_{\substack{y \rightarrow x^+ \\ y \neq x}} \inf \frac{(x - y)(x + y) - 2(x - y) + d(x - y)}{-(x - y)}$$

$$\hat{\partial}f(x) = \lim_{\substack{y \rightarrow x^+ \\ y \neq x}} \inf(-x - y + 2 - d) \geq 0$$

$$\hat{\partial}f(x) = \sup_{\delta > 0} \left\{ \inf(-x - y + 2 - d) \mid y \in B(x^+, \delta) \cap \langle 1, +\infty \rangle \right\}$$

$$\hat{\partial}f(x) = \sup_{\delta > 0} \{-\delta - 2(x - 1) - d\} \geq 0$$

$$\hat{\partial}f(x) = \sup_{\delta > 0} \{-\delta - 2(x - 1) - d\} \geq 0 \Rightarrow d \leq -2(x - 1) \dots (**)$$

Ahora de (*) y (**) se concluye que el subdiferencial de f para $x < 1$ es

$$\hat{\partial}f(x) = -2(x - 1), \quad \forall x < 1$$

Finalmente, el sub-diferencial de la función f en todo su dominio estará dado por:

$$\hat{\partial}f(x) = \begin{cases} \emptyset & ; x = 1 \\ -2(x - 1); & x > 1 \\ 1; & x < 1 \end{cases}$$

Ejemplo 2.2.4.5

Sea la función f definida mediante

$$f(x, y) = |x - 1|$$

Cuyo grafico de f se muestra en la siguiente figura

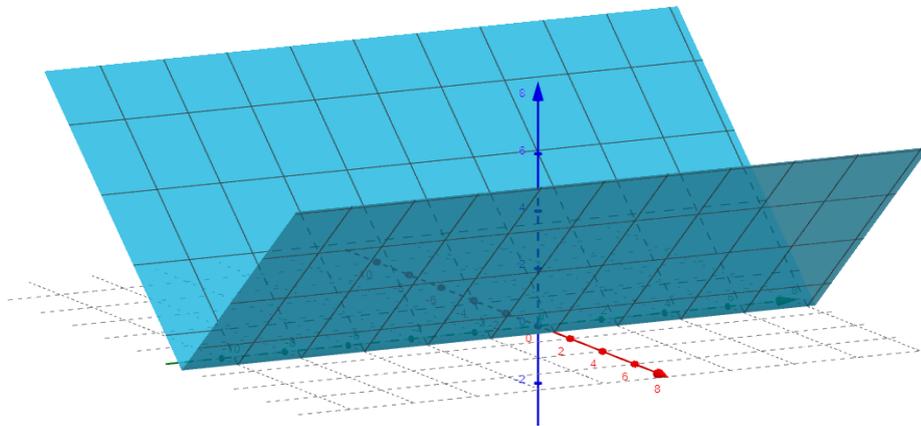


Figura 2.16. Función convexa de dos variables no diferenciable en todo \mathbb{R}^2

A continuación, analizaremos el subdiferencial de f en el punto (1,0)

$$\hat{\partial}f(1,0) = \left\{ d \in \mathbb{R}^2 : \lim_{\substack{\bar{y} \rightarrow \bar{x} \\ \bar{y} \neq \bar{x}}} \inf \frac{f(x, y) - f(1,0) - \langle (d_1, d_2), (x, y) - (1,0) \rangle}{\|(x, y) - (1,0)\|} \geq 0 \right\}$$

$$\hat{\partial}f(1,0) = \liminf_{\substack{\bar{y} \rightarrow \bar{x} \\ \bar{y} \neq \bar{x}}} \frac{|x-1| - 0 - \langle (d_1, d_2), ((x-1), y) \rangle}{\|(x, y) - (1, 0)\|} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{\substack{\bar{y} \rightarrow (1,0) \\ \bar{y} \neq (1,0)}} \frac{|x-1| - d_1(x-1) - d_2y}{\sqrt{y^2 + (x-1)^2}} \geq 0$$

Si $\bar{y} = (x, y)$, $\bar{x} = (1, 0)$ y haciendo $(x-1) = a$, $y = b$ tenemos que si $x \rightarrow 1$, $y \rightarrow 0 \Rightarrow a \rightarrow 0$ y $b \rightarrow 0$, por lo tanto, de la definición

$$\hat{\partial}f(1,0) = \liminf_{(a,b) \rightarrow (0,0)} \frac{|a| - d_1(a) - d_2b}{\sqrt{b^2 + a^2}} \geq 0$$

Ahora analizaremos el subdiferencial $\hat{\partial}f(1,0)$, para tal efecto se considera los siguientes casos

i) Sean $a > 0$, $b > 0$ y el conjunto $A = \{(a, b) \in \mathbb{R}^2: b = 0\}$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a - d_1(a) - d_2(0)}{\sqrt{0 + a^2}} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a(1 - d_1)}{a} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} (1 - d_1) \geq 0 \Rightarrow (1 - d_1) \geq 0 \Rightarrow d_1 \leq 1$$

Por lo tanto $\hat{\partial}f(1,0) = \{(d_1, d_2) \in \mathbb{R}^2: d_1 \leq 1, d_2 \in \mathbb{R}\}$

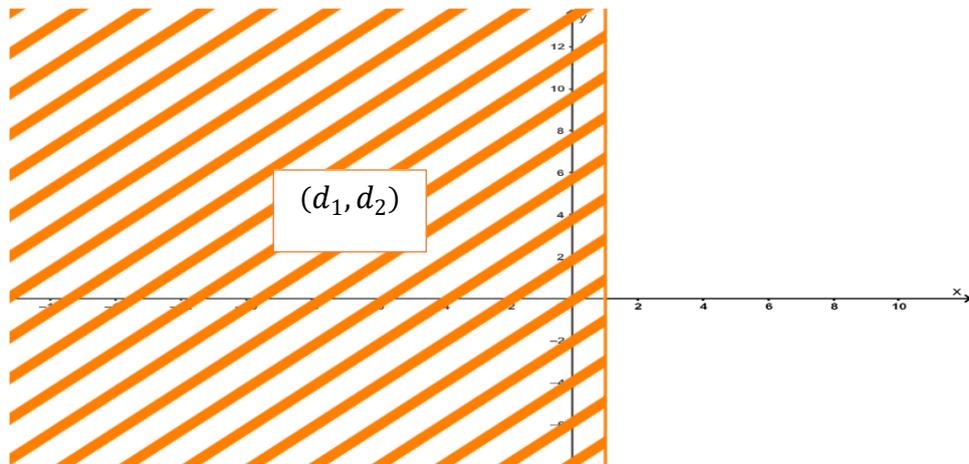


Figura 2.17. Conjunto de subdiferenciales de f para $a > 0$, $b = 0$

ii) Sean $a > 0$, $b > 0$ y el conjunto $A = \{(a, b) \in \mathbb{R}^2: a = 0\}$

$$\hat{\partial}f(1,0) = \liminf_{b \rightarrow 0} \frac{0 - d_1(0) - d_2 b}{\sqrt{0 + b^2}} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{b \rightarrow 0} \frac{-d_2 b}{b} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{b \rightarrow 0} (-d_2) \geq 0 \Rightarrow -d_2 \geq 0 \Rightarrow d_2 \leq 0$$

Por lo tanto $\hat{\partial}f(1,0) = \{(d_1, d_2) \in \mathbb{R}^2: d_2 \leq 0, d_1 \in \mathbb{R}\}$

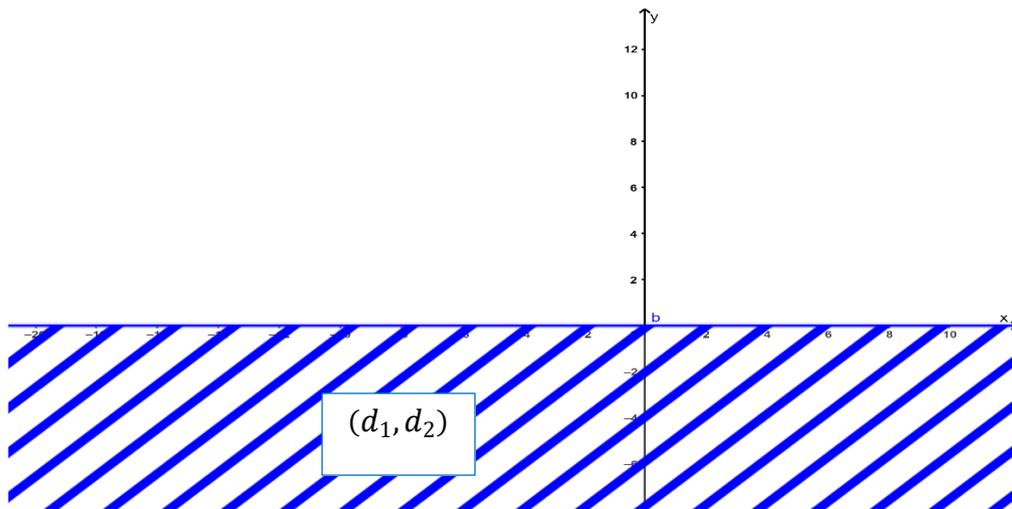


Figura 2.18. Conjunto de subdiferenciales de f para $a = 0, b > 0$

iii) Sean $a > 0, b > 0$ y el conjunto $A = \{(a, b) \in \mathbb{R}^2: a = b\}$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a - d_1(a) - d_2 a}{\sqrt{a^2 + a^2}} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a(1 - d_1 - d_2)}{\sqrt{2}a} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{(1 - d_1 - d_2)}{\sqrt{2}} \geq 0 \Rightarrow (1 - d_1 - d_2) \geq 0 \Rightarrow d_1 + d_2 \leq 1$$

Por lo tanto $\hat{\partial}f(1,0) = \{(d_1, d_2) \in \mathbb{R}^2: d_1 + d_2 \leq 1, \}$

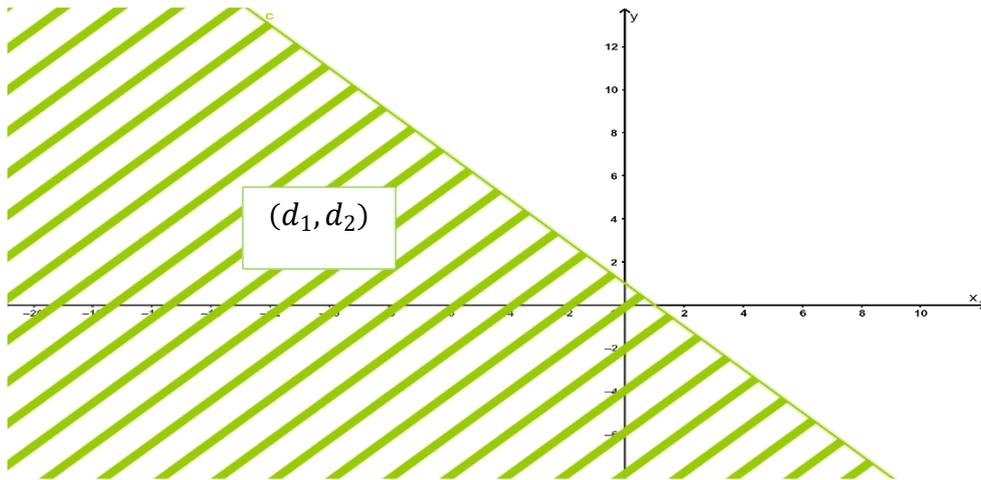


Figura 2.19. Conjunto de subdiferenciales de f cuando $a = b$

iv) Sean $a > 0$, $b > 0$ y el conjunto $A = \{(a, b) \in \mathbb{R}^2 : b = ma, m \in \mathbb{R}\}$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a - d_1(a) - d_2(ma)}{\sqrt{a^2 + m^2 a^2}} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{a(1 - d_1 - md_2)}{a\sqrt{1 + m^2}} \geq 0$$

$$\hat{\partial}f(1,0) = \liminf_{a \rightarrow 0} \frac{(1 - d_1 - md_2)}{\sqrt{1 + m^2}} \geq 0 \Rightarrow \frac{(1 - d_1 - md_2)}{\sqrt{1 + m^2}} \geq 0$$

$$\Rightarrow (1 - d_1 - md_2) \geq 0 \Rightarrow d_1 + md_2 \leq 1$$

Por lo tanto $\hat{\partial}f(1,0) = \{(d_1, d_2) \in \mathbb{R}^2 : d_1 + md_2 \leq 1, m \in \mathbb{R}\}$

El grafico del subdiferencial para valores $m \in \mathbb{R}$

Para $m > 0$ el grafico será

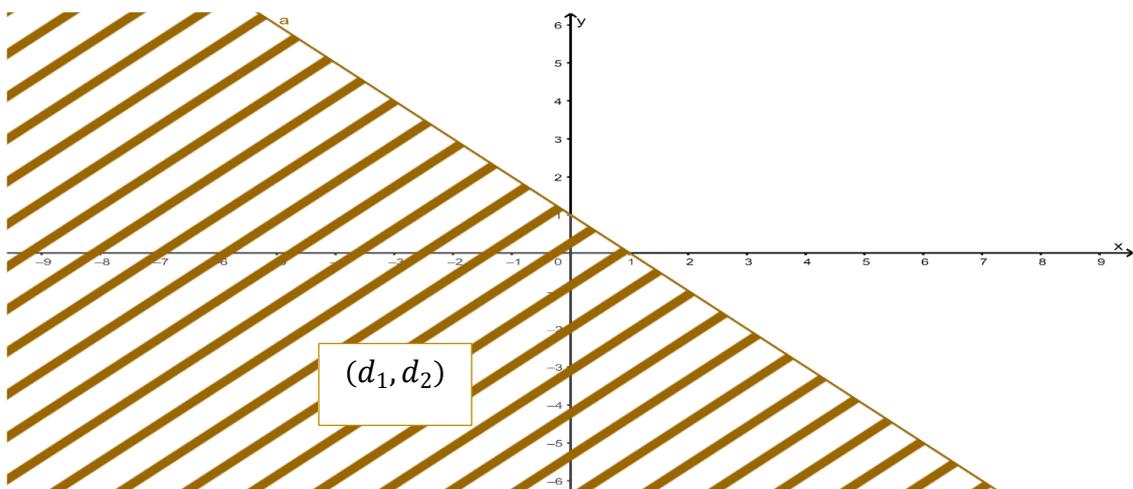


Figura 2.20. Conjunto de subdiferenciales de f cuando $b = ma, m > 0$

Para $m < 0$ el grafico será

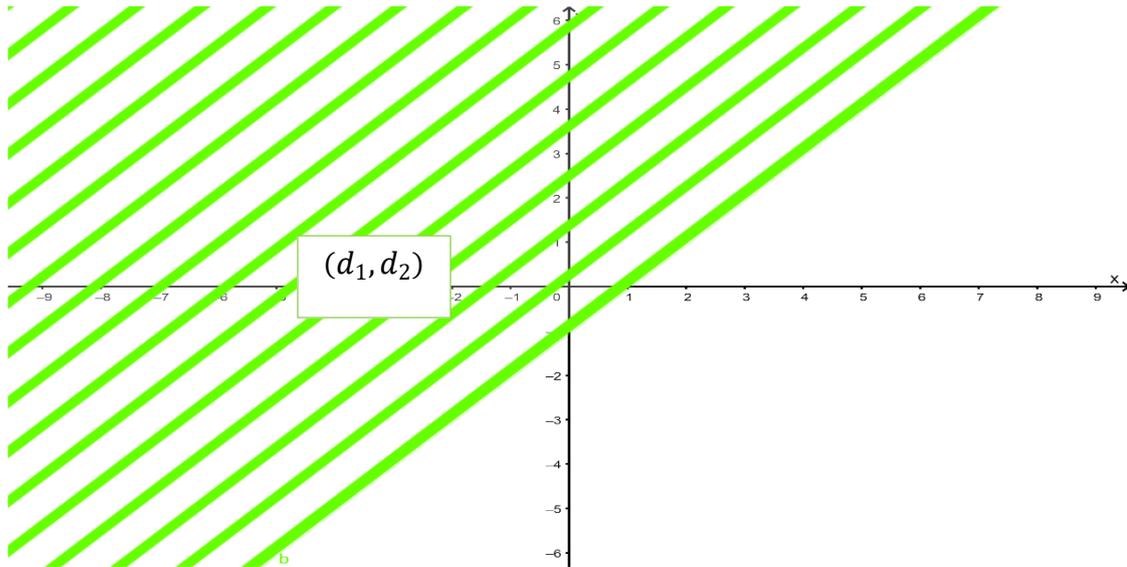


Figura 2.21. Conjunto de subdiferenciales de f cuando $b = ma, m < 0$

Proposición 2.2.4.6. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ y $\bar{x} \in \text{dom}f$. Entonces tenemos que los conjuntos subgradiantes $\hat{\partial}f(\bar{x})$ y $\partial^{Lim}f(x)$ son cerrados, con $\hat{\partial}f(\bar{x})$ convexo y $\hat{\partial}f(\bar{x}) \subset \partial^{Lim}f(x)$.

Demostración. a) Probaremos que $\hat{\partial}f(\bar{x}) \subset \partial^{Lim}f(x)$.

Considerando que $z \in \hat{\partial}f(\bar{x})$ y tomando $y^k = \bar{y}$ y $z^k = z$, para todo $k \in \mathbb{N}$, tenemos que $y^k \rightarrow \bar{y}, f(y^k) \rightarrow f(\bar{y})$ y $z^k \rightarrow z$, donde $z^k \in \hat{\partial}f(x^k)$, implicando que $z \in \partial^{Lim}f(x)$.

b) Probaremos que $\hat{\partial}f(\bar{x})$ es convexo.

En efecto, sean $u_1, u_2 \in \hat{\partial}f(\bar{x})$ y haciendo $u_t = (1-t)u_1 + tu_2$ entonces $\forall t \in [0,1]$ se tiene que

$$\begin{aligned} & f(z) - f(\bar{x}) - \langle u_t, z - \bar{x} \rangle + o(\|z - \bar{x}\|) = \\ & = f(z) - f(\bar{x}) - \langle (1-t)u_1 + tu_2, z - \bar{x} \rangle + o(\|z - \bar{x}\|) \\ & = (1-t)[f(z) - f(\bar{x}) - \langle u_1, z - \bar{x} \rangle + o(\|z - \bar{x}\|)] + \\ & \quad + t[f(z) - f(\bar{x}) - \langle u_2, z - \bar{x} \rangle + o(\|z - \bar{x}\|)], \text{ y dado que} \\ & u_1, u_2 \in \hat{\partial}f(\bar{x}), \text{ entonces } f(z) - f(\bar{x}) - \langle u_1, z - \bar{x} \rangle + o(\|z - \bar{x}\|) \geq 0 \text{ y} \end{aligned}$$

$f(z) - f(\bar{x}) - \langle u_2, z - \bar{x} \rangle + o(\|z - \bar{x}\|) \geq 0$, con $\lim_{\substack{y \rightarrow \bar{x} \\ y \neq \bar{x}}} \frac{\theta(\|z - \bar{x}\|)}{\|z - \bar{x}\|} = 0$, por lo tanto,

tenemos que $f(z) - f(\bar{x}) - \langle u_t, z - \bar{x} \rangle + o(\|z - \bar{x}\|) \geq 0$, implicando que $u_t \in \hat{\partial}f(\bar{x})$, para todo $t \in [0,1]$.

c) Probaremos que $\hat{\partial}f(\bar{x})$ es cerrado.

Sea $\{v^k\}$ una sucesión en $\hat{\partial}f(\bar{x})$ tal que $v^k \rightarrow v$. Como $v^k \in \hat{\partial}f(\bar{x})$, entonces

$$f(y) \geq f(\bar{x}) + \langle v^k, y - \bar{x} \rangle + o(\|y - \bar{x}\|)$$

Luego tomando el límite en la desigualdad con $k \rightarrow +\infty$, obtenemos

$$f(y) \geq f(\bar{x}) + \langle v, y - \bar{x} \rangle + o(\|y - \bar{x}\|), \forall y \in \mathbb{R}^n$$

Así concluimos que $v \in \hat{\partial}f(\bar{x})$.

d) Probaremos que $\partial^{Lim}f(x)$, es cerrado.

Sea $\{u^i\}$ una sucesión en $\partial^{Lim}f(x)$, tal que $u^i \rightarrow u$ y como $u^i \in \partial^{Lim}f(x)$, entonces existe una sucesión $\{x^i_k\}$ tal que $x^i_k \rightarrow \bar{x}$ y $u^i_k \in \hat{\partial}f(x^i_k)$ tal que $u^i_k \rightarrow u^i$. Por otro lado dado que $u^i_k \in \hat{\partial}f(x^i_k)$, entonces

$f(y) \geq f(x^i_k) + \langle u^i_k, y - x^i_k \rangle + o(\|y - x^i_k\|), \forall y \in \mathbb{R}^n$, y haciendo que $k \rightarrow +\infty$ tenemos que

$$f(y) \geq f(\bar{x}) + \langle u^i, y - \bar{x} \rangle + o(\|y - \bar{x}\|), \forall y \in \mathbb{R}^n, \text{ lo cual implica que } u^i \in \hat{\partial}f(\bar{x}).$$

Ahora haciendo $x^i = \bar{x}$, tenemos que $x^i \rightarrow \bar{x}$ y considerando que $u^i \rightarrow u$, con $u^i \in \hat{\partial}f(x^i)$, se tiene que $u \in \partial^{Lim}f(x)$.

Proposición 2.2.4.7. Si la función propia $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, tiene un mínimo local $\bar{x} \in \text{dom}(f)$, entonces $0 \in \hat{\partial}f(\bar{x})$

Demostración. Véase [22], página. 24.

Proposición 2.2.4.8. Sea $C \subset \mathbb{R}^n, C \neq \emptyset$, un conjunto convexo cerrado y $x \in C$, entonces $\partial^{Lim}(\delta_C)(x) = N_C(x)$.

Demostración. Véase [22], página. 30.

Proposición 2.2.4.9. Sean $f, g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funciones propias, tales que f es localmente Lipschitz en $\bar{x} \in \text{dom}(f) \cap \text{dom}(g)$, y g es semicontinua inferior en dicho punto, entonces

$$\partial^{Lim}(f + g)(\bar{x}) \subset \partial^{Lim}f(\bar{x}) + \partial^{Lim}g(\bar{x})$$

Demostración. Véase [52], Teorema 2.33

Proposición 2.2.4.10. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función convexa talque $\bar{x} \in \text{dom}f$.

Entonces $\hat{\partial}f(\bar{x}) = \{d \in \mathbb{R}^n: f(y) - f(\bar{x}) \geq \langle d, y - \bar{x} \rangle, \forall y \in \mathbb{R}^n\} = \partial^{Lim}f(\bar{x})$

Demostración. Véase [13], página 26.

Proposición 2.2.4.11. Si $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función diferenciable en \bar{x} entonces

$$\hat{\partial}f(\bar{x}) = \{\nabla f(\bar{x})\}$$

Demostración. Véase [13], página 28.

Proposición 2.2.4.12. Si $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función continuamente diferenciable en \bar{x} entonces

$$\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$$

Demostración. Véase [13], página 29.

2.2.5. Subdiferencial de Clarke

Definición 2.2.5.1. Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función localmente lipschitziana en el punto $\bar{x} \in \mathbb{R}^n$. La derivada direccional generalizada de la función f en la dirección $u \in \mathbb{R}^n$, es definida por:

$$f^0(\bar{x}, u) = \liminf_{\substack{y \rightarrow \bar{x} \\ t \downarrow 0}} \sup \frac{f(y + tu) - f(y)}{t}$$

Definición 2.2.5.2 Sea $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función localmente lipschitziana en el punto $\bar{x} \in \mathbb{R}^n$ entonces el sub-diferencial de f en el sentido de Clarke para cada x es dado por el conjunto.

$$\partial^0 f(\bar{x}) = \{\xi \in \mathbb{R}^n: f^0(\bar{x}, u) \geq \langle \xi; u \rangle, \forall u \in \mathbb{R}^n\}$$

Cada elemento $\xi \in \partial^0 f(x)$ es llamado sub-gradiente de f para cada x en el sentido de Clarke

Ejemplo 2.2.5.3. Sea la función f definida mediante $f(x) = |x|$, y por la definición 2.2.5.2 y 2.2.5.1 se tiene que

$$\partial^0 f(x) = \{\xi \in \mathbb{R}: f^0(x, v) \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}\}$$

Por lo tanto

$$\partial^0 f(x) = \left\{ \xi \in \mathbb{R}^n: \lim_{t \downarrow 0} \sup_{y \rightarrow x} \frac{f(y + tv) - f(y)}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R} \right\}$$

Tomando para $x = 0$

$$\partial^0 f(0) = \lim_{t \downarrow 0} \sup_{y \rightarrow 0} \frac{|y + tv| - |y|}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}$$

$$\begin{aligned} \partial^0 f(0) &= \lim_{t \downarrow 0} \sup_{y \rightarrow 0} \frac{|tv|}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R} \\ &= |v| \geq \xi \cdot v \end{aligned}$$

(i) Si $v \geq 0$, entonces $\xi \leq 1$

(ii) Si $v < 0$, entonces $\xi \geq -1$

Considerando (i) y (ii) se tiene $\xi \in [-1, 1]$, por lo tanto $\partial^0 f(0) = [-1, 1]$

Por otro lado, para los casos $x < 0$, y $x > 0$, como la función f es diferenciable entonces

$$\partial^0 f(x) = -1, \text{ si } x < 0 \quad \text{y} \quad \hat{\partial} f(x) = 1, \text{ si } x > 0$$

Finalmente

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

Ejemplo 2.2.5.4

Sea la función f definida mediante $f(x) = -x^2 + 1$, y por la definición 2.2.5.2 y 2.2.5.1 se tiene que

$$\partial^0 f(x) = \{\xi \in \mathbb{R}: f^0(x, v) \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}\}$$

Por lo tanto

$$\partial^0 f(x) = \left\{ \xi \in \mathbb{R}^n: \lim_{t \downarrow 0} \sup_{y \rightarrow x} \frac{f(y + tv) - f(y)}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R} \right\}$$

Tomando para $x = 0$

$$\partial^0 f(0) = \lim_{t \downarrow 0} \sup_{y \rightarrow 0} \frac{(y + tv)^2 + 1 - (-y^2 + 1)}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}$$

$$\partial^0 f(0) = \lim_{t \downarrow 0} \sup_{y \rightarrow 0} \frac{t^2 v^2 - 2y tv}{t} \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}$$

$$\begin{aligned}\partial^0 f(0) &= \lim_{t \downarrow 0} (tv^2 - 2 \sup_{y \rightarrow 0} (yv)) \geq \langle \xi; v \rangle, \quad \forall v \in \mathbb{R} \\ &= 0 \geq \xi \cdot v\end{aligned}$$

(i) Si $v \geq 0$, entonces $\xi \leq 0$

(ii) Si $v < 0$, entonces $\xi \geq 0$

Considerando (i) y (ii) se tiene $\partial^0 f(0) = 0$

Por otro lado, para los casos $x < 0, y x > 0$, como la función f es diferenciable entonces

$$\partial^0 f(x) = -2x, \text{ si } x < 0 \text{ y } x > 0.$$

Finalmente

$$\partial^0 f(x) = \begin{cases} 0 & ; x = 0 \\ -2x & ; x < 0 \\ -2x & ; x > 0 \end{cases}$$

Observación 2.2.5.5. De la definición 2.2.5.2 se tiene directamente que para todo $x \in \text{dom}f$ se tiene que $\hat{\partial}f(x) \subset \partial^{\text{lim}}f(x) \subset \partial^0 f(x)$ (Ver [15], inclusión (7))

Lema 2.2.5.6. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, funciones propias y localmente Lipschitz en $x \in \text{dom}f \cap \text{dom}g$. Luego para todo $d \in \mathbb{R}^n$ se tiene que:

$$\text{i) } (f + g)^0(x, d) \leq f^0(x, d) + g^0(x, d)$$

$$\text{ii) } (\lambda f)^0(x, d) = \lambda f^0(x, d), \text{ para todo } \lambda \geq 0$$

$$\text{iii) } f^0(x, \lambda d) = \lambda f^0(x, d), \text{ para todo } \lambda \geq 0$$

Demostración. Véase [22], pág. 32.

Lema 2.2.5.7. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función propia localmente Lipschitz en $x \in \text{dom}(f)$, y λ un escalar arbitrario, entonces

$$\partial^0(\lambda f) = \lambda \partial^0 f(x)$$

Demostración. Si la función f es localmente Lipschitz entonces λf es también localmente Lipschitz. A continuación, probaremos que:

1.- $\partial^0(\lambda f)(x) \subset \lambda \partial^0 f(x)$. Sea $w \in \partial^0(\lambda f)(x)$, entonces por la Definición 2.2.5.2 se tiene que

$$\langle w, d \rangle \leq (\lambda f)^0(x, d) \quad \forall d \in \mathbb{R}^n.$$

Dado que λ un escalar arbitrario, analizaremos los siguientes casos:

i) Si $\lambda \geq 0$, por el Lema 2.2.5.6 (ii) se tiene que $(\lambda f)^0 = \lambda f^0$, entonces de la desigualdad anterior se tiene

$$\langle w, d \rangle \leq \lambda f^0(x, d) \quad \forall d \in \mathbb{R}^n.$$

Ahora si $\lambda = 0$ en $\partial^0(\lambda f) = \lambda \partial^0 f(x)$ se verifica la igualdad y si $\lambda > 0$ se tiene $\langle \frac{w}{\lambda}, d \rangle \leq f^0(x, d) \quad \forall d \in \mathbb{R}^n$ lo cual implica que $\frac{w}{\lambda} \in \partial^0 f(x)$, de donde se concluye que $w \in \lambda \partial^0 f(x)$

ii) Para probar el caso cuando $\lambda < 0$, será suficiente estudiar $\lambda = -1$. Sea $w \in \partial^0(-f)(x)$, entonces de forma análoga por el Lema 2.2.5.6 (ii) tenemos que

$$\langle w, d \rangle \leq (-f)^0(x, d) \quad \forall d \in \mathbb{R}^n$$

Pero por la Definición 2.2.5.1 se tiene

$$f^0(x, -d) := \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y - td) - f(y)}{t}, \text{ hacemos } y - td = v, \text{ luego}$$

$$f^0(x, -d) := \limsup_{\substack{v \rightarrow x \\ t \downarrow 0}} \frac{(-f)(v + td) - (-f)(v)}{t} = (-f)^0(x, d)$$

entonces $\langle -w, -d \rangle = \langle w, d \rangle \leq f^0(x, -d) \quad \forall d \in \mathbb{R}^n$, implicando que $-w \in \partial^0 f(x)$. Por lo tanto, de (i) e (ii) se tiene que $\partial^0(\lambda f)(x) \subset \lambda \partial^0 f(x)$ para todo $\lambda \in \mathbb{R}$.

2.- $\lambda \partial^0 f(x) \subset \partial^0(\lambda f)(x)$. Tomando un $w \in \partial^0 f(x)$, haciendo uso de la Definición 2.2.5.2 y siguiendo de forma inversa lo realizado en el ítem (1), se probará fácilmente dicha inclusión.

Lema 2.2.5.8. Sean $f_i: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ $i = 1, \dots, m$, funciones propias y localmente Lipschitz en $x \in \bigcap_{i=1}^m \text{dom}(f_i)$, entonces.

$$(\sum_{i \in I}^m f_i)^0(x) \subset \sum_{i \in I}^m f_i^0(x)$$

Demostración. La prueba se da forma inmediata a partir del lema 2.2.5.6 (i)

Proposición 2.2.5.9. Sean $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función propia localmente Lipschitz en \mathbb{R}^n , entonces f^0 es semicontinua superior (scs), es decir si $(x, d) \in$

$\mathbb{R}^n \times \mathbb{R}^n$ y $\{(x^k, d^k)\}$ es una sucesión en $\mathbb{R}^n \times \mathbb{R}^n$, tal que $\lim_{n \rightarrow +\infty} (x^k, d^k) = (x, d)$, entonces

$$\limsup_{n \rightarrow +\infty} f^0(x^k, d^k) \leq f^0(x, d)$$

Demostración. Véase [22], páginas 33-34.

Proposición 2.2.5.10. Sean $g: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función propia, cuasi-convexa y localmente Lipschitz en \mathbb{R}^n . Si $d \in g^0(x)$ tal que $\langle d, y - x \rangle > 0$, entonces

$$g(y) \geq g(x).$$

Demostración. Véase [22], página. 35.

Proposición 2.2.5.11. Sean $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función propia, convexa y localmente Lipschitz, entonces $\partial^0 f(x)$ coincide con el sub-diferencial en x y $f^0(x, d)$ coincide con la derivada direccional $f'(x, d)$ para cada d .

Demostración. Véase [22], páginas 36-37.

Teorema 2.2.5.12. Si una función $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es coerciva y semicontinua inferior en un conjunto $X \neq \emptyset$, entonces existe un punto que representa un mínimo global de f en X .

Demostración. Véase [57], corolario 3.4.

Lema 2.2.5.13. Sean $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, una función localmente Lipschitz y $g: \mathbb{R}^n \rightarrow \mathbb{R}$, una función convexa, entonces

$$(f + g)^0(x, d) = f^0(x, d) + g^0(x, d) \quad \forall x, d \in \mathbb{R}^n$$

Además, si función g es diferenciable se tiene que

$$\partial^0(f + g)(x, d) = \partial^0 f(x) + \nabla g(x) \quad \forall x \in \mathbb{R}^n$$

Demostración. Véase [22], página. 38.

Definición 2.2.5.14. Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función localmente lipschitziana en el punto $x \in \mathbb{R}^n$ entonces el sub-diferencial de F en el sentido de Clarke para cada x es dado por el conjunto.

$$\partial F(x) = \{w^t \in \mathbb{R}^{m \times n}: F^0(x, d) \geq \langle w^t; d \rangle, \forall d \in \mathbb{R}^n\}$$

Cada elemento $w^t \in \partial F(x)$ es llamado sub-gradiente de F para cada x en el sentido de Clarke y $F^0(x, d) = (f_1^0(x, d), f_2^0(x, d), \dots, f_n^0(x, d))$, con $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, n$.

Observación 2.2.5.15. Si consideramos $m = 1$ obtenemos la definición 2.2.5.2, es decir

$$\partial F(x) = \partial^0 f(x) = \{\xi \in \mathbb{R}^n: f^0(x, v) \geq \langle \xi; v \rangle, \forall v \in \mathbb{R}^n\}$$

2.2.7 Condiciones de optimalidad y direcciones de descenso

Definición 2.2.7.1 Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$, con $F(\bar{x}) = (F_1(\bar{x}), \dots, F_m(\bar{x}))$, una función localmente Lipschitz en \mathbb{R}^n . Diremos que \hat{x} es un punto crítico Pareto-Clarke de F si para todas las direcciones $v \in \mathbb{R}^n$, existe $i_0 \in \{1, \dots, m\}$ tal que $F_{i_0}^0(\hat{x}, v) \geq 0$.

Observación 2.2.7.2. Considerando la Definición 2.2.7.1 un punto \bar{x} no es un punto crítico de Pareto-Clarke, si existe una dirección $v \in \mathbb{R}^n$ que satisfice.

$$F_i^0(x^k, v) < 0, \forall i \in \{1, \dots, m\}$$

Implicando, que, para cada $i \in \{1, \dots, p\}$, v^k es una dirección de descenso, para cada función F_i , es decir, existe $\varepsilon > 0$, tal que

$$F_i(x^k + tv) < F_i(x^k), \forall t \in \langle 0, \varepsilon \rangle \text{ y } \forall i \in \{1, \dots, m\}$$

Definición 2.2.7.3 Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $v \in \mathbb{R}^n$, se dice que d es una dirección de descenso para la función multiobjetivo F en x^k , si existe $\varepsilon > 0$ tal que

$$F(x^k + tv) < F(x^k), \forall t \in \langle 0, \varepsilon \rangle.$$

Definición 2.2.7.4. Sea $F: C \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, con C conjunto convexo no vacío y $F(\bar{x}) = (F_1(\bar{x}), \dots, F_m(\bar{x}))$, una función localmente Lipschitz en C , Diremos que $\hat{x} \in C$ es un punto crítico Pareto-Clarke de F en C si para cualquier $y \in C$, existe $i_0 \in \{1, \dots, m\}$ tal que $F_{i_0}^0(\hat{x}, y - \hat{x}) \geq 0$.

Observación 2.2.7.5. Considerando la Definición 2.2.7.4, un punto \hat{x} no es un punto crítico de Pareto-Clarke, si existe $y \in C$ que satisfice.

$$F_i^0(\hat{x}, y - \hat{x}) < 0 \quad \forall i \in \{1, \dots, m\}$$

Implicando que para todo $i \in \{1, \dots, m\}$, existe $y \in C$, para cada función F_i , es decir, existe $\varepsilon > 0$, tal que

$$F_i(\hat{x} + t(y - \hat{x})) < F_i(\hat{x}), \quad \forall t \in \langle 0, \varepsilon \rangle \quad y \quad \forall i \in \{1, \dots, m\}$$

2.2.8 Convergencia de sucesiones

Lema 2.2.8.1 Sean las sucesiones no negativas $\{v_k\}$, $\{\tau_k\}$ y $\{\beta_k\}$ de números reales que satisfacen

$$v_{k+1} \leq (1 + \tau_k)v_k + \beta_k$$

donde $\sum_{k=0}^{+\infty} \tau_k < +\infty$, $\sum_{k=0}^{+\infty} \beta_k < +\infty$. Entonces la sucesión v_k es convergente.

Definición 2.2.8.2 Sea $\{x^k\} \subset \mathbb{R}^n$ tal que $\{x^k\}$ converge a un punto $\hat{x} \in \mathbb{R}^n$ y $(d, H) \in \mathcal{F}_+(\bar{C})$. Entonces se dice que la convergencia es:

i. H -lineal, si existe $\theta \in \mathbb{R}$, $\theta < 1$ y un número positivo $n \in \mathbb{N}$ tal que

$$H(\hat{x}, x^{k+1}) \leq \theta H(\hat{x}, x^k) \quad \forall k > n.$$

ii. H -superlineal, si existe $\{\beta_k\}$ una sucesión convergente a cero y un número positivo $\bar{n} \in \mathbb{N}$ tal que

$$H(\hat{x}, x^{k+1}) \leq \beta_k H(\hat{x}, x^k) \quad \forall k > \bar{n}.$$

iii. H -orden n , si existe $\theta > 0$, $n \geq 1$ y un número positivo $\bar{n} \in \mathbb{N}$ tal que

$$H(\hat{x}, x^{k+1}) \leq \theta [H(\hat{x}, x^k)]^n \quad \forall k > \bar{n}.$$

CAPITULO III

MÉTODO DEL PUNTO PROXIMAL ESCALARIZADO INEXACTO CON DISTANCIA PROXIMAL GENERALIZADA (ISPPMR)

3.1 El algoritmo

Estamos interesados en resolver el problema de optimización multiobjetivo (POM) con restricciones

$$Opt\{(F_1(x), \dots, F_m(x)) : x \in C \subset \mathbb{R}^n\} \quad (3.1)$$

donde el conjunto C es convexo con $int(C) \neq \emptyset$ y $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ satisface las siguientes hipótesis:

(H1) $F_i(\bar{x}) \geq 0$, para todo $i = 1, 2, \dots, m$ y para todo $\bar{x} \in \mathbb{R}^n$.

(H2) F es localmente Lipschitz en \mathbb{R}^n .

(H3) F es estrictamente o semi-estricta quasi-conveva en \mathbb{R}^n .

Para solucionar el problema (3.1) proponemos un método de punto proximal escalarizado inexacto con regularización de distancia proximal, denotado por ISPPMR

Algoritmo ISPPMR

Inicio: Elección de un punto inicial arbitrario

$$x^0 \in C \quad (3.2)$$

Paso iterativo: Para $k = 0, 1, 2, \dots$ y dado un x^k , encontrar $x^{k+1} \in \Omega_k$ y $e^{k+1} \in \mathbb{R}^n$ tal que:

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \lambda_k d(\cdot, x^k) \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1}) \quad (3.3)$$

donde ∂^0 es el subdiferencial de Clarke, $\Omega_k = \{x \in C: F(x) \preceq F(x^k), \forall k \in \mathbb{N}\}$, $\lambda_k > 0$, $\{z_k\} \subset \mathbb{R}_+^n$, $\|z_k\| = 1$ y $N_{\Omega_k}(x^{k+1})$ es el cono normal a Ω_k y x^{k+1} , $d(\cdot, \cdot)$ es una distancia proximal sobre C .

Criterio de parada: Si $x^{k+1} = x^k$ o si x^{k+1} es un punto crítico Pareto-clarke, detener el algoritmo. De lo contrario hacer $k \leftarrow k + 1$ y retornar al paso iterativo.

Observación 3.1. En la práctica para obtener el punto x^{k+1} que satisfaga (3.3), solamente se debe encontrar un punto crítico aproximado (mínimo local, máximo local o punto silla), del siguiente problema de optimización

$$\min \left\{ \sum_{i=1}^m z_i^k F_i(\cdot) + \lambda_k d(\cdot, x^k) : F_i(x) \leq F_i(x^k), i = 1, \dots, m \right\}$$

3.2. Existencia de los iterados

Teorema 3.2.1. Si $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una función que satisface las hipótesis **(H1)** **(H2)**, **(H3)** y $(d, H) \in \mathcal{F}(\bar{C})$, entonces la sucesión $\{x^k\}$ generada por el algoritmo ISPPMR está bien definido.

Demostración. Sea $x^0 \in C$ un punto arbitrario de inicialización. Asumamos que x^k exista, y definamos la función $\varphi_k(x) = \langle F(x), z_k \rangle + \lambda_k d(x, x^k) + \delta_{\Omega_k}(x)$, donde $\delta_{\Omega_k}(\cdot)$, es la función indicadora de Ω_k . Así tendremos la siguiente equivalencia

$$\min\{\varphi_k(x) : x \in \mathbb{R}^n\} \cong \min\{\langle F(x), z_k \rangle + \lambda_k d(x, x^k) : x \in \Omega_k\}.$$

Ahora analizando cada una de las funciones que componen $\varphi_k(x)$, tenemos que. $\langle F(x), z_k \rangle$, es acotada inferiormente, dado que $F_i(\bar{x}) \geq 0, \forall i = 1, 2, \dots, m, \forall \bar{x} \in \mathbb{R}^n$ (ver (H1)), y $\{z_k\} \subset \mathbb{R}^m_{\geq}$, por otro lado $\langle F(x), z_k \rangle$ es continua, también que $d(\cdot, x^k)$, es continua y coerciva (ver (Definición 2.2.2.1)) y $\delta_{\Omega_k}(x)$ es semicontinua inferior y acotada superiormente, por lo tanto $\varphi_k(x)$ es una función es semicontinua inferior, coerciva y propia y por la Teorema 2.2.5.12. se tiene que existe un $x^{k+1} \in \Omega_k$ que representa un mínimo global de $\varphi_k(x)$, y la Proposición 2.2.4.7. se tiene que

$$0 \in \partial^0(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k) + \delta_{\Omega_k}(\cdot))(x^{k+1})$$

Ahora usando la proposición Proposición 2.2.3.6, de la expresión anterior implica que

$$0 \in \partial^{Lim}(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k) + \delta_{\Omega_k}(\cdot))(x^{k+1})$$

Luego dado que $\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k)$ es localmente Lipschitz y $\delta_{\Omega_k}(\cdot)$ es sci (ver Proposición 2.2.1.9), por la Proposición 2.2.4.9, de la expresión anterior se tiene que

$$0 \in \partial^{Lim}(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k}(\cdot))(x^{k+1})$$

Por otro lado como Ω_k es un conjunto convexo cerrado y $x^{k+1} \in \Omega_k$, entonces por la Proposición 2.2.4.8 de la expresión anterior tenemos que

$$0 \in \partial^{Lim}(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1}) + N_{\Omega_k}(x^{k+1}) \quad (3.4)$$

Finalmente haciendo uso de la Observación 2.2.5.5, y considerando $e^{k+1} = 0$, la iteración (3.3) es obtenida de (3.4).

Observación 3.2.2. En este método estamos interesados en la convergencia asintótica del algoritmo ISPPM por lo que asumimos a lo largo de esta tesis que en cada iteración x^k no es un punto crítico de Pareto-Clarke. Es decir, si existe un $y \in C$ que satisface.

$$F^0_i(x^k, y - x^k) < 0 \quad \forall i \in \{1, \dots, m\}$$

implicando que para todo $i \in \{1, \dots, m\}$, existe $y \in C$, para cada función F_i , es decir, existe $\varepsilon > 0$, tal que

$$F_i(x^k + t(y - x^k)) < F_i(x^k), \quad \forall t \in \langle 0, \varepsilon \rangle \text{ y } \forall i \in \{1, \dots, m\}$$

En otras palabras, existe un elemento de C para la función multiobjetivo F en x^k , es decir, existe $\varepsilon > 0$ tal que

$$F(x^k + t(y - x^k)) \prec F(x^k), \quad \forall t \in \langle 0, \varepsilon \rangle$$

además, implica que el interior de Ω_{k+1} , denotado por $int(\Omega_{k+1})$, no es vacío.

Por otro lado, si la condición $x^{k+1} \neq x^k$ no satisface, es decir, si hay un k_0 tal que $x^{k_0+1} = x^{k_0}$, entonces fácilmente se puede probar que este punto es un punto crítico de Pareto-Clarke de F .

A continuación, presentamos otra hipótesis sobre la función F y el punto inicial $x^0 \in \mathbb{R}^m_{\geq}$, el cual es citado en varios trabajos relacionados con algoritmos proximales vectoriales, ver, por ejemplo, Bonnel et al. (2005), Ceng y Yao (2007) y Villacorta y Oliveira (2011). Por tanto, consideramos la siguiente hipótesis adicional:

(H4) $(F(x^0) - \mathbb{R}^m_{\geq}) \cap F(C)$, es \mathbb{R}^m_{\geq} completo

lo que significa que para toda sucesión $\{a^k\} \subset \mathbb{R}^n$ con $a^0 = x^0$, tal que

$F(a^{k+1}) \leq F(a^k)$, $\forall k \in \mathbb{N}$, existe un $a \in \mathbb{R}^n$ tal que $F(a) \leq F(a^k)$, para todo $k \in \mathbb{N}$.

Denotemos el conjunto

$$E := \{x \in C : F(x) \leq F(x^k), \forall k \in \mathbb{N}\}$$

Este conjunto junto con las hipótesis **(H1)** - **(H4)** cumple las siguientes condiciones:

1. $E \subseteq \Omega_k$, para cualquier $k \in \mathbb{N}$, y $E \neq \emptyset$, en efecto por el Teorema 3.2.1, la secuencia $\{x^{k+1}\}$ generado por el algoritmo ISPPM existe y satisface $F(x^{k+1}) \leq F(x^k)$. Entonces, por la afirmación **(H4)** existe $a \in \mathbb{R}^n$ tal que $F(a) \leq F(x^k)$, así $a \in E$.

2.-Además, por **(H2)** y **(H3)** está claro que el conjunto E también es cerrado y convexo.

3.3 Resultados de convergencia

El siguiente resultado es fundamental en el análisis de convergencia de la sucesión generada por el algoritmo ISPPMR.

Lema 3.3.1. Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función que satisface **(H1)**, **(H2)**, **(H3)** **(H4)**, $(d, H) \in \mathcal{F}_+(\bar{C})$, y $\{x^k\}$ una sucesión generada por el algoritmo ISPPMR. Si \hat{x} es un punto de acumulación de $\{x^k\}$, entonces $\hat{x} \in E \cap C$.

Prueba. Sea $\hat{x} \in \mathbb{R}^n$ es un punto de acumulación de $\{x^k\}$, entonces existe $\{x^{k_j}\} \subseteq \{x^k\}$ tal que $x^{k_j} \rightarrow \hat{x}$. Esto implica que $\hat{x} \in C$. Probaremos que $\hat{x} \in E$. Por la hipótesis **(H2)**, F es localmente Lipschitz en \mathbb{R}^n , la función $\langle F(\cdot), z \rangle$ es también localmente Lipschitz, por lo tanto, es continua en \mathbb{R}^n en particular para todo $z \in \mathbb{R}^m_{\geq} \setminus \{0\}$. Resultando que $\lim_{j \rightarrow \infty} \langle F(x^{k_j}), z \rangle = \langle F(\hat{x}), z \rangle$.

Por otro lado, dado que $x^{k+1} \in \Omega_k$, entonces $F(x^{k+1}) \leq F(x^k)$ donde $F_i(x^{k+1}) \leq F_i(x^k)$ para todo $i \in \{1, \dots, p\}$ y como $z \in \mathbb{R}^m_{\geq} \setminus \{0\}$, entonces

$$\langle F(x^{k+1}), z \rangle \leq \langle F(x^k), z \rangle.$$

Así, $\langle F(\cdot), Z \rangle$ es decreciente y también está acotado inferiormente ($\langle F(\cdot), Z \rangle \geq 0$), por la hipótesis **(H1)**, y además $z \in \mathbb{R}^m_{\geq} \setminus \{0\}$, luego la sucesión $\{\langle F(x^k), Z \rangle\}$ es convergente. Entonces

$$\langle F(\hat{x}), z \rangle = \lim_{k \rightarrow +\infty} \langle F(x^k), z \rangle = \inf_{k \in \mathbb{N}} \{\langle F(x^k), z \rangle\} \leq \langle F(x^k), z \rangle.$$

De la desigualdad anterior tenemos

$$\langle F(x^k) - F(\hat{x}), z \rangle \geq 0 \quad \forall k \in \mathbb{N} \quad y \quad z \in \mathbb{R}^m_{\geq} \setminus \{0\}$$

lo que implica que $F(x^k) - F(\hat{x}) \in \mathbb{R}^m_{\geq}$ donde $F(\hat{x}) \preceq F(x^k) \quad \forall k \in \mathbb{N}$ y por lo tanto $\hat{x} \in E$.

Proposición 3.3.2. Sean $\{x^k\}$ y $\{e^k\}$ sucesiones generadas por el algoritmo ISPPMR. Si se cumplen las afirmaciones (H1), (H4) y $(d, H) \in \mathcal{F}_+(\bar{C})$, entonces para todo $\bar{x} \in E$ y para todo $k \in \mathbb{N}$ tenemos.

$$H(\bar{x}, x^{k+1}) \leq H(\bar{x}, x^k) - \gamma H(x^{k+1}, x^k) - \frac{1}{\lambda_k} \langle e^{k+1}, \bar{x} - x^{k+1} \rangle$$

Demostración

De la relación (3.3) y dado que $\langle F(\cdot), Z_k \rangle = \sum_{i=1}^m z^i_k F_i(\cdot)$ es una función localmente Lipschitz y $\nabla_1 d(\cdot, x^k)$ es una función convexa diferenciable, tenemos

$$e^{k+1} \in \partial^0 \left(\sum_{i=1}^m z^i_k F_i(\cdot) \right) (x^{k+1}) + \lambda_k \nabla_1 d(x^{k+1}, x^k) + N_{\Omega_k}(x^{k+1}) \quad (3.5)$$

Ahora aplicando la propiedad de subdiferencial de Clarke, Lema 2.2.5.7 y Lema 2.2.5.8 tenemos

$$\partial^0 \left(\sum_{i=1}^m z^i_k F_i(\cdot) \right) (x^{k+1}) \subset \sum_{i=1}^m \partial^0 \left(z^i_k F_i(x^{k+1}) \right) = \sum_{i=1}^m z^i_k \partial^0 F_i(x^{k+1})$$

Así que en (5)

$$e^{k+1} \in \sum_{i=1}^m z^i_k \partial^0 F_i(x^{k+1}) + \lambda_k \nabla_1 d(x^{k+1}, x^k) + N_{\Omega_k}(x^{k+1})$$

La expresión anterior indica la existencia de vectores

$g^k_i \in \partial^0 F_i(x^{k+1})$ y $v_k \in N_{\Omega_k}(x^{k+1})$, talque

$$-\nabla_1 d(x^{k+1}, x^k) = \frac{1}{\lambda_k} \left(\sum_{i=1}^m z^i_k g^k_i + v_k - e^{k+1} \right) \quad (3.6)$$

Ahora tomando un punto arbitrario $\bar{x} \in E$, de la definición de E , tenemos que $\bar{x} \in \Omega_k$ para todos k , en particular $\bar{x} \in \Omega_{k+1}$. Dado que $\text{int}\Omega_{k+1} \neq \emptyset$, entonces existe $\{x^\ell\} \subset \text{int}\Omega_{k+1}$ talque $x^\ell \rightarrow \bar{x}$.

Por otro lado, está claro que dado que la función F es estrictamente cuasi-convexa y localmente Lipschitz, $g^k_i \in \partial^0 F_i(x^{k+1})$ y $F_i(x^\ell) < F_i(x^{k+1})$ para todo $i \in \{1, \dots, m\}$ (esto es cierto porque $x^\ell \in \text{int}\Omega_{k+1}$), de la propiedad de cuasiconvexidad para cada i tenemos

$$\langle g^k_i, x^{k+1} - x^\ell \rangle \geq 0, \quad \forall i \in \{1, \dots, m\} \quad (3.7)$$

Ahora, dado que $v_k \in N_{\Omega_k}(x^{k+1})$, entonces $\langle v_k, \mathcal{W} - x^{k+1} \rangle \leq 0$ para todo $\mathcal{W} \in \Omega_k$, y como $x^\ell \in \Omega_k \subset \Omega_{k+1}$ entonces

$$\langle v_k, x^{k+1} - x^\ell \rangle \geq 0 \quad (3.8)$$

Tomando $c = x^\ell$ y $b = x^{k+1}$ y $a = x^k$, en la siguiente desigualdad:

$$\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a) \quad \forall c \in C, \gamma \in \langle 0, 1 \rangle$$

(ver Definición 2.2.2.2)

Obtenemos

$$\langle x^\ell - x^{k+1}, \nabla_1 d(x^{k+1}, x^k) \rangle \leq H(x^\ell, x^k) - H(x, x^{k+1}) - \gamma H(x^{k+1}, x^k) \quad \forall c \in C, \gamma \in \langle 0, 1 \rangle \quad (3.9)$$

De la ecuación (3.6) y usando las desigualdades (3.7) y (3.8) en (3.9), tenemos

$$H(x^\ell, x^{k+1}) \leq H(x^\ell, x^k) - \gamma H(x^{k+1}, x^k) - \frac{1}{\lambda_k} \langle e^{k+1}, x^\ell - x^{k+1} \rangle,$$

para todos $k \in \mathbb{N}$ y tomando $l \rightarrow +\infty$ obtenemos el resultado.

Proposición 3.3.3. Sean $\{x^k\}$ y $\{e^k\}$ sucesiones generadas por el algoritmo ISPPMR. Si se cumplen las hipótesis **(H1)**, **(H3)**, **(H4)**, $(d, H) \in \mathcal{F}(\bar{C})$. Si las siguientes condiciones se cumplen

$$i. \frac{\|e^{k+1}\|}{\lambda_k} \leq \eta_k \sqrt{H(x^{k+1}, x^k)} \quad (3.10)$$

$$ii. -\sum_{k=0}^{+\infty} \eta_k < +\infty \quad (3.11)$$

iii. -Existe $\theta > 0$, talque

$$\|x - y\|^2 \leq \theta H(x, y), \quad \forall x \in \bar{C} \quad y \quad \forall y \in C \quad (3.12)$$

Entonces se tiene que

a. Existe un número natural k_0 talque para todo $k \geq k_0$ y para todo $\bar{x} \in E$, entonces

$$H(\bar{x}, x^{k+1}) \leq \left(1 + \frac{\theta \eta_k}{1 - \theta \eta_k}\right) H(\bar{x}, x^k) + \left(\frac{\eta_k}{4} - \gamma\right) H(x^{k+1}, x^k);$$

b. $\{H(\bar{x}, x^{k+1})\}$ converge para todo $\bar{x} \in E$;

c. $\{x^k\}$ es acotada;

d. $\lim_{k \rightarrow +\infty} H(x^{k+1}, x^k) = 0$.

e. Si $(d, H) \in \mathcal{F}_+(C)$, entonces $\{x^k\}$ converge en E .

Demostración

a. Sea $\bar{x} \in E$, y de la desigualdad

$$\begin{aligned} 0 &\leq \left\| \frac{e^{k+1}}{\sqrt{2\lambda_k \eta_k}} + \sqrt{2\lambda_k \eta_k}(\bar{x} - x) \right\|^2 \\ \left\| \frac{e^{k+1}}{\sqrt{2\lambda_k \eta_k}} + \sqrt{2\lambda_k \eta_k}(\bar{x} - x) \right\|^2 &= \frac{\|e^{k+1}\|^2}{2\lambda_k \eta_k} + 2\lambda_k \eta_k \|\bar{x} - x\|^2 + 2 \left\langle \frac{e^{k+1}}{\sqrt{2\lambda_k \eta_k}}, \sqrt{2\lambda_k \eta_k}(\bar{x} - x) \right\rangle \\ 0 &\leq \frac{\|e^{k+1}\|^2}{2\lambda_k \eta_k} + 2\lambda_k \eta_k \|\bar{x} - x\|^2 + 2 \left\langle \frac{e^{k+1}}{\sqrt{2\lambda_k \eta_k}}, \sqrt{2\lambda_k \eta_k}(\bar{x} - x) \right\rangle \\ -2 \left\langle \frac{e^{k+1}}{\sqrt{2\lambda_k \eta_k}}, \sqrt{2\lambda_k \eta_k}(\bar{x} - x) \right\rangle &\leq \frac{\|e^{k+1}\|^2}{2\lambda_k \eta_k} + 2\lambda_k \eta_k \|\bar{x} - x\|^2 \\ -\langle e^{k+1}, (\bar{x} - x) \rangle &\leq \frac{\|e^{k+1}\|^2}{2\lambda_k \eta_k} + \lambda_k \eta_k \|\bar{x} - x\|^2 \end{aligned}$$

$$-\left(\frac{1}{\lambda_k}\right)\langle e^{k+1}, (\bar{x} - x) \rangle \leq \frac{\|e^{k+1}\|^2}{4\lambda_k^2\eta_k} + \eta_k \|\bar{x} - x\|^2.$$

Haciendo $x = x^{k+1}$ obtenemos

$$-\left(\frac{1}{\lambda_k}\right)\langle e^{k+1}, (\bar{x} - x^{k+1}) \rangle \leq \frac{\|e^{k+1}\|^2}{4\lambda_k^2\eta_k} + \eta_k \|\bar{x} - x^{k+1}\|^2.$$

Usando la proposición 3.3.2, tenemos

$$H(\bar{x}, x^{k+1}) \leq H(\bar{x}, x^k) - \gamma H(x^{k+1}, x^k) + \frac{\|e^{k+1}\|^2}{4\lambda_k^2\eta_k} + \eta_k \|\bar{x} - x^{k+1}\|^2$$

Considerando las afirmaciones (3.10) y (3.12) obtenemos

$$(1 - \theta\eta_k)H(\bar{x}, x^{k+1}) \leq H(\bar{x}, x^k) + \left(\frac{\eta_k}{4} - \gamma\right)H(x^{k+1}, x^k) \text{ de (3.11) } \eta_k \rightarrow 0 \text{ y } \theta > 0$$

entonces existe k_0 tal que $0 < 1 - \theta\eta_k \leq 1$ para todo $k \geq k_0$. Luego en la desigualdad anterior obtenemos

$$H(\bar{x}, x^{k+1}) \leq \left(\frac{1}{(1-\theta\eta_k)}\right)H(\bar{x}, x^k) + \left(\frac{1}{(1-\theta\eta_k)}\right)\left(\frac{\eta_k}{4} - \gamma\right)H(x^{k+1}, x^k)$$

Por otro lado como $0 < 1 - \theta\eta_k \leq 1$, y $\left(\frac{\eta_k}{4} - \gamma\right) < 0 \rightarrow \left(\frac{1}{(1-\theta\eta_k)}\right)\left(\frac{\eta_k}{4} - \gamma\right) \leq \left(\frac{\eta_k}{4} - \gamma\right)$

Finalmente, en la desigualdad anterior obtenemos el resultado deseado

$$H(\bar{x}, x^{k+1}) \leq \left(\frac{1}{(1-\theta\eta_k)}\right)H(\bar{x}, x^k) + \left(\frac{\eta_k}{4} - \gamma\right)H(x^{k+1}, x^k)$$

b. Como $\left(\frac{\eta_k}{4} - \gamma\right) < 0$ y de la desigualdad del ítem a), tenemos que

$$H(\bar{x}, x^{k+1}) \leq \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right)H(\bar{x}, x^k), \text{ para todo } k \geq k_0.$$

Por otro lado $\eta_k \rightarrow 0^+$ y $\theta > 0$, entonces para todo $0 < \varepsilon < 1$, existe \widetilde{k}_0 tal que $\theta\eta_k < \varepsilon$, para todo $k \geq \widetilde{k}_0$, luego si $1 - \varepsilon < 1 - \theta\eta_k \leq 1$, tenemos

$$\frac{\theta\eta_k}{1 - \theta\eta_k} < \frac{\theta\eta_k}{1 - \varepsilon}, \quad \forall k \geq \widetilde{k}_0$$

Aplicando sumatoria a la desigualdad anterior y por la condición (3.11) tenemos que

$$\sum_{k=0}^{+\infty} \frac{\theta\eta_k}{1-\theta\eta_k} < +\infty \text{ y haciendo } v_{k+1} = H(\bar{x}, x^{k+1}), v_k = H(\bar{x}, x^k), \tau_k = \frac{\theta\eta_k}{1-\theta\eta_k}$$

y $\beta_k = 0$ y por el lema 2.2.7.1 y considerando que $\sum_{k=0}^{+\infty} \tau_k < +\infty$, obtenemos que la sucesión $\{H(\bar{x}, x^{k+1})\}$ converge para todo $\bar{x} \in E$.

c. Probaremos que la sucesión $\{x^k\}$ es acotada.

En efecto del ítem b) tenemos que $\{H(\bar{x}, x^{k+1})\}$, converge para todo $\bar{x} \in E$, donde se cumple las condiciones del Lema 2.2.7.1 es decir dado que

$$H(\bar{x}, x^{k+1}) \leq \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\bar{x}, x^k), \text{ para todo } k \geq k_0, \text{ y haciendo}$$

$$v_{k+1} = H(\bar{x}, x^{k+1}), v_k = H(\bar{x}, x^k), \tau_k = \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right) \text{ y } \beta_k = 0$$

$$\text{Satisface } \sum_{k=0}^{+\infty} \frac{\theta\eta_k}{1-\theta\eta_k} < +\infty \text{ y } \sum_{k=0}^{+\infty} \beta_k < +\infty.$$

Por lo tanto, la sucesión $\{x^k\}$ es acotada

d. Probaremos que $\lim_{k \rightarrow +\infty} H(x^{k+1}, x^k) = 0$

En efecto de los ítems b) y c) se tiene que

$$\left(\frac{\eta_k}{4} - \gamma\right) H(x^{k+1}, x^k) \leq \left(\frac{1}{1-\theta\eta_k}\right) H(\bar{x}, x^k) - H(\bar{x}, x^{k+1}), \text{ la sucesión } \{x^k\} \text{ es acotada}$$

y $\{H(\bar{x}, x^{k+1})\}$ converge para todo $\bar{x} \in E$.

Por lo tanto, tomado limite cuando $k \rightarrow +\infty$, en la desigualdad anterior y usando la definición 2.1.4.2 (vi) tendremos que

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

Por lo tanto, el resultado

e. Probaremos que si $(d, H) \in \mathcal{F}_+(C)$, entonces $\{x^k\}$ converge en E .

En efecto de los ítems b) y c) se tiene que la sucesión $\{x^k\}$ es acotada y $\{H(\bar{x}, x^k)\}$ converge para todo $\bar{x} \in E$ y usando la Definición 2.2.2.2 (vi) se tiene que

$$\lim_{k \rightarrow +\infty} H(\bar{x}, x^k) = 0$$

en consecuencia

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

por lo tanto, el resultado.

Teorema 3.3.4. Si la función $F: R^n \rightarrow R^m$, talque $F_i: R^n \rightarrow R \cup \{+\infty\}$, satisface las hipótesis **(H1)**, **(H2)**, **(H3)** y **(H4)**, con $0 < \lambda_k \leq \bar{\lambda}$, entonces la sucesión generada por ISPPMR converge hacia algún punto crítico Pareto Clarke del POM (3.1)

Demostración. Probaremos que \hat{x} es un punto crítico Pareto-Clarke, (Ver definición 2.2.7.4) en efecto por la Proposición 3.3.3 ítem e) existe un $\hat{x} \in E$ talque $x^k \rightarrow \hat{x}$, considerando que \hat{x} no es un punto crítico Pareto-Clarke, entonces existe $y \in C$ talque $F^0_i(\hat{x}, y - \hat{x}) < 0 \quad \forall i \in \{1, \dots, m\}$, lo cual implica que existe $y \in C$ para la función F en \hat{x} es decir existe $\delta > 0$ talque

$$F_i(\hat{x} + \lambda(y - \hat{x})) < F_i(\hat{x}), \quad \forall \lambda \in \langle 0, \delta \rangle, \quad i = 1, 2, \dots, m \quad (3.13)$$

Po lo tanto $(\hat{x} + \lambda(y - \hat{x})) \in \Omega_k$.

Así mismo por el Teorema 3.2.1 la sucesión $\{x^{k+1}\}$ generada por el ISPPM existe y $x^{k+1} \in \Omega_k$, además por la relación (3.5) se tiene que

$$e^{k+1} \in \partial^0(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1}) + N_{\Omega_k}(x^{k+1}), \text{ implicando que}$$

$$(-\lambda_k \nabla_1 d(x^{k+1}, x^k) - v_k + e^{k+1}) \in \partial^0(\langle F(\cdot), Z_k \rangle)(x^{k+1}),$$

donde $v_k \in N_{\Omega_k}(x^{k+1})$. Haciendo uso de la definición 2.2.4.2 correspondiente a $\partial^0(\cdot)$ tenemos

$$\langle -\lambda_k \nabla_1 d(x^{k+1}, x^k) - v_k + e^{k+1}, r \rangle \leq \langle F(\cdot), Z_k \rangle^0(x^{k+1}, r), \quad \forall r \in C \quad (3.14)$$

donde $\langle F(\cdot), Z_k \rangle^0(x^{k+1}, r)$, representa la derivada direccional de Clarke. Ahora por propiedades de producto interno de vectores y $\lambda_k > 0$, en la desigualdad anterior (3.14) tenemos

$$\langle -\lambda_k \nabla_1 d(x^{k+1}, x^k) + e^{k+1}, r \rangle - \langle v_k, r \rangle \leq \langle F(\cdot), Z_k \rangle^0(x^{k+1}, r), \quad \forall r \in C \quad (3.15)$$

Considerando que $r = (\hat{x} + \lambda(y - \hat{x})) - x^{k+1}$, y dado que $v_k \in N_{\Omega_k}(x^{k+1})$, se tiene que $\langle v_k, r \rangle \leq 0$, y de forma equivalente $0 \leq -\langle v_k, r \rangle$ donde $(\hat{x} + \lambda(y - \hat{x})) \in \Omega_k$.

Luego en la desigualdad anterior (3.15)

$$\langle -\lambda_k \nabla_1 d(x^{k+1}, x^k) + e^{k+1}, r \rangle \leq \langle F(\cdot), Z_k \rangle^0(x^{k+1}, r) \quad (3.16)$$

Ahora aplicando el Lema 2.2.5.6, ítems (i), (ii) y (iii) a la desigualdad (3.16) término de la derecha se tiene

$$\langle F(\cdot), z_k \rangle^0(x^{k+1}, r) = \left(\sum_{i=0}^m z_k^i F_i(\cdot) \right)^0(x^{k+1}, r) \leq \sum_{i=0}^m \left(z_k^i F_i(x^{k+1}, r) \right)^0$$

$$\langle F(\cdot), z_k \rangle^0(x^{k+1}, r) \leq \sum_{i=0}^m F_i^0(x^{k+1}, z_k^i r),$$

donde z_k^i son las componentes del vector z_k y reemplazando en la expresión (3.16) se tiene

$$\langle e^{k+1} - \lambda_k \nabla_1 d(x^{k+1}, x^k), r \rangle \leq \sum_{i=0}^m F_i^0(x^{k+1}, z_k^i r) \quad (3.17)$$

Por otro lado dado que $\langle F(\cdot), Z_k \rangle$ es una función localmente Lipschitz y $d(\cdot, x^k)$ es estrictamente diferenciable (ver definición 2.2.2.1 (i)) y por el Lema 2.2.5.2 en la expresión (17) existe, g^k talque $g^k = e^{k+1} - \lambda_k \nabla_1 d(x^{k+1}, x^k)$, y también considerando $\{x^{k_j}\} \subseteq \{x^k\}$ se tiene

$$\langle e^{k_j+1} - \lambda_k \nabla_1 d(x^{k_j+1}, x^{k_j}), r \rangle \leq \sum_{i=0}^m F_i^0(x^{k_j+1}, z_{k_j}^i r') \quad (3.18)$$

Donde $r' = (\hat{x} + \lambda(y - \hat{x})) - x^{k_j+1}$ y como $\{Z_k\}$ es acotada existe una sub-secuación $\{z_k^i\}$, talque $\lim_{j \rightarrow +\infty} z_k^i = \bar{z}$. Ahora tomando límite superior cuando $j \rightarrow +\infty$ en el término izquierdo de la desigualdad (3.18) tiende a cero, donde se tiene

$$0 \leq \lim_{j \rightarrow +\infty} \sup \sum_{i=0}^m F_i^0(x^{k_j+1}, z_{k_j}^i r') = \sum_{i=0}^m \lim_{j \rightarrow +\infty} \sup F_i^0(x^{k_j+1}, z_{k_j}^i r')$$

Ahora dado que la función F es propia y localmente Lipschitz por la Proposición 2.2.5.9 se tiene que F_i^0 es semicontinua superior. Entonces

$$\lim_{j \rightarrow +\infty} \sup F_i^0(x^{k_j+1}, z_{k_j}^i r') \leq F_i^0(\hat{x}, \bar{z}^i \lambda(y - \hat{x})) \quad (3.19)$$

Donde $\{(x^{k_j+1}, z_{k_j}^i r')\}$ es una sucesión de $\mathbb{R}^n \times \mathbb{R}^n$, en la cual por las condiciones anteriormente mencionadas satisface $\lim_{j \rightarrow +\infty} (x^{k_j+1}, z_{k_j}^i r') = (\hat{x}, \bar{z}^i \lambda(y - \hat{x}))$ y reemplazando en la desigualdad (19) se tiene

$$0 \leq \sum_{i=0}^m \lim_{j \rightarrow +\infty} \sup F_i^0(x^{k_{j+1}}, z_{k_j}^i r') \leq \sum_{i=0}^m F_i^0(\hat{x}, \bar{z}^i \lambda(y - \hat{x}))$$

Luego haciendo uso del Lema 2.2.5.6 (iii) y dado que $\lambda > 0$, se tiene que

$$\begin{aligned} \sum_{i=0}^m \bar{z}^i F_i^0(\hat{x}, (y - \hat{x})) &= \bar{z}^1 F_1^0(\hat{x}, (y - \hat{x})) + \bar{z}^2 F_2^0(\hat{x}, (y - \hat{x})) + \dots + \\ &+ \bar{z}^m F_m^0(\hat{x}, (y - \hat{x})) \end{aligned} \quad (3.20)$$

Además, como $\bar{z} \in \mathbb{R}_{\geq}^m \setminus \{0\}$, entonces $\bar{z} \neq 0$. Así existe un conjunto de índices en los cuales $\bar{z} > 0$. Sin pérdida de generalidad consideremos el conjunto $J = \{i \in \{1, \dots, m\}: \bar{z}^i > 0\}$, entonces en (3.20) existe $i_0 \in J$ talque $F_{i_0}^0(\hat{x}, (y - \hat{x})) \geq 0$, lo cual contradice (13). Por lo tanto \hat{x} es un punto crítico Pareto Clarke del POM (3.1).

CAPITULO IV

TASA DE CONVERGENCIA DEL ALGORITMO ISPPMR

En este capítulo se analizará la convergencia del algoritmo ISPPMR, considerando las condiciones para el error de aproximación e^{k+1} , la cual estará dado por.

Algoritmo ISPPMR

Inicio: Elección de un punto inicial arbitrario

$$x^0 \in C \quad (4.1)$$

Paso iterativo: Para $k = 0, 1, 2, \dots$ y dado un x^k , encontrar $x^{k+1} \in \Omega_k$ y $e^{k+1} \in \mathbb{R}^n$ tal que:

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \lambda_k d(\cdot, x^k) \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1}) \quad (4.2)$$

donde ∂^0 es el subdiferencial de Clarke, $\Omega_k = \{x \in C: F(x) \leq F(x^k), \forall k \in \mathbb{N}\}$, $\lambda_k > 0$, $\{z_k\} \subset \mathbb{R}_+^n$, $\|z_k\| = 1$, $N_{\Omega_k}(x^{k+1})$ es el cono normal a Ω_k y x^{k+1} , $d(\dots)$ es una distancia proximal sobre C y el término del error e^{k+1} satisfaciendo las siguientes condiciones

$$\frac{\|e^{k+1}\|}{\lambda_k} \leq \eta_k \sqrt{H(x^{k+1}, x^k)} \quad (4.3)$$

$$\sum_{k=0}^{+\infty} \eta_k < +\infty \quad (4.4)$$

Criterio de parada: Si $x^{k+1} = x^k$ o si x^{k+1} es un punto crítico Pareto-clarke, detener el criterio. De lo contrario hacer $k \leftarrow k + 1$ y retornar al paso iterativo.

El algoritmo ISPPMR tiene como objetivo resolver el problema de optimización multiobjetivo cuasi-convexo, como se vio en el capítulo tres, es decir

$$Opt\{(F_1(x), \dots, F_m(x)) : x \in C \subset \mathbb{R}^n\} \quad (4.5)$$

donde el conjunto C es convexo con $int(C) \neq \emptyset$ y la función $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ satisface las siguientes hipótesis:

(H1) $F_i(\bar{x}) \geq 0$, para todo $i = 1, 2, \dots, m$ y $\bar{x} \in dom(F_i)$

(H2) F es localmente Lipschitz en \mathbb{R}^n

(H3) F es estrictamente o semi-estricta quasi-conveva en \mathbb{R}^n

(H4) $(F(x^0) - \mathbb{R}^m_{\geq}) \cap F(C)$, es \mathbb{R}^m_{\geq} completo

Además del conjunto convexo cerrado no vacío descrito en el capítulo anterior, dado por

$$E = \{x \in C: F(x) \leq F(x^k), \forall k \in \mathbb{N}\} \quad (4.6)$$

Por otro lado, también denotemos al conjunto

$$W = E \cap P_C \quad (4.7)$$

donde P_C representa al conjunto de puntos críticos Pareto-clarke del (POM) (4.5) (ver definición 2.2.7.4), dado por:

$$P_C = \{\hat{x} \in C: F_{i_0}^0(\hat{x}, y - \hat{x}) \geq 0, \forall y \in C \text{ y para algun } i_0 \in \{1, \dots, m\}\}.$$

Observación 4.0.1 Considerando (3) y del hecho que $\partial^{Lim}(\delta_{\Omega_k})(x^{k+1}) = N_{\Omega_k}(x^{k+1})$ (ver Proposición 2.2.4.8), se tiene que

$$e^{k+1} \in \partial^0(\langle F(\cdot), z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1})$$

Por otro lado, dado que $\langle F(\cdot), z_k \rangle$ es una función localmente Lipschitz (ver Definición 2.2.1.30) y $d(\cdot, x^k)$ es convexa diferenciable (ver Definición 2.2.2.1). Aplicando el Lema 2.2.2.5 a la expresión anterior tenemos que

$$e^{k+1} \in \partial^0(\langle F(\cdot), z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}) =$$

$$e^{k+1} \in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \lambda_k \nabla_1 d(\cdot, x^k)(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}),$$

donde $e^{k+1} - \lambda_k \nabla_1 d(\cdot, x^k)(x^{k+1}) \in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1})$. (4.8)

Así existe un $v^k \in \left(\partial^0(\langle F(\cdot), z_k \rangle + \lambda_k d(\cdot, x^k))(x^{k+1})\right)$, tal que

$$v^k = e^{k+1} - \lambda_k \nabla_1 d(x^{k+1}, x^k) \quad (4.9)$$

4.1 Análisis de la tasa de convergencia

Para hacer el análisis de tasa de convergencia se va a considerar las siguientes hipótesis:

(H5) Para $\hat{x} \in W$ tal que $x^k \rightarrow \hat{x}$, existen $\delta = \delta(\hat{x}) > 0$ y $\rho_k = \rho_k(\hat{x}) > 0$, tales

que para todo $\omega \in B(0, \delta) \subset \mathbb{R}^n$ y para todo x^k con $\omega \in \left(\partial^0 \left(\langle F(\cdot), z_k \rangle + \partial^{Lim}(\delta_{\Omega_k})(\cdot) \right) (x^{k+1}) \right)$, se tiene que

$$H(\hat{x}, x^{k+1}) \leq \rho_k \|\omega\|^2, \text{ donde } (d, H) \in \mathcal{F}_+(\bar{C}) \quad (4.10)$$

(H6) Dado que $(d, H) \in \mathcal{F}_+(\bar{C})$ (ver Definición 2.1.4.2), la función $\nabla_1 d(\cdot, u)$ satisface la siguiente condición:

Para cualquier $x_0 \in \bar{C}$, existen $L > 0$ y $r > 0$ tal que

$$\|\nabla_1 d(x, u) - \nabla_1 d(\hat{x}, u)\| \leq L \|x - \hat{x}\|, \quad \forall x, \hat{x} \in B(x_0, r) \cap C, \forall u \in C \quad (4.11)$$

Observación 4.1.1 La hipótesis mencionada anteriormente es llamada la condición de crecimiento en el punto de convergencia $\hat{x} \in \mathbb{R}^n$ (ver [22], página. 83).

El conjunto W (ver relación 4.7) y las hipótesis **(H5)**, **(H6)** que se están proponiendo para el análisis de la tasa de convergencia del algoritmo ISPPMR, representa una generalización del trabajo propuesto por Cruzado [22] para la optimización multiobjetivo.

Lema 4.1.2 Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el algoritmo ISPPMR, $(d, H) \in \mathcal{F}_+(\bar{C})$, y supongamos que las hipótesis **(H1)**, **(H2)**, **(H3)**, **(H4)**, **(H5)**, **(H6)** son satisfechas con la condición $0 < \lambda_k \leq \bar{\lambda}$, entonces

i) Existe $\tilde{k} \in \mathbb{N}$ tal que

$$\|v^k\| \leq \delta \quad \forall k \geq \tilde{k}, \text{ donde } v^k \text{ es dado por (3.9)} \quad (4.12)$$

ii) Se tiene que

$$H(\hat{x}, x^{k+1}) \leq \rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2 H(x^k, x^{k+1}) \quad \forall k \geq \tilde{k} \quad (4.13)$$

Demostración.

i) Sean $x^k \rightarrow \hat{x}$ y $v^k \in \left(\partial^0 \left(\langle F(\cdot), z_k \rangle + \partial^{Lim}(\delta_{\Omega_k})(\cdot) \right) (x^{k+1}) \right)$ (ver relación 4.9).

Dado que $\frac{\|e^{k+1}\|}{\lambda_k} \leq \eta_k \sqrt{H(x^{k+1}, x^k)}$, $\lambda_k \leq \bar{\lambda}$, y **(H6)** se tiene que

$$\|v^k\| \leq \|e^{k+1}\| + \lambda_k \|\nabla_1 d(x^{k+1}, x^k)\|$$

$$\|v^k\| \leq \lambda_k \eta_k \sqrt{H(x^{k+1}, x^k)} + \lambda_k L \|x^{k+1} - x^k\|,$$

por la condición iii) de la Proposición 3.3.3, en la desigualdad anterior se tiene

$$\begin{aligned} \|v^k\| &\leq \lambda_k \eta_k \sqrt{H(x^{k+1}, x^k)} + \lambda_k L \sqrt{\theta} \sqrt{H(x^{k+1}, x^k)} \\ \|v^k\| &\leq \lambda_k (\eta_k + L\sqrt{\theta}) \sqrt{H(x^{k+1}, x^k)} \\ \|v^k\| &\leq \bar{\lambda} (\eta_k + L\sqrt{\theta}) \sqrt{H(x^{k+1}, x^k)} \end{aligned} \quad (4.14)$$

Como $\eta_k \rightarrow 0$, y $H(x^{k+1}, x^k) \rightarrow 0$ (ver Proposición 3.3.3 (d)), entonces existe un $\tilde{k} \in \mathbb{N}$ tal que $\|v^k\| < \delta$ para todo $k \geq \tilde{k}$.

ii) En la relación (4.10) haciendo $\omega = v^k$ para todo $k \geq \tilde{k}$ y considerando la relación (4.14), es decir $\|v^k\| \leq \lambda_k (\eta_k + L\sqrt{\theta}) \sqrt{H(x^{k+1}, x^k)}$ se tiene que:

$$H(\hat{x}, x^{k+1}) \leq \rho_k \|\omega\|^2 = \rho_k \|v^k\|^2 \leq \rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2 H(x^{k+1}, x^k)$$

$$H(\hat{x}, x^{k+1}) \leq \rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2 H(x^{k+1}, x^k) \quad \forall k \geq \tilde{k}.$$

Teorema 4.1.3 Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el algoritmo ISPPMR, $(d, H) \in \mathcal{F}_+(\bar{C})$, y supongamos que las hipótesis **(H1)**, **(H2)**, **(H3)**, **(H4)**, **(H5)**, **(H6)** son satisfechas con la condición $0 < \lambda_k \leq \bar{\lambda}$, entonces la sucesión $\{x^k\}$ converge H -linealmente hacia $\hat{x} \in W$. Además, si $\{\lambda_k\}, \{\rho_k\}$ convergen a cero, entonces la convergencia es H -superlineal.

Demostración. Sean $\hat{x} \in W$ un punto limite de $\{x^k\}$ y $v^k \in \left(\partial^0 \left(\langle F(\cdot), z_k \rangle + \right. \right.$

$\left. \partial^{Lim}(\delta_{\Omega_k})(\cdot) \right)(x^{k+1})$ dado por la relación (4.9). Definimos

$$\omega^k = v^k$$

considerando la relación (4.12) tenemos que

$$\|\omega^k\| < \delta \quad \forall k \geq \tilde{k}$$

Por lo tanto $\omega^k \in B(0, \delta)$ para todo $k \geq \tilde{k}$.

Por otro lado, por la Proposición 3.3.2 (a), para todo $k \geq k_0$ se tiene que

$$H(\hat{x}, x^{k+1}) \leq \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\hat{x}, x^k) + \left(\frac{\eta_k}{4} - \gamma\right) H(x^{k+1}, x^k) \quad \forall \hat{x} \in E \quad (4.15)$$

Por la relación (4.13) también tenemos que

$$\begin{aligned} H(\hat{x}, x^{k+1}) &\leq \rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2 H(x^k, x^{k+1}) \quad \forall k \geq \tilde{k} \\ -H(x^k, x^{k+1}) &\leq -\left(\frac{1}{\rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2}\right) H(\hat{x}, x^{k+1}) \end{aligned} \quad (4.16)$$

de donde para todo $k \geq \max\{k_0, \tilde{k}\}$ reemplazando la relación (4.16) en (4.15) se tiene

$$H(\hat{x}, x^{k+1}) \leq \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\hat{x}, x^k) - \left(\gamma - \frac{\eta_k}{4}\right) \left(\frac{1}{\rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2}\right) H(\hat{x}, x^{k+1})$$

de donde

$$\left(1 + \frac{4\gamma - \eta_k}{4\rho_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2}\right) H(\hat{x}, x^{k+1}) \leq \left(\frac{1}{1-\theta\eta_k}\right) H(\hat{x}, x^k)$$

Como $\rho_k > 0$ y $(4\gamma - \eta_k) > 0$, entonces para todo $k \geq \max\{k_0, \tilde{k}\}$ tenemos que

$$H(\hat{x}, x^{k+1}) \leq \beta_k H(\hat{x}, x^k) \quad (4.17)$$

donde

$$\beta_k = \left(\frac{4\rho_k(\eta_k + L\sqrt{\theta})^2}{4\rho_k(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\lambda_k^2}}\right) \left(\frac{1}{1-\theta\eta_k}\right) \quad (4.18)$$

Dado que $0 < \lambda_k \leq \bar{\lambda}$ para todo $k \in \mathbb{N}$, se tiene que

$$\beta_k \leq r_k$$

donde

$$r_k = \left(\frac{4\rho_k(\eta_k + L\sqrt{\theta})^2}{4\rho_k(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\bar{\lambda}^2}}\right) \left(\frac{1}{1-\theta\eta_k}\right) \quad (4.19)$$

como $\eta_k \rightarrow 0$, en la expresión anterior tenemos que

$$r_k \rightarrow \left(\frac{4\rho L^2 \theta}{4\rho L^2 \theta + \frac{4\gamma}{\lambda^2}} \right)$$

Por lo tanto, existe un numero positivo $k_1 \in \mathbb{N}$ con $k \geq k_1$, talque

$$\beta_k \leq r_k < \frac{1}{2} \left(1 + \frac{4\rho L^2 \theta}{4\rho L^2 \theta + \frac{4\gamma}{\lambda^2}} \right) < 1 \quad \forall k \geq k_1$$

Entonces en (4.17) se tiene que

$$H(\hat{x}, x^{k+1}) \leq \bar{\theta} H(\hat{x}, x^k) \quad \forall k \geq \max\{k_0, \tilde{k}, k_1\}$$

donde

$$\bar{\theta} = \left(\frac{4\rho L^2 \theta}{4\rho L^2 \theta + \frac{4\gamma}{\lambda^2}} \right)$$

Por lo tanto, la sucesión $\{x^k\}$ converge H -linealmente a \hat{x} dado que $\bar{\theta} < 1$ (ver Definición 2.2.8.2 (i)).

En el caso de la convergencia H -superlineal de la sucesión generada por el algoritmo ISPPMR, si $\{\rho_k\}$ y $\{\lambda_k\}$ converge a cero en la relación (4.18), tenemos que $\beta_k \rightarrow 0$. Entonces la sucesión $\{x^k\}$ converge H -superlineal a \hat{x} . (ver Definición 2.2.8.2 (ii)).

CAPITULO V

EXPERIMENTACIÓN NUMÉRICA

En esta sección damos algunos ejemplos numéricos que muestran la funcionalidad del algoritmo propuesto. Toda la experimentación numérica fue realizada en una laptop personal (*core i7-7500 de 2.90 GHz de 64 bits con 2 núcleos y 16 GB de RAM*) usando Matlab 2022. Además, las soluciones referenciales a los problemas planteados sobre los cuales serán comparados nuestro algoritmo también fueron calculadas usando el toolbox de Optimización de Matlab 2022.

Ejemplo 5.1. Sea el problema de minimización multiobjetivo definido mediante

$$\min\{(F_1(x_1, x_2), F_2(x_1, x_2)): x_1 \geq 0, x_2 \geq 0\}$$

donde $F_1(x_1, x_2) = -2e^{-2x_1^2 - x_1^2} + 2$ y $F_2(x_1, x_2) = 2(x_1 - 1)^2 + (x_2 - 1)^2$.

La función F_1 es una función diferenciable cuasi-convexa y F_2 es una función cuasi-convexa suave, por lo tanto $(F_1(\cdot), F_2(\cdot))^T$ función vectorial cuasi-convexa de clase $C^1(\mathbb{R}_+^2)$ (ver Teorema 2.2.1.23). Este problema satisface las hipótesis **(H1)**, **(H2)** y **(H3)** y, además, se puede observar claramente en las figuras 5.1,5.2 que los mínimos para cada una de las funciones componentes F_1 y F_2 son alcanzados en los puntos (0,0) y (1,1) respectivamente. Cabe mencionar que, en el contexto del análisis multiobjetivo, las soluciones de Pareto también incluyen a estas soluciones, como se observa en la figura 5.3.

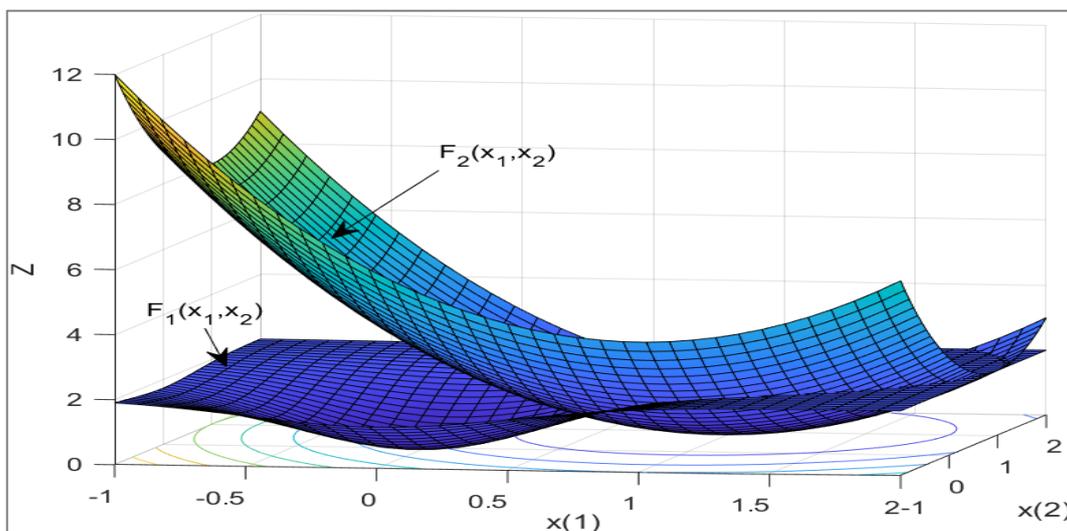


Figura 5.1. Funciones componentes objetivo ejemplo 5.1

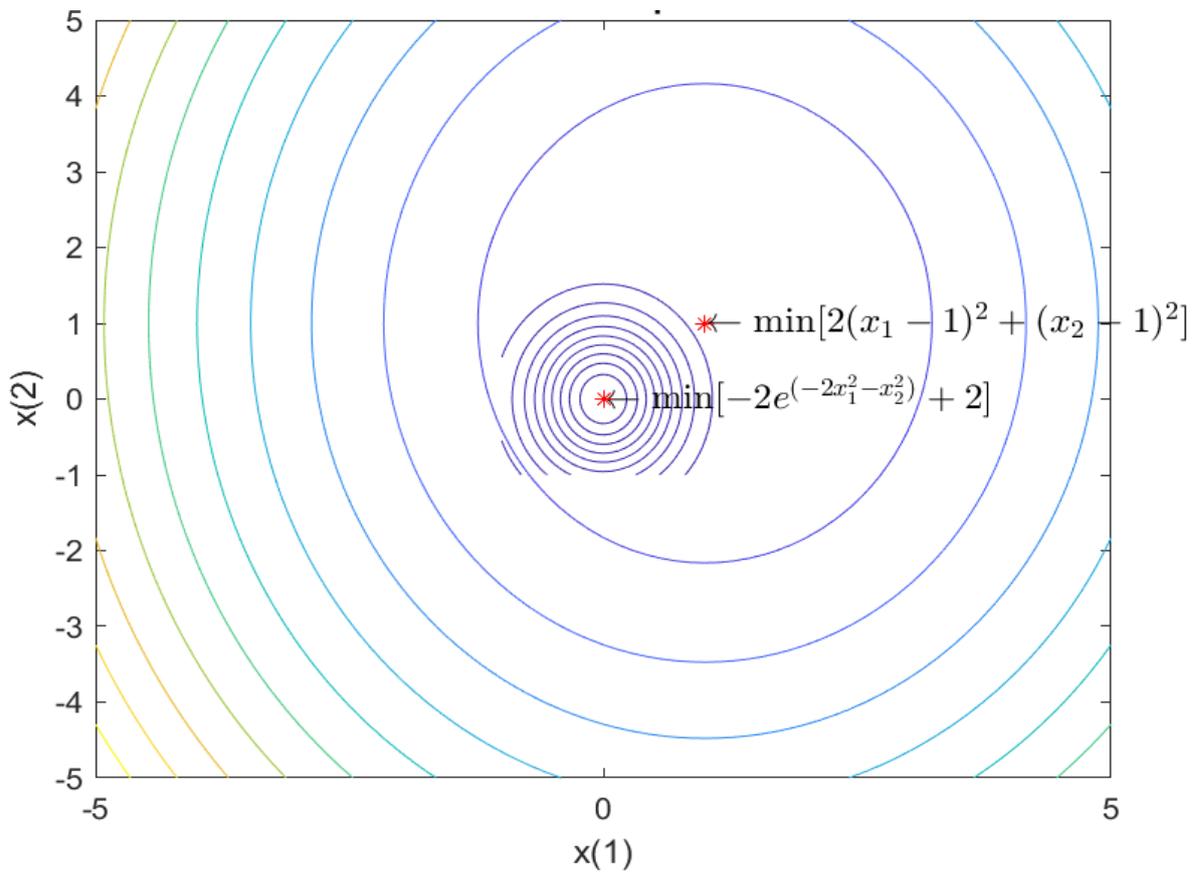


Figura 5.2. Curvas de nivel de las funciones objetivo ejemplo 5.1

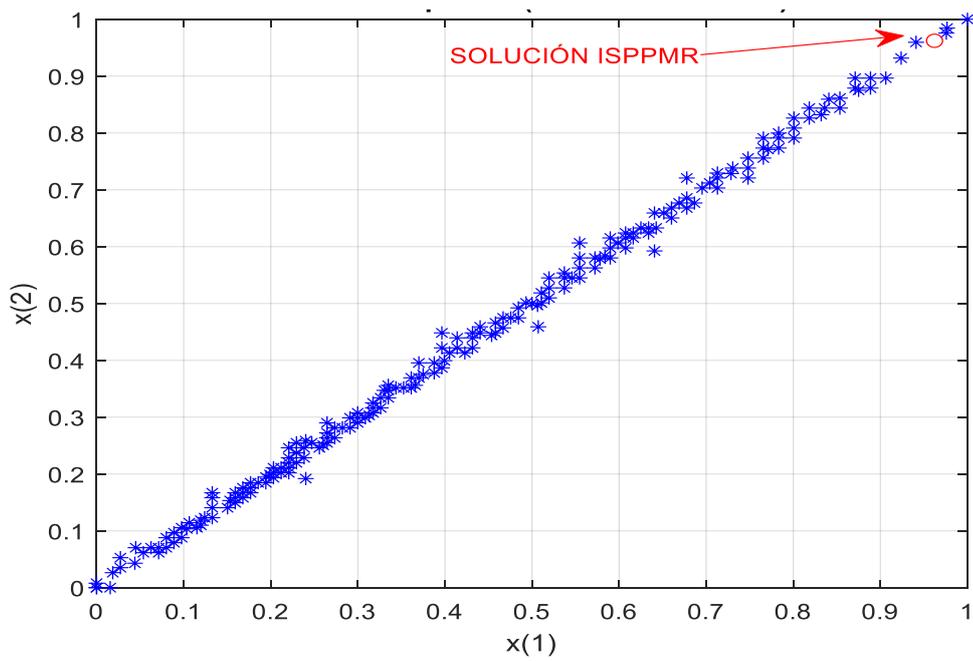


Figura 5.3. Soluciones de Pareto ejemplo 5.1

Ahora tomando como punto inicial $x^0 = (x_1^0, x_2^0)$ y dado $x^k \in \mathbb{R}_+^2$, el paso principal de la propuesta del algoritmo es encontrar un punto crítico (mínimo local, máximo local o un punto de silla) del siguiente problema, considerando los parámetros fijos $z_k = (z_1^k, z_2^k)$, λ_k para cada k .

$$\left\{ \begin{array}{l} \text{ming}(x_1, x_2) = (-2e^{-2x_1^2 - x_1^2} + 2)z_1^k + (2(x_1 - 1)^2 + (x_2 - 1)^2)z_2^k + \\ \quad \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sujeto a:} \\ \quad 2x_1^2 + x_2^2 \leq 2(x_1^k)^2 + (x_2^k)^2 \\ \quad (x_1 - 1)^2 + (x_2 - 1)^2 \leq (x_1^k - 1)^2 + (x_2^k - 1)^2 \end{array} \right.$$

la cual de forma equivalente o denominada forma estándar se escribe como:

$$\left\{ \begin{array}{l} \text{ming}(x_1, x_2) = (-2e^{-2x_1^2 - x_1^2} + 2)z_1^k + (2(x_1 - 1)^2 + (x_2 - 1)^2)z_2^k + \\ \quad \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sujeto a:} \\ \quad 2x_1^2 + x_2^2 - 2(x_1^k)^2 - (x_2^k)^2 \leq 0 \\ \quad (x_1 - 1)^2 + (x_2 - 1)^2 - (x_1^k - 1)^2 - (x_2^k - 1)^2 \leq 0 \end{array} \right.$$

donde consideramos como distancias proximales $d(\cdot, \cdot)$, para la experimentación numérica de este ejemplo, a las siguientes distancias.

i.-Distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln\left(\frac{x_i}{y_i}\right) + y_i - x_i$$

ii.-Distancia proximal homogénea de segundo orden definida por

$$d(u, v) := \sum_{j=1}^2 v_j^2 \left(\frac{u_j}{v_j} - \ln\left(\frac{u_j}{v_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2, \text{ con } \sigma = 0.01$$

A continuación, presentamos las experimentaciones numéricas, tomando las distancias proximales mencionadas anteriormente, los parámetros fijos $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{10}}, \frac{\sqrt{9}}{\sqrt{10}}\right)$, $\lambda_k = 0.5, \lambda_k = 0.3$, además el punto $x^0 = (3, 2)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas usando la funcionalidad de MATLAB con el algoritmo SQP y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.00001$ para finalizar el algoritmo ISPPMR.

1.-Iteraciones numéricas del algoritmo usando la distancia Kullback-Leibler Bregman

Tabla 5.1. Resultados de la experimentación numérica con la distancia Kullback-Leibler Bregman

k	$N[x^k]$	λ_k	$\ x^{k+1} - x^k\ $	$x^k = (x_1^k, x_2^k)$	$F_1(x_1^k, x_2^k)$	$F_2(x_1^k, x_2^k)$	$\sum F_i(x^k)z^k$
1	8	0.5	2.07660	(1.113348 1.13235)	1.95349	0.04321	0.65874
2	11	0.5	0.18862	(0.98138 0.99757)	1.89228	0.00069	0.59905
3	12	0.5	0.01215	(0.98425 0.98576)	1.89096	0.00069	0.59863
4	13	0.5	0.00103	(0.98469 0.98482)	1.89095	0.00069	0.59863
5	14	0.5	0.00008	(0.98473 0.98472)	1.89095	0.00069	0.59863
6	14	0.5	0.00000	(0.98473 0.98473)	1.89095	0.00069	0.59863

En la Tabla 1 se muestra que con $k = 6$, iteraciones internas se aproxima a la solución del problema, por otro lado $N[x^k]$ representa el numero de iteraciones internas de cada subproblema, necesarias para obtener el punto (x_1^k, x_2^k) . Por ejemplo, para obtener el punto $x^4 = (0.9846, 0.9848)$, se necesita $N[x^4] = 13$ iteraciones internas. Asimismo, podemos observar en la tabla que la sucesión x^k converge hacia el punto $(1,1)$, la cual es una solución de Pareto como se muestra en la Figura 5.3. Por otro lado, en la Figura 5.4 se muestran las soluciones que forman la frontera de Pareto, en particular observamos la solución del problema alcanzada por el algoritmo planteado.

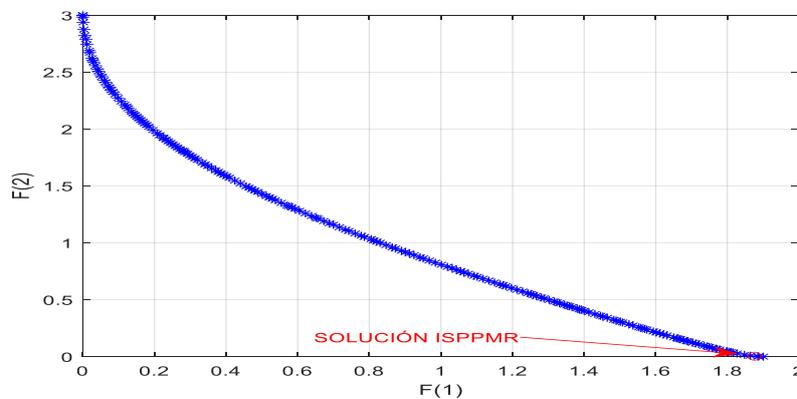


Figura 5.4. Frontera de Pareto ejemplo 5.1 con distancia Kullback-Leibler Bregman

2.- Iteraciones numéricas del algoritmo usando la distancia proximal homogénea de segundo orden.

Tabla 5.2. Resultados de la experimentación numérica con la distancia proximal homogénea de segundo orden.

k	$N[x^k]$	λ_k	$\ x^{k+1} - x^k\ $	$x^k = (x_1^k, x_2^k)$	$F_1(x_1^k, x_2^k)$	$F_2(x_1^k, x_2^k)$	$\sum F_i(x^k)z^i$
1	6	0.3	2.28688	(0.96220 0.96208)	1.87558	0.00429	0.59718
2	12	0.3	0.00009	(0.96216 0.96216)	1.87558	0.00429	0.59718
3	14	0.3	0.00000	(0.96216 0.96216)	1.87558	0.00429	0.59718

En la Tabla 2 las iteraciones usando la distancia proximal homogénea de segundo orden, se observa que con $k = 3$, iteraciones internas se aproxima a la solución del problema, con $\lambda_k = 0.3$, a diferencia de usar la distancia Kullback-Leibler Bregman que necesita $k = 6$ iteraciones internas, con $\lambda_k = 0.5$. Asimismo, podemos observar en la tabla que la sucesión x^k converge hacia una solución de Pareto (0.9621 0.9621), como el caso anterior, la cual es una solución de Pareto como se muestra en la Figura 5.3. Por otro lado, en la Figura 5.5 se muestran las soluciones que forman la frontera de Pareto, en particular observamos la solución del problema, alcanzada por el algoritmo planteado

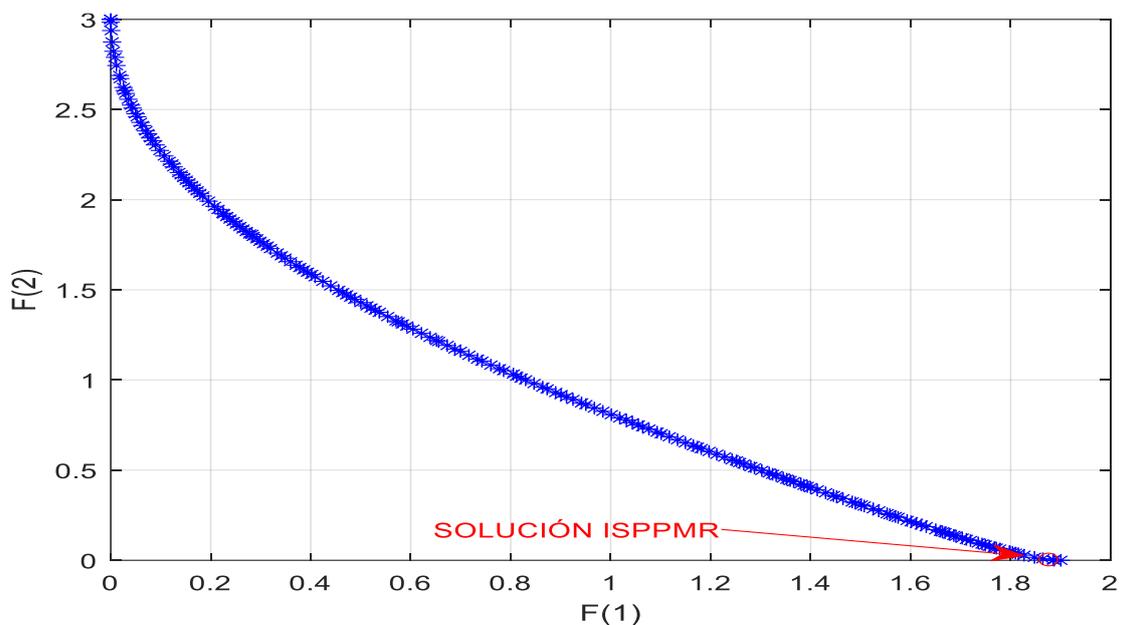


Figura 5.5. Frontera de Pareto Ejemplo 5.1 con distancia proximal homogénea de segundo orden

Ejemplo 5.2.- El siguiente problema de minimización multiobjetivo a diferencia del problema anterior está definido sobre un conjunto convexo cerrado, es decir

$$\min\{(F_1(x), F_2(x)) : x \in C \subset \mathbb{R}^2\}$$

donde $F_1(x_1, x_2) = -x_1^{1/3}x_2^{2/3}$, $F_2(x_1, x_2) = -x_1^{0.25}x_2^{0.75}$, y

$$C = \{(x_1, x_2) \in \mathbb{R}^2: -x_1 + x_2 \leq 2; 3x_1 + x_2 \leq 18, 3x_1 - x_2 \leq 12 \quad x_1 \geq 0, x_2 \geq 0\}$$

Las funciones F_1, F_2 son funciones diferenciables cuasi-convexas, por lo tanto $(F_1(\cdot), F_2(\cdot))^T$ función vectorial cuasi-convexa de clase $C^1(\mathbb{R}_+^2)$ (ver Teorema 2.2.1.23). Este problema satisface las hipótesis **(H1)**, **(H2)** y **(H3)**, igual que el problema del ejemplo 5.1. Además, mencionar que este problema solamente tiene al punto (4,6) como única solución de Pareto la cual se encuentra en el conjunto C , como se observa en la figura 5.6.

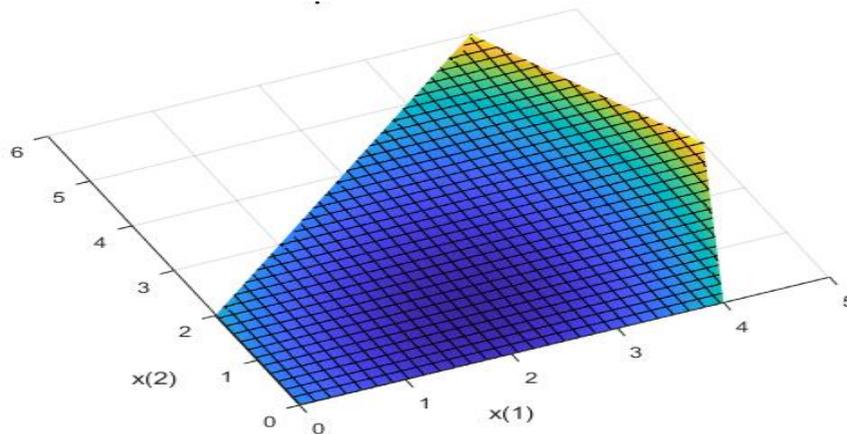


Figura 5.6. Conjunto de restricciones ejemplo 5.2

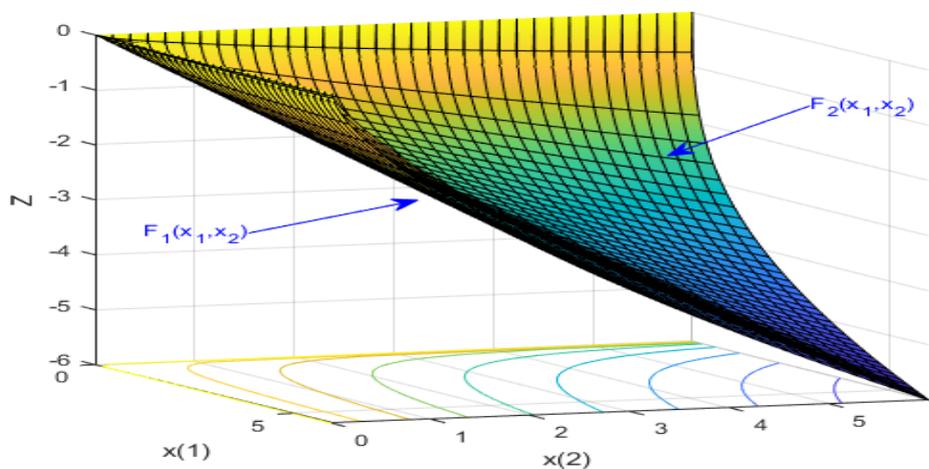


Figura 5.7. Funciones componentes objetivo ejemplo 5.2

Tomamos x^0 como punto inicial y dado $x^k \in C$, el paso principal de la propuesta del algoritmo como en el ejemplo 5.1, es encontrar un punto crítico (mínimo local, máximo local o un punto de silla) del siguiente problema.

$$\left\{ \begin{array}{l} \text{ming}(x_1, x_2) = (-x_1^{1/3} x_2^{2/3})_{z_1^k} + (-x_1^{0.25} x_2^{0.75})_{z_2^k} + \\ \lambda_k d \left(((x_1, x_2), (x_1^k, x_2^k)) \right) \\ \text{Sa:} \\ -x_1^{1/3} x_2^{2/3} + (x_1^k)^{1/3} (x_2^k)^{2/3} \leq 0 \\ -x_1^{0.25} x_2^{0.75} + (x_1^k)^{0.25} (x_2^k)^{0.75} \leq 0 \end{array} \right.$$

donde consideramos como distancia proximal $d(\cdot, \cdot)$, para la experimentación numérica de este ejemplo, a la distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln \left(\frac{x_i}{y_i} \right) + y_i - x_i$$

A continuación, presentamos las experimentaciones numéricas, tomando la distancia proximal mencionada anteriormente, y haciendo una variación de los parámetros fijos.

1.-Sea $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$, $\lambda_k = 0.2$, el punto $x^0 = (1, 1)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas usando la funcionalidad de MATLAB con el algoritmo SQP y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.0000001$ para finalizar el algoritmo ISPPMR.

Tabla 5.3. Resultados de la experimentación numérica con la distancia Kullback-Leibler Bregman

k	$N[x^k]$	λ_k	$\ x^{k+1} - x^k\ $	$x^k = (x_1^k, x_2^k)$	$F_1(x_1^k, x_2^k)$	$F_2(x_1^k, x_2^k)$	$\sum F_i(x^k) z_i^k$
1	9	0.2	4.03477	(2.85303 4.58408)	-3.91384	-4.07160	-5.64657
2	11	0.2	1.46093	(3.75158 5.73600)	-4.97902	-5.15835	-7.16821
3	14	0.2	0.29987	(3.95622 5.95520)	-5.19626	-5.37641	-7.47601
4	14	0.2	0.05220	(3.99266 5.99258)	-5.23396	-5.41410	-7.52931
5	18	0.2	0.00869	(3.99877 5.99877)	-5.24023	-5.42036	-7.53818
6	16	0.2	0.00142	(3.99978 5.99978)	-5.24126	-5.42139	-7.53963
7	11	0.2	0.00025	(3.99996 5.99996)	-5.24144	-5.42157	-7.53989
8	14	0.2	0.00003	(3.99999 5.99999)	-5.24147	-5.42160	-7.53993
9	14	0.2	0.00000	(3.99999 5.99999)	-5.24148	-5.42161	-7.53994
10	15	0.2	0.00000	(3.99999 5.99999)	-5.24147	-5.42160	-7.53993

En la Tabla 5.3 se muestra que con $k = 10$, iteraciones se aproxima a la solución del problema. Asimismo, podemos observar en la tabla que la sucesión x^k converge hacia el punto $(4,6)$, la cual es la única solución de Pareto, implicando que el problema tenga una única solución en la frontera de Pareto, determinada por el algoritmo propuesto, además indicar que la Figura 5.8 nos muestra la diferencia de dos pasos consecutivos que van convergiendo a cero, de acuerdo a las iteraciones.

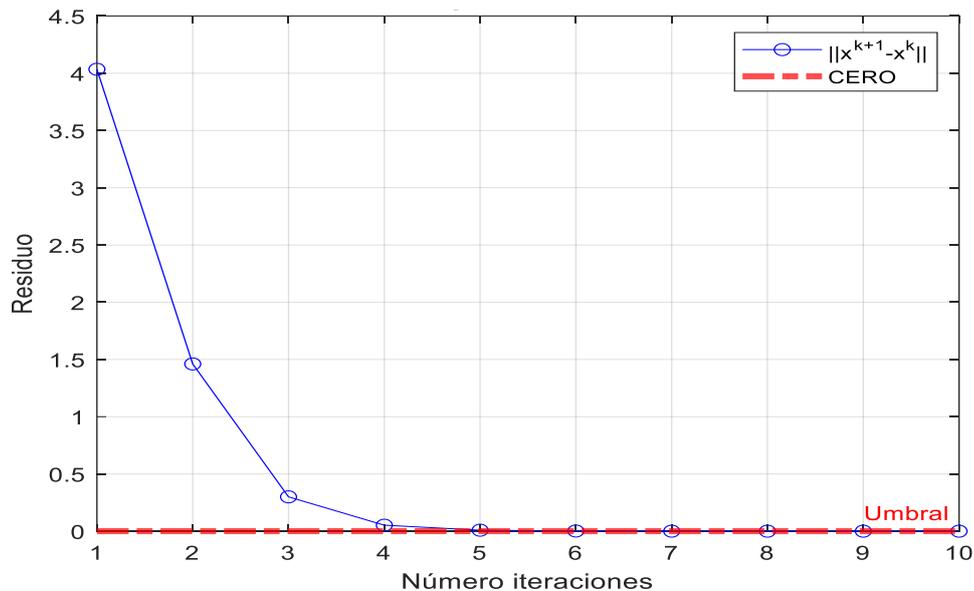


Figura 5.8. Distancias de los puntos consecutivos ejemplo 5.2, con $\lambda_k = 0.2$

2.-Sea $z_k = (z^k_1, z^k_2) = \left(\frac{1}{\sqrt{3}}, \sqrt{\frac{2}{3}}\right)$, $\lambda_k = 0.4$, el punto $x^0 = (1, 2)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas usando la funcionalidad de MATLAB con el algoritmo SQP y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.000001$ para finalizar el algoritmo ISPPMR.

Tabla 5.4. Resultados de la experimentación numérica con la distancia Kullback-Leibler Bregman

k	$N[x]$	λ_k	$\ x^{k+1} - x^k\ $	$x^k = (x_1^k, x_2^k)$	$F_1(x_1^k, x_2^k)$	$F_2(x_1^k, x_2^k)$	$\sum F_i(x^k)$
1	6	0.4	2.07462	(2.08783 3.76654)	-3.09405	-3.24998	-4.43995
2	9	0.4	1.44541	(2.98041 4.90343)	-4.15360	-4.32956	-5.93316
3	13	0.4	0.79883	(3.51488 5.49713)	-4.73581	-4.91563	-6.74782
4	12	0.4	0.39152	(3.78491 5.78064)	-5.01960	-5.19991	-7.14378
5	19	0.4	0.17674	(3.90828 5.90720)	-5.14734	-5.32762	-7.32180
6	12	0.4	0.07600	(3.96162 5.96134)	-5.20220	-5.38240	-7.39821
7	11	0.4	0.03189	(3.98407 5.98400)	-5.22521	-5.40537	-7.43024
8	18	0.4	0.01324	(3.99341 5.99339)	-5.23475	-5.41490	-7.44353
9	16	0.4	0.00546	(3.99727 5.99726)	-5.23869	-5.41883	-7.44902
10	15	0.4	0.00225	(3.99886 5.99886)	-5.24032	-5.42045	-7.45128
11	17	0.4	0.00092	(3.99952 5.99951)	-5.24099	-5.42112	-7.45221
12	12	0.4	0.00049	(3.99987 5.99987)	-5.24135	-5.42148	-7.45271
13	11	0.4	0.00009	(3.99994 5.99993)	-5.24142	-5.42154	-7.45281
14	14	0.4	0.00004	(3.99997 5.99997)	-5.24145	-5.42158	-7.45285
15	11	0.4	0.00002	(3.99999 5.99999)	-5.24147	-5.42160	-7.45288
16	16	0.4	0.00000	(3.99999 5.99998)	-5.24147	-5.42160	-7.45288

En la Tabla 5.4 se muestra que con $k = 16$, iteraciones internas se aproxima a la solución del problema, habiendo una diferencia con respecto a la primera experimentación numérica en la cual solamente fueron $k = 10$ iteraciones internas. Asimismo, podemos indicar que la convergencia de la sucesión y la solución en la frontera de Pareto, es análogo que la primera experimentación numérica, salvo el comportamiento de la diferencia de dos pasos consecutivos que converge a cero, tiene más iteraciones como se apreciar en la Figura 5.9.

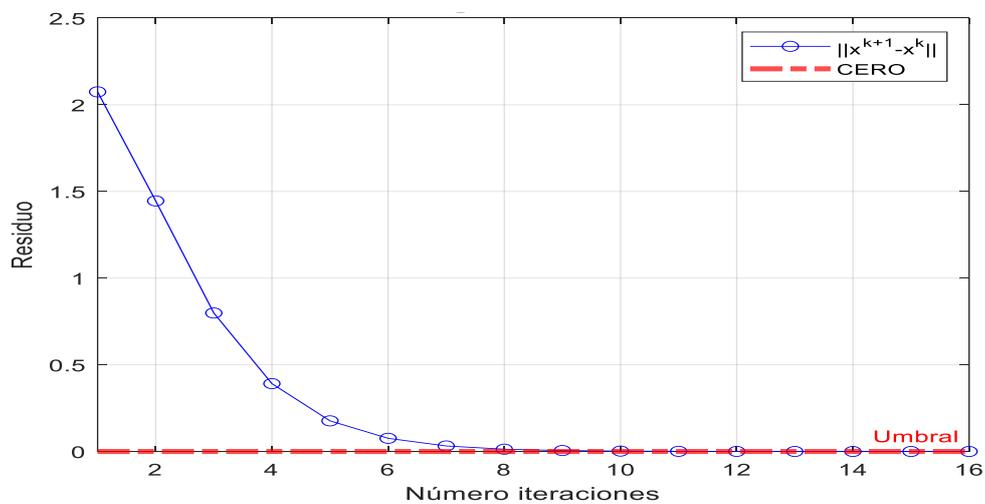


Figura 5.9. Distancias de los puntos consecutivos ejemplo 5.2, con $\lambda_k = 0.4$

CAPITULO VI

APLICACION A LA ECONOMIA

En la teoría económica siempre está en la búsqueda de optimizar los recursos, por ejemplo en el campo de la producción es necesario utilizar modelos matemáticos para optimizar la producción (ver [44]) una de las funciones que es usada ampliamente en esta área es la función Cobb-Douglas propuesta inicialmente por Knut Wicksell (1851-1926), luego Charles Coob y Paul Douglas en 1928 evidenciaron estudios estadísticos concretos, consideraron que la función de producción descrita por $Q = AL^\alpha K^\beta$ estaba vinculada al factor trabajo denotado por L , al factor de capital denotado por K , y al factor total de la producción denotado por A , donde $\alpha, \beta \in \langle 0,1 \rangle$ y $A > 0$. A continuación, mostraremos el siguiente ejemplo, donde aplicaremos el método propuesto.

La empresa RENACER O&Y S.A.C produce un bien. La función de producción de un bien Q es dado por la función, $Q = f(K, L) = 80K^{1/4}L^{1/2}$ y su precio de venta es $P = 40$ dólares. Los precios de los factores capital y trabajo son 80 y 40 respectivamente. Calcular los niveles de capital y trabajo, con la cual maximicen la utilidad y minimicen los costos de la empresa, considerando que la empresa tiene un monto de inversión de 10240 dólares.

Para formular el modelo matemático se tiene que definir las funciones utilidad y costo de la empresa, en efecto, la función ingreso $I(K, L) = PQ = (40)80K^{1/4}L^{1/2}$, la función costo $C(K, L) = 80K + 40L$ y la función utilidad $U(K, L) = (40)80K^{1/4}L^{1/2} - 80K - 40L$ y además considerando las restricciones $80K + 40L \leq 10240, K > 0, L > 0$.

Por lo tanto, el problema a optimizar este dado por:

$$\min\{-U(K, L), C(K, L)\}$$

$$\text{S.a } 80K + 40L \leq 10240$$

$$K > 0, L > 0$$

donde $U(K, L) = (40)80K^{1/4}L^{1/2} - 80K - 40L$, $C(K, L) = 80K + 40L$

$$C = \{(K, L) \in \mathbb{R}^2: 80K + 40L \leq 10240; K > 0, L > 0\}$$

La función U es una función diferenciable cuasiconvexa y C es una función cuasiconvexa suave, por lo tanto $(U(\cdot), C(\cdot))^T$ función vectorial cuasi-convexa suave en C (ver Teorema 2.2.1.23). Este problema satisface las hipótesis **(H1)**, **(H2)** y **(H3)**, igual que los ejemplos del capítulo anterior. Además, se puede observar en la Figura 6.2 que el punto (70,70) es una solución de Pareto.

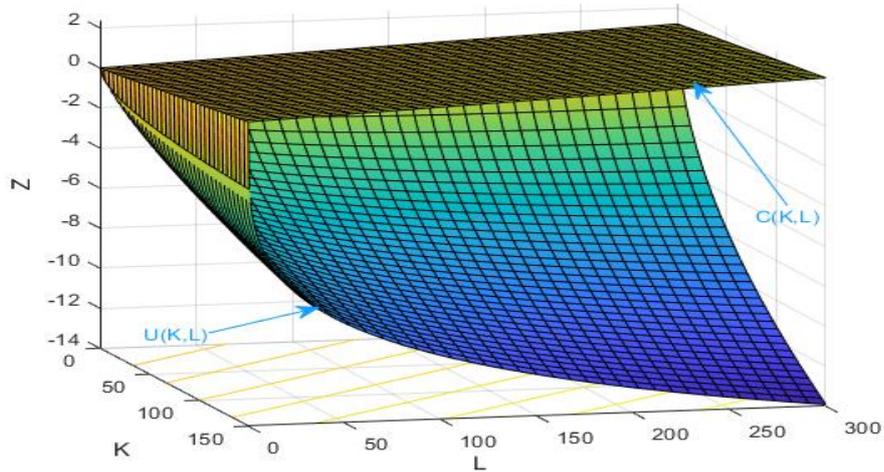


Figura 6.1. Funciones componentes objetivo utilidad y costo

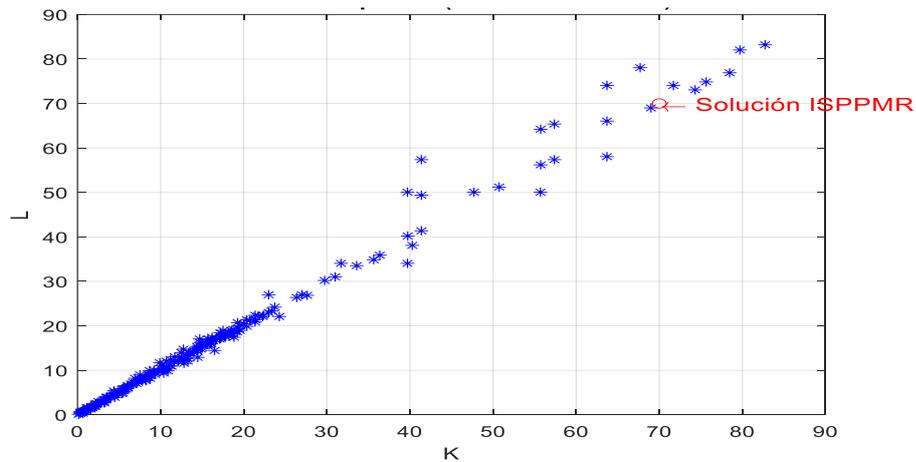


Figura 6.2. Soluciones de Pareto en el conjunto C

Ahora tomando como punto inicial $x^0 = (x_1^0, x_2^0)$ y dado $x^k \in C$, el paso principal de la propuesta del algoritmo es encontrar un punto crítico (mínimo local) del siguiente problema.

$$\left\{ \begin{array}{l} \text{ming}(K, L) = (3200K^{1/4}L^{1/2} - 80K - 40L)z_1^k + (80K + 40L)z_2^k + \\ \quad \lambda_k d\left(\left((K, L), (K^k, L^k)\right)\right) \\ \text{Sa:} \\ 3200K^{1/4}L^{1/2} - 80K - 40L \leq 3200(K^k)^{1/4}(L^k)^{1/2} - 80K^k - 40L^k \\ 80K + 40L \leq 80K^k + 40L^k \end{array} \right.$$

considerando los parámetros fijos $z_k = (z_1^k, z_2^k)$, λ_k para cada k . Además, haciendo $(K, L) = (x_1, x_2)$, se tiene de forma equivalente o denominada forma estándar el problema descrito como:

$$\begin{cases} \text{ming}(x_1, x_2) = (3200x_1^{1/4}x_2^{1/2} - 80x_1 - 40x_2)z_1^k + (80x_1 + 40x_2)z_2^k + \\ \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sa:} \\ 3200x_1^{1/4}x_2^{1/2} - 80x_1 - 40x_2 - 3200(x_1^k)^{1/4}(x_2^k)^{1/2} + 80x_1^k + 40x_2^k \leq 0 \\ 80x_1 + 40x_2 - 80x_1^k - 40x_2^k \leq 0 \end{cases}$$

donde consideramos como distancia proximal $d(\cdot, \cdot)$, para la experimentación numérica de este ejemplo, a la distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln\left(\frac{x_i}{y_i}\right) + y_i - x_i$$

A continuación, presentamos las experimentaciones numéricas, tomando la distancia proximal mencionada anteriormente, los parámetros fijos $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\lambda_k = 0.1$, además el punto $x^0 = (100, 10)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (10, 10)$ para resolver todos los subproblemas usando la funcionalidad de MATLAB con el algoritmo SQP y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.000001$ para finalizar el algoritmo ISPPMR.

Tabla 6.1. *Resultados de la experimentación numérica con distancia Kullback-Leibler Bregman*

k	$N[x^k]$	λ_k	$\ x^{k+1} - x^k\ $	$x^k = (x_1^k, x_2^k)$	$U(x_1^k, x_2^k)$	$C(x_1^k, x_2^k)$	$\sum F_i(x^k)z_i^k$
1	6	0.1	67.02546	(70.02529, 69.94940)	-69041.44817	8400.00000	- 42879.97922
2	6	0.1	0.056600	(69.99998, 70.00002)	-69041.45575	8400.00000	- 42879.98458
3	8	0.1	0.00006	(70.00001, 69.99996)	-69041.45575	8400.00000	- 42879.98458
4	6	0.1	0.00009	(69.99997, 70.00005)	-69041.45575	8400.00000	- 42879.98458
5	8	0.1	0.00007	(70.00000, 69.99998)	-69041.45575	8400.00000	- 42879.98458
6	6	0.1	0.00007	(69.99997, 70.00005)	-69041.45575	8400.00000	- 42879.98458
7	6	0.1	0.00000	(69.99997, 70.00005)	-69041.45575	8400.00000	- 42879.98458

8	6	0.1	0.00000	(69.99997 70.00005)	-69041.45575	8400.00000	-42879.98458
9	6	0.1	0.00000	(69.99997 70.00005)	-69041.45575	8400.00000	-42879.98458
10	6	0.1	0.00000	(69.99997 70.00005)	-69041.45576	8400.00000	-42879.98458
11	6	0.1	0.00000	(69.99997 70.00005)	-69041.45576	8400.00000	-42879.98458
12	6	0.1	0.00000	(69.99997 70.00005)	-69041.45575	8400.00000	-42879.98458

En la Tabla 6.1 se muestra que con $k = 12$, iteraciones internas se aproxima a la solución del problema. Asimismo, podemos observar en la tabla que la sucesión x^k converge hacia el punto $(70,70)$, la cual es una solución de Pareto como se indicó en la figura 6.2. Por otro lado, en la figura 6.3 se muestran las soluciones que forman la frontera de Pareto, en particular la solución del problema determinada por el algoritmo propuesto.

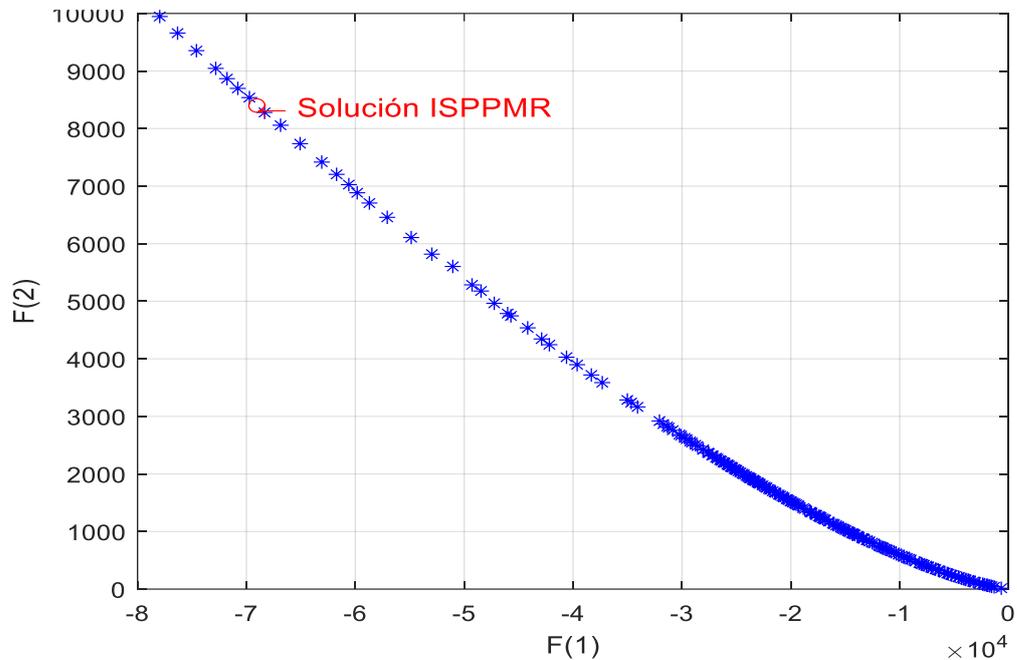


Figura 6.3. Frontera de Pareto función utilidad y costo

Además, además indicar que la Figura 6.4 nos muestra la diferencia de dos pasos consecutivos que van convergiendo a cero, de acuerdo a las iteraciones.

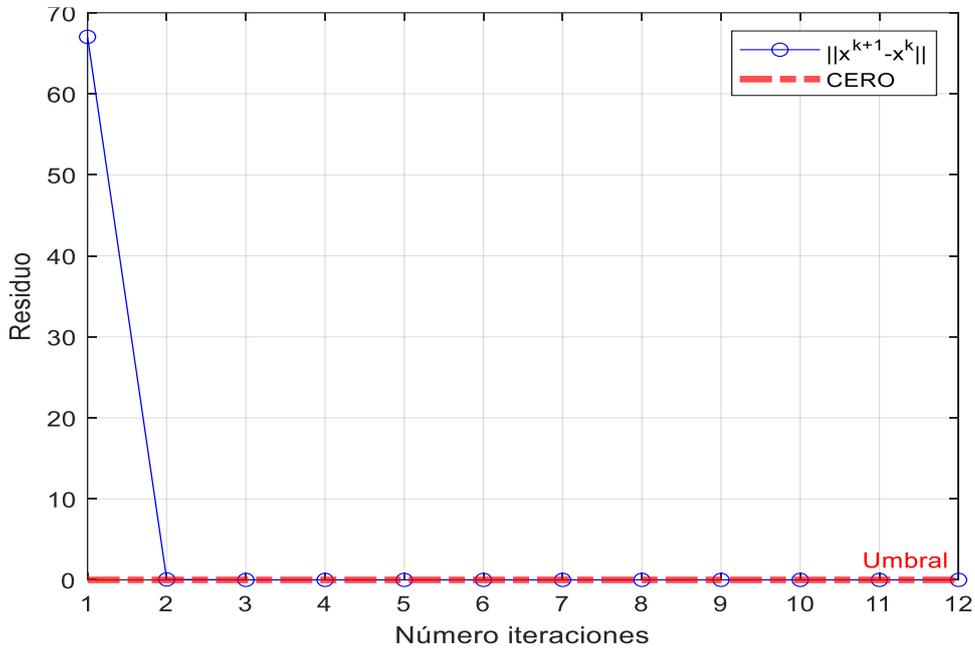


Figura 6.4. Distancias de los puntos consecutivos

Finalmente podemos concluir que la empresa RENACER O&Y S.A.C al producir el bien tendrá una utilidad máxima aproximada de 69 041 dólares y un costo mínimo aproximado de 8400 dólares, cuando los precios de los factores capital y trabajo sean 69,99 y 70 respectivamente.

CAPÍTULO VII

MARCO METODOLÓGICO

7.1 Hipótesis central de la investigación

7.1.1. Hipótesis general

Usando la metodología de optimización continua se demuestra la extensión del Método Proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones y su aplicación a la economía.

i). -Considerando las siguientes condiciones para el error de aproximación e^{k+1}

$$\frac{\|e^{k+1}\|}{\lambda_k} \leq \eta_k \sqrt{H(x^{k+1}, x^k)}$$

$$\sum_{k=0}^{+\infty} \eta_k < +\infty$$

el algoritmo ISPPM para minimización multiobjetivo cuasi-convexa, puede ser extendido a la siguiente versión inexacta

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), Z_k \rangle + \lambda_k d(\cdot, x^k) \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1})$$

donde ∂^0 es el subdiferencial de Clarke, $\Omega_k = \{x \in \mathbb{R}^n : F(x) \leq F(x^k)\}$, $\lambda_k > 0$, $\{Z_k\} \subset \mathbb{R}_+^n$, $\|Z_k\| = 1$ y $N_{\Omega_k}(x^{k+1})$ es el cono normal a Ω_k , en x^{k+1} $d(\cdot, \cdot)$ es una distancia proximal sobre C y e^{k+1} es el error de aproximación.

7.1.2. Hipótesis específicas

1. Se demuestra la introducción de un algoritmo de punto proximal para Optimización Multiobjetivo Cuasi-convexa con restricciones en espacios euclidianos.

Si el error de aproximación e^{k+1} satisface las condiciones indicadas en la hipótesis general, por lo tanto, el ISPPMR está bien definido en espacios Euclidianos.

2. Se prueba y analiza que la sucesión generada por el algoritmo converge a un punto crítico o a un mínimo.

Usando la técnica de convergencia de sucesiones se obtiene que sucesión generada por algoritmo converge a un punto crítico.

3.- Se analiza y demuestra la velocidad de convergencia del algoritmo.

El algoritmo ISPPMR inexacto presenta una velocidad de convergencia lineal y superlineal.

4. Se aplica el algoritmo a la solución de un problema en economía.

El algoritmo ISPPMR inexacto puede ser usado para resolver problemas en economía.

7.2 Variables e indicadores de la investigación

Dado que la presente investigación es netamente teórica, implica que no se presenten indicadores, solamente se describirán las variables de estudio, las variables son las siguientes:

Variable independiente: Hipótesis (H1), (H2), y (H3)

Variable dependiente: Convergencia del algoritmo

7.3 Método de la investigación

En la presente investigación usamos la metodología de Optimización matemática, en la cual consiste en dar una buena definición del algoritmo del método de proximal, que la sucesión de puntos generada por el algoritmo exista y además que la sucesión converja a una solución del problema

7.4 Diseño o esquema de la investigación

La presente investigación del trabajo de tesis es de tipo no experimental debido a que no se hace ninguna manipulación de las variables independientes, donde su diseño es descriptivo correlacional pues su propósito es indagar la convergencia de valores que presentan las propiedades del método proximal de optimización.

7.5 Población y muestra

Por ser nuestro trabajo teórico, no existe población que estudiar, sin embargo, nuestro estudio se encuentra inmerso dentro de \mathbb{R}^n y optimización multiobjetivo cuasi-convexa.

7.6 Actividades del proceso investigativo

presente trabajo de investigación cuyo avance fue elaborado siguiendo las pautas de la Escuela de Pos-Grado de la Universidad Nacional del Santa y coordinación con mi asesor, el resumen de actividades se muestra en la siguiente tabla:

Actividad	Fecha
Revisión de artículos y tesis del tema de investigación	Febrero, marzo de 2021
Formulación del problema de investigación	Abril de 2021
Formulación de la hipótesis de investigación	Mayo de 2021
Redacción del estado de arte del tema de investigación	Junio 2021
Revisión de metodologías referidas al tipo de investigación	Julio de 2021
Elección de la metodología de investigación	Agosto de 2021
Descripción de la metodología para la contratación de hipótesis	Setiembre de 2021
Recolección de datos de la investigación (proposiciones y teoremas)	Octubre de 2021
Demostración de las proposiciones y teoremas	Noviembre de 2021
Análisis e interpretación de los teoremas	Diciembre de 2021
Contrastación de los resultados obtenidos	Enero, febrero, marzo de 2022
Redacción del trabajo de investigación	Abril, mayo, junio y julio de 2022

7.7 Técnicas e instrumentos de la investigación

Para la realización de este trabajo de investigación se revisó bibliografía especializada, libros, artículos científicos, página web, tesis y además de la asistencia a foros, eventos, así como la comunicación de investigadores referidos al tema.

7.8 Técnicas de procesamiento y análisis de los datos.

Por la naturaleza de la investigación no se realiza ningún análisis estadístico

CAPÍTULO VIII

RESULTADOS Y CONCLUSIONES

8.1. RESULTADOS

El aporte del siguiente trabajo de investigación es de carácter teórico y aplicativo.

Los resultados de convergencia obtenidos por el algoritmo ISPPMR en el Capítulo 3 en un subconjunto del espacio Euclidiano son novedosos, cabe indicar que aún no han sido publicados por ninguna revista nacional e internacional. Los resultados más relevantes son los siguientes.

- 1.-Se ha revisado resultados de convergencia de algunos métodos de punto proximal escalarizados para minimización mono y multiobjetivo de funciones cuasi-convexas, definidos en el espacio Euclidiano y se adaptado a su versión multiobjetivo con restricciones definida en el espacio Euclidean.
- 2.-Considerando las hipótesis necesarias se ha probado que cualquier punto de acumulación de la sucesión $\{x^k\}$, generada por el algoritmo ISPPMR, es un punto crítico Pareto-Clarke del problema de optimización multiobjetivo (1).
- 3.-Se presentan resultados de experimentación numérica mediante dos ejemplos, la cual se ha usado distintas distancias proximales y diferentes valores para los parámetros fijos.
- 4.-Se aplica el algoritmo propuesto a un problema en economía.

8.2. DISCUSIÓN

- 1.-El método presentado en este trabajo de investigación puede ser considerado como un avance importante, en los métodos proximales para optimización multiobjetivo con restricciones en el espacio Euclidiano, la cual modelan de una forma más aproximada los problemas que se presentan en el contexto real.

2.-Los problemas que se modelan en la teoría económica donde se utiliza funciones cuasi-convexas presentan una gran complejidad, donde son difíciles de resolver usando las técnicas de optimización conocidas, por ello la presente tesis es una contribución para la obtención de algoritmos de aproximación más eficientes.

3.-Apartir de considerar las hipótesis necesarias, cualquier punto de acumulación de una sucesión generada por el algoritmo ISPPMR converge a un punto crítico Pareto-Clarke del problema multiobjetivo cuasi-convexo con restricciones. Estos resultados guardan relación con Cruzado Acuña (2018), donde presento un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexo, donde prueba que las sucesiones generadas por el algoritmo convergen a un punto crítico Pareto-Clarke, pero para problemas de minimización multiobjetivo irrestrictos, además cabe mencionar que el autor utilizo la distancia euclidiana para regularizar la función objetivo y en cambio en la presente investigación se utilizó distancias proximales para regularizar la función objetivo.

CAPÍTULO IX

CONCLUSIONES Y RECOMENDACIONES

9.1.- Conclusiones

1.- En esta tesis presentamos la introducción de un método proximal para optimización multiobjetivo cuasi-convexa con restricciones y su aplicación a la economía definido en el espacio Euclidiano, utilizando distancias proximales.

2.-Demostramos algunos resultados teóricos de convergencia, considerando condiciones para el error de aproximación.

3.-El algoritmo puede utilizarse para resolver modelos matemáticos en la teoría de la producción en economía.

9.2.- Recomendaciones

Un futuro trabajo puede ser la implementación computacional del método estudiado para algunos problemas del contexto real en la teoría de la decisión y establecer una comparación con otros métodos.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Apolinario, H. C. F., Papa Quiroz, E., and Oliveira, P. R. (2016). *A scalarization proximal point method for quasiconvex multiobjective minimization*. *Journal of Global Optimization* 64, 79–96.
- [2] Arsenio Pecha C. (2005). *Optimización estática y dinámica en economía*, Universidad Nacional de Colombia, Facultad de Ciencias.
- [3] Ávila Acosta, R. (1992) *Introducción a la Metodología de la Investigación*. Tesis. Estudios y Ediciones R.A. Lima.
- [4] Barbolla R, Cerdá E, Sanz, P (2001). *Optimización: Cuestiones, Ejercicios y aplicaciones a la economía*, Pearson Educación, S. A., Madrid.
- [5] Baygorrea N, Papa Quiroz EA, Maculan N. (2017). *On the Convergence Rate of an Inexact Proximal Point Algorithm for Quasiconvex Minimization on Hadamard Manifolds*; *Journal of the Operations Research Society of China*, Volume 5 (4):457-467.
- [6] Baygorrea N, Papa Quiroz EA, Maculan N (2016). *Inexact Proximal Point Methods for Quasiconvex Minimization on Hadamard Manifolds*; *Journal of the Operations Research Society of China*, 4 (4): 397-424.
- [7] Bazaraa, M.S., Sherali, H.D. e Shetty, C.M. (2006). *Nonlinear Programming: Theory and Algorithms*. 3 ed., John Wiley and Sons, Inc., New York.
- [8] Benazic, R. T. (2000). *Topología en espacios euclidianos*. Sociedad Matemática Peruana Second edition, XVIII Coloquio.
- [9] Benazic, R. T. (2001) *Homeomorfismos en espacios euclidianos*. Sociedad Matemática Peruana, XIX Coloquio.
- [10] Bento G, Cruz Neto J, Lopez G, Soubeyran A, Souza J. (2018). *The proximal point method for locally Lipschitz functions in multiobjective optimization with application to the compromise problem*. *Society for Industrial and Applied Mathematics*.
- [11] Bento G, Cruz Neto J, Lopez G, Soubeyran A, Souza J. (2020). *A proximal point method for difference of convex functions in multi-objective optimization with application to group dynamic problems*. 263–290.
<https://doi.org/10.1007/s10589-019-00139-0>

- [12] Bermeo Carrasco, O.A. (2010). *Programacion multiobjetivo conveva*. Universidad Nacional del Callao Lima-Peru.
- [13] Berkovitz L, (2003). *Convexity and optimization in R^n* ; Ney York, John Wiley & Sons.
- [14] Bonfin de Assunção, P. (2015). *Um Algoritmo Proximal com Quase-distância*. Universidad federal de Goiás-Brasil.
- [15] Bolintineanu, S. (2001). *Vector variational principles; e-efficiency and scalar stationarity*. Journal of Convex Analysis 8 71–85.
- [16] Bolte, J., Daniilidis, A., Lewis, A., and Shiota, M. (2007). *Clarke subgradients of stratifiable functions*. SIAM Journal on Optimization 18, 2, 556–572.
- [17] Bonnel H., Iusem A.N. Svaiter B. F. (2019). *Metodos proximales en optimización vectorial*. Artículo Universite de la Nouvelle-calédonie.
- [18] Bonnel, H., Iusem, AN, Svaiter, BF. (2005). *Métodos proximales en la optimización vectorial*. SIAM J. Optim. 15 (4), 953–970.
- [19] Borda Maticorena D.C. (2010). *convergencia del método de punto proximal con distancia homogénea de orden r en optimización convexa*. Universidad Nacional del Callao Lima-Peru.
- [20] Ceng, L., Yao, J. (2007). *Approximate proximal methods in vector optimization*. Eur. J. Oper. Res. 183, 1–19.
- [21] Clarke F (1990). *Optimization and nonsmooth analysis*, Philadelphia, SIAM.
- [22] Clarke F (2013). *Functional analysis, calculus of variations and optimal control*, New York, Springer Science & Business Media.
- [23] Coello, C. (2005): “*Introducción a la Optimización Multiobjetivo usando Metaheurísticas*”, Sección de Computación.
- [24] Cruzado Acuña, C. (2019). *Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa*. Universidad Nacional de san marcos- Lima-Peru.
- [25] Cunha, F. G. M., da Cruz Neto, J. X., Oliveira, P. R. (2010). *A proximal point algorithm with φ -divergence to quasiconvex programming*, 59: 777-792.
- [26] Chen, J. S., Pan, S. (2008). *A proximal-like algorithm for a class of nonconvex programming*. Pacific Journal of Optimization, 4: 319-333.

- [27] Chen, G., Teboulle, M. (1993). *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, 3: 538-543.
- [28] Chuan L. , Chih J. (2007). *Approximate proximal methods in vector optimization*. European Journal of Operational Research 183 1–19.
- [29] Ehrgott M. (2005). *Multicriteria optimización*. Springer, Verlag verlin.
- [30] Fernandez Apolinario, H. (2014). *Método de ponto proximal para minimização multiobjetivo quase-convexa*. Tesis Dctoral PESC-COPP Universidad Federal do Rio de Janeiro, Brasil.
- [31] Francis Ysidro Edgeworth, (1881). *Mathematical Psychics*, McMaster University Archive for the History of Economic Thought.
- [32] Greenberg, H.J., Pierskalla, W.P. (1973). *Quasi-conjugate functions and surrogate duality*. Cah. Cent. Étud. Rech. Opér. 15, 437–448.
- [33] Guler, O. (1992). *New proximal point proximal algorithms for convex minimization*. SIAM Journal on Optimization, 2, 649–664
- [34] Hernandez R; Fernandez C; Bauatista P. (2010). *Metodología de la Investigación*. Quinta edición México.
- [35] Hu, Y., Yang, X., & Sim, C. K. (2015). *Inexact subgradient methods for quasi-convex optimization problems*. *European Journal of Operational Research*, 240(2), 315–327. <https://doi.org/10.1016/j.ejor.2014.05.017>.
- [36] Iusem, A. N., Sosa, W. (2010). *A proximal point method for equilibrium problems in Hilbert spaces*. Optimization, 59: 1259-1274.
- [37] Jahn J. (2004). *Vector Optimization*. Theory, applications and extensions. Springer, Berlin.
- [38] Trullo, L. Romero, M. (2005) *Microeconomía, Problemas y ejercicios*.
- [39] L. M. Graña Drummond Y Ellen H. Fukuda. (2009). *Convergencia del método de gradiente proyectado para optimización vectorial*. Universidade Federal do Rio de Janeiro, Sao Paulo Brasil.
- [40] Luc, D.T. (1989). *Theory of vector optimization, Lecture Notes in Economics and Mathematical Systems*. Springer, Berlin.
- [41] Kin Keung Lai¹, Shashi Kant Mishra, Geetanjali Panda, Md Abu Talhamainuddin Ansary, y Bhagwat Ram. (2020). *En el método de descenso*

más pronunciado para problemas optimización multiobjetivo sin restricciones, J. Appl. Math. Comput.

- [42] Kenneth H, Rosen. (2004). *Matemática discreta y sus aplicaciones*, quinta edición, España.
- [43] Klimovsky, G. (2001). *Las desventuras del conocimiento científico*. A-Z editora S.A. Argentina.
- [44] Kunh, H.W., A.W. Toker: "*Nonlinear Programing*. In J. Neyman (ed). (1951). *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley.
- [45] Mendez M, (2008). *Algoritmos Evolutivos y Preferencias del Decisor Aplicados a Problemas de Optimización Multiobjetivo Discretos*, Tesis Doctoral, Universidad de Las Palmas de Gran Canaria
- [46] Madden P. (1987). *Concavidad y optimización en Microeconomía*. Madrid editorial Alianza.
- [47] Maino M., Pittet J., Kodrich C. (1993). *Programación multicriterio; un instrumento para el diseño de producción* Santiago de Chile.
- [48] Mallma Ramirez, L., Papa Quiroz. E. A., Oliveira, P. R. (2017). *An Inexact Proximal Method with Proximal Distances for Quasimonotone Equilibrium Problems*. Journal of the Operations Research Society of China 5 (4), pp 545-561.
- [49] Mangasarian, O.L. (1969). *Nonlinear Programming*. McGraw-Hill, New York.
- [50] Martinet, B. (1970). *Regularisation d'inéquations variationnelles par approximations successives*, Revue Française Informatique Recherche Opérationnelle, 4, 154-158.
- [51] Martinet, B. (1970). *Breve communication. regularization inequations variationally par approximations successive*. Revue française d'informatique et de recherche operationnelle. Serie rouge 4, R3, 154–158.
- [52] Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford University Press, New York, NY, USA.

- [53] Materano C. (2015). *Funciones Convexas, Cuasi-Convexas, Lipschitz y la Desigualdad de Simpson*. Universidad Central de Venezuela Facultad de Ciencias,.
- [54] Miettinen, K.M. (2012). *Nolinear multiobjective optimization*. Volumen.12. Springer Science & Business Media.
- [55] Mordukhovich, B. S. (2006). *Variational analysis and generalized differentiation I: Basic theory*, vol. 330. Springer Science & Business Media.
- [56] Nava, Rafael. (2015). *Funciones convexas no diferenciables*, Mexico.
- [57] Papa Quiroz, E.A., Mallma Ramirez, L., Oliveira, P.R. (2015). *An inexact proximal method for quasiconvex minimization*. Eur. J. Oper. Res. 246, 721–729.
- [58] Papa Quiroz E., Borda D., Collantes F. (2020). *Un método Proximal para minimización multiobjetivo cuasiconvexo en el ortante no negativo y su aplicación a la teoría de la demanda en microeconomía*. UPN, UNSM, Lima.
- [59] Papa Quiroz, E. Cruzado S. (2019). *Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa en espacios Euclidianos*. PESQUIMAT 22(1): 31–50 UNSM, Lima-Peru.
- [60] Papa Quiroz, E. (2017). *Optimización continua: teoría y ejercicios*. Lima, Perú: Editorial Académica Española.
- [61] Papa Quiroz, E., Cruzado Acuña, S. (2021). *Linear and Superlinear Convergence of an Inexact Algorithm with Proximal Distances for Variational Inequality Problems*. Privada del Norte, U., & Nacional Mayor de San Marcos, U.
- [62] Papa Quiroz, & Cruzado, E. A. (n.d.). *An Inexact Scalarization Proximal Point Method for Multiobjective Quasiconvex Minimization*.
- [63] Papa Quiroz, E. A., Oliveira, P. R. (2012). *An extension of proximal methods for quasiconvex minimization on the nonnegative orthant*. European Journal of Operational Research 216: 26-32.
- [64] Papa Quiroz, E. A., Mallma Ramirez, L., Oliveira, P. R. (2015). *An inexact proximal method for quasiconvex minimization*. European Journal of Operational Research 246 (3), pp 721-729.

- [65] Polyak BT (1987). *Introduction to Optimization*. Optimization Software, New York.
- [66] Robledo J. (2020). *Análisis Convexo en la Economía y su relación con la Programación no Lineal*. Universidad Nacional San Martín.
- [67] Rockefeller, R. (1976). *Monotone operators and the proximal point algorithm*. SIAM Journal on Optimization, 14, 877–898.
- [68] Rockefeller, R. (1998). *convex Analysis*, Princeton University.
- [69] Romero C. (1993) *Teoría de la Decisión Multicriteria*. Alianza Editorial. Madrid.
- [70] Sawaragi Y., Nakayama H., T.T. (1985). *Theory of multiobjective optimization*. Mathematics in science and Engineering, vol. 176. Academic Press Inc.
- [71] Souza, S. S., Oliveira, P. R., da Cruz Neto, J. X., Soubeyran, A. (2010). *A proximal method with separable Bregman distance for quasiconvex minimization on the nonnegative orthant*. European Journal of Operational Research, 201: 365-376.
- [72] Suzuki, S, Kuroiwa, D. (2014). *Characterizations of the solution set for quasiconvex programming in terms of Greenberg–Pierskalla subdifferentia*. Springer Science+Business Media New York.
- [73] Wilfredo Pareto. (1906). *Manuale di economia política con una introduzione alla scienza sociale*, Società Editrice Libreria.
- [74] Van Tiel J (1984). *Convex Analysis: An Introductory Text*, Chichester, UK, John Wiley & Sons.
- [75] Villacorta, K.D.V., Oliveira, P.R. (2011). *An interior proximal method in vector optimization*. Eur. J. Oper. Res. 214, 485–492.
- [76] Von Neumann, John, y Oskar Morgenstern (1944). *Theory of games and economic behavior*. Princeton, NJ, Princeton University Press.
- [77] Y. Collette and P. Siarry (2002). *Optimisation multiobjectif*. Editions Eyrolles, Paris.
- [78] L.A. Zadeh. Fuzzy Sets (1965). *Information and Control*, 8:338–353.

ANEXOS

DESCRIPCION DE LA IMPLEMENTACION DEL ALGORITMO EJEMPLO 1

El problema a optimizar este dado por:

$$\left\{ \begin{array}{l} \text{ming}(x_1, x_2) = (-2e^{-2x_1^2 - x_1^2} + 2)z_1^k + (2(x_1 - 1)^2 + (x_2 - 1)^2)z_2^k + \\ \quad \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sujeto a:} \\ \quad 2x_1^2 + x_2^2 - 2(x_1^k)^2 - (x_2^k)^2 \leq 0 \\ \quad (x_1 - 1)^2 + (x_2 - 1)^2 - (x_1^k - 1)^2 - (x_2^k - 1)^2 \leq 0 \end{array} \right.$$

Donde se consideró dos distancias proximales $d(\cdot, \cdot)$, dado por:

i.-Distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln\left(\frac{x_i}{y_i}\right) + y_i - x_i$$

ii.-Distancia proximal homogénea de segundo orden definida por

$$d(u, v) := \sum_{j=1}^2 v_j^2 \left(\frac{u_j}{v_j} - \ln\left(\frac{u_j}{v_j}\right) - 1 \right) + \frac{\sigma}{2} \|x - y\|^2, \text{ con } \sigma = 0.01$$

Los parámetros fijos $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{10}}, \frac{\sqrt{9}}{\sqrt{10}}\right)$, $\lambda_k = 0.5, \lambda_k = 0.3$, además el punto $x^0 = (3, 2)$ de la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas y el criterio de parada $\|x^{k+1} - x^k\| < 0.00001$ para finalizar el algoritmo ISPPMR.

A continuación, se presenta el algoritmo

```
function ejercicio_1_tesis
    clc; clear;
    warning('off','all');
    % CONDICIONES INICIALES DEL PROBLEMA
    % -----
    z1=1/sqrt(10);
    z2=sqrt(1-z1^2);
    lambda = 0.5;
    p0 = [2,3];
    x0 = [3,2];
    op_dis = 2; % {1,2,3,4} No se recomienda la Opción 1 (distancia Euclidiana)
                % ya que trabaja sobre todo Rn
    GRAF = 0; % (0) Muestra gráfico (1) No muestra gráfico
    QUIET = 0; % (0) Muestra iteraciones (1) No muestra iteraciones
    MAX_ITER = 200;
    TOL = 10^-5;

    % DEFINIMOS LAS COMPONENTES DE LA FUNCIÓN OBJETIVO
    % -----
    syms x1 x2
```

```

F1 = -2*exp(-2*x1^2-x2^2)+2;
F2 = 2*(x1-1)^2+(x2-1)^2;

F_obj = matlabFunction(z1*F1+z2*F2, 'vars', {x1 x2});
F1 = matlabFunction(F1, 'vars', {x1 x2});
F2 = matlabFunction(F2, 'vars', {x1 x2});

% DEFINIMOS LAS RESTRICCIONES
% -----
lb = [0 0];

% INICIAMOS EL ALGORITMO
% -----

t_start = tic;

%-----
% CONDICIONES INICIALES PARA EL ALGORITMO
%-----
%options = optimoptions('fmincon','Algorithm','sqp','Display','off');
options =
optimoptions('fmincon','Algorithm','sqp','ConstraintTolerance',1.0000e-
12,'Display','off');
%options = optimset('Display','iter','algorithm','sqp');

fprintf('\n%5s\n','ALGORITMO INICIADO:');
fprintf('%5s\n','-----');
fprintf('%5s\n','FUNCIÓN OPT (SUBPROBLEMA): "fmincon"');
fprintf('%5s\n','ALGORITMO : "sqp"');
format short;
fprintf('%5s\n','TOLERANCIA RESTRICCIONES : "1.0000e-18"');
fprintf('%5s\n','DISPLAY : "off"');

%-----
% INICIO DEL ALGORITMO
%-----
for k = 1:MAX_ITER

    %lambda = 1/k;

    [x,fval, exitflag, output] = fmincon(@(x) (spreadargs(x,F_obj) +
lambda*distancia(x,x0,op_dis)),...
p0,[],[],[],[],[],[],@x)
Restricciones(x,x0),options);

% Cargamos el historial
% -----
history.num_iter(k) = k;
history.lambda(k) = lambda;
history.sub_iter(k) = output.iterations;
history.exitflag(k) = exitflag;
history.x1(k) = x(1);
history.x2(k) = x(2);
history.residuo(k) = norm(x-x0,2);
history.F1(k) = F1(x(1),x(2));
history.F2(k) = F2(x(1),x(2));

```

```

history.objval_prox(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2)) +
lambda*distancia(x,x0,op_dis);
history.objval(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2));

% Checkeamos el criterio de finalización
% -----
if (history.residuo(k) < TOL)
    fprintf('\n%5s\t%10.7f\n', 'ALGORITMO FINALIZADO: Tolerancia
alcanzada: |x - x0| <', TOL);
    fprintf('%5s\n', '-----')
    break;
end

if(history.exitflag(k) == -2)
    fprintf('\n%5s\n', 'ALGORITMO FINALIZADO: Punto no fealizable
encontrado, considere posible punto critico de pareto a la iteración
anterior');
    break;
end

% Actualizamos la variable
% -----
x0 = x;
end

toc(t_start);

% #####
% GENERAMOS LA TABLA DE RESULTADOS
% #####
if ~QUIET
    fprintf('%5s\n', '-*-' )
    Num_iter = history.num_iter';
    Sub_iter = history.sub_iter';
    Exit_flag = history.exitflag';
    Lambda = history.lambda';
    Residuo = history.residuo';
    Solution = [history.x1' history.x2'];
    Fun1 = history.F1';
    Fun2 = history.F2';
    Fun_obj = history.objval';
    %Validation = history.eval';
    result =
table(Num_iter,Sub_iter,Exit_flag,Lambda,Residuo,Solution,...
        Fun1,Fun2,Fun_obj);
    format long
    disp(result);
end

% #####
% Encontramos los puntos y la frontera de pareto usando GA (gamultiobj)
% #####

options = optimoptions("gamultiobj", "PopulationSize", 250);
fit = @(x)[-2*exp(-2*x(1)^2-x(2)^2)+2, 2*(x(1)-1)^2+(x(2)-1)^2];
[sol1 fval1] = gamultiobj(fit,2,[],[],[],[],lb,[],[],options);

figure(1);

```

```

plot(sol1(:,1),sol1(:,2),'ms')
hold on
plot(history.x1(k),history.x2(k),'ro')
text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPMR','Color','red','FontSize',12)
xlabel('x(1)')
ylabel('x(2)')
title('Puntos de pareto (con Gamultiobj)')
grid on

figure(2);
plot(fval1(:,1),fval1(:,2),'ms')
hold on
plot(history.F1(k),history.F2(k),'ro')
text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPMR','Color','red','FontSize',12)
xlabel('F(1)')
ylabel('F(2)')
title('Frontera de pareto (con Gamultiobj)')
grid on

% #####
% Encontramos los puntos y la frontera de pareto usando (paretosearch)
% #####
fun = @(x)[-2*exp(-2*x(1)^2-x(2)^2)+2,2*(x(1)-1)^2+(x(2)-1)^2];
options = optimoptions('paretosearch','ParetoSetChangeTolerance',1e-
8,'ParetoSetSize',200);
[sol2 fval2] = paretosearch(fun,2,[],[],[],[],lb,[],[],options);

figure(3);
plot(sol2(:,1),sol2(:,2),'b*')
hold on
plot(history.x1(k),history.x2(k),'ro')
text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPMR','Color','red','FontSize',12)
xlabel('x(1)')
ylabel('x(2)')
title('Puntos de pareto (con Paretosearch)')
grid on

figure(4);
plot(fval2(:,1),fval2(:,2),'b*')
hold on
plot(history.F1(k),history.F2(k),'ro')
text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPMR','Color','red','FontSize',12)
xlabel('F(1)')
ylabel('F(2)')
title('Frontera de pareto (con Paretosearch)')
grid on

% #####
% GENERAMOS LA REGIÓN FEASIBLE
% #####
h1 = 2; h2 = 1;
dist = @(x1,x2) x1*log(x1/h1)-x1+h1 + ...
x2*log(x2/h2)-x2+h2;

```

```

figure(5);
fsurf(dist);
title('Espacio de restricciones')
xlabel('x(1)');
ylabel('x(2)');

% #####
% GENERAMOS LOS GRÁFICOS DE LAS FUNCIONES
% #####
figure(6);
fsurf(F1,[-1 2 -1 2]);
hold on
fsurf(F2,'ShowContours','on');
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
zlabel('Z');

% #####
% GENERAMOS LAS CURVAS DE NIVEL DE LAS FUNCIONES COMPONENTES
% #####
figure(7);
fcontour(F1,[-1 2 -1 2]);
hold on
fcontour(F2);
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
xt = [0 1];
yt = [0 1];
str = {'$\leftarrow\min [-2e^{(-2x_{1}^{2}-x_{2}^{2})}+2]$', '$\leftarrow\min [2(x_{1}-1)^{2}+(x_{2}-1)^{2}]$'};
t = text(xt,yt,str,'Color','black');
set(t,'Interpreter','latex','fontSize',12)
plot([0 1],[0 1],'r*')

% #####
% GENERAMOS LOS GRÁFICOS
% #####
if ~GRAF
    figure(8);
    plot(history.num_iter,history.residuo,'b-o');
    title('Convergencia "RESIDUO"')
    hold on
    yline(0,'r-','Umbral','LineWidth',3);
    xlabel('Número iteraciones');
    ylabel('Residuo');
    %ylim([-0.5 4]);
    xlim([1 history.num_iter(k)]);
    legend('||x^{k+1}-x^{k}||','CERO')
    grid on
end
end

function [C,Ceq] = Restricciones(x,x0)

% Restricciones de Pareto (desigualdad)
% -----
C1 = (-2*exp(-2*x(1)^2-x(2)^2)+2) - ((-2*exp(-2*x0(1)^2-x0(2)^2)+2));

```

```

C2 = (2*(x(1)-1)^2+(x(2)-1)^2) - (2*(x0(1)-1)^2+(x0(2)-1)^2);

C = [C1;C2];

% Restricciones de Igualdad
% -----
Ceq = [];
end

function d = distancia(x,x0,op_dis)

% DEFINIMOS LAS COMPONENTES PARA LAS DISTANCIAS PROXIMALES
% -----

x1 = x(1);
x2 = x(2);

y1 = x0(1);
y2 = x0(2);

% -----

n = length(x);

if op_dis == 1 % Euclidean (vale para Rn)
    d01 = norm(x1-y1,2)^2;
    d02 = norm(x2-y2,2)^2;
end

if op_dis == 2 % Kullback-Leibler divergence (Vale para Octante positivo)
    d01 = sum(x1.*log(x1./y1)-x1+y1);
    d02 = sum(x2.*log(x2./y2)-x2+y2);
end

if op_dis == 3 % Itakura-Saito distance (Vale para Octante positivo)
    d01 = sum(x1./y1 - log(x1./y1) - ones(n,1));
    d02 = sum(x2./y2 - log(x2./y2) - ones(n,1));
end

if op_dis == 4 % Second-order homogeneous (Vale para Octante positivo)
    sigma = 0.001; v = 0.01;
    d01 = sum((v/2)*((x1-y1).*(x1-y1))+sigma*((y1.*y1).*log(y1./x1) +
x1.*y1 - y1.*y1));
    d02 = sum((v/2)*((x2-y2).*(x2-y2))+sigma*((y2.*y2).*log(y2./x2) +
x2.*y2 - y2.*y2));
end

d = d01 + d02;
end

function r = spreadargs(x, fhandle)
    t = num2cell(x);
    r = fhandle(t{:});
end

```

DESCRIPCION DE LA IMPLEMENTACION DEL ALGORITMO EJEMPLO 2

El problema a optimizar este dado por:

$$\left\{ \begin{array}{l} \text{ming}(x_1, x_2) = (-x_1^{1/3} x_2^{2/3})_{z_1^k} + (-x_1^{0.25} x_2^{0.75})_{z_2^k} + \\ \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sa:} \\ -x_1^{1/3} x_2^{2/3} + (x_1^k)^{1/3} (x_2^k)^{2/3} \leq 0 \\ -x_1^{0.25} x_2^{0.75} + (x_1^k)^{0.25} (x_2^k)^{0.75} \leq 0 \end{array} \right.$$

donde consideramos como distancia proximale $d(\cdot, \cdot)$, a la distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln\left(\frac{x_i}{y_i}\right) + y_i - x_i$$

Variación de los parámetros fijos.

1.- $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\lambda_k = 0.2$, el punto $x^0 = (1, 1)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.0000001$ para finalizar el algoritmo ISPPMR.

2.- $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{3}}, \sqrt{\frac{2}{3}}\right)$, $\lambda_k = 0.4$, el punto $x^0 = (1, 2)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (2, 3)$ para resolver todos los subproblemas y consideramos el criterio de parada $\|x^{k+1} - x^k\| < 0.0000001$ para finalizar el algoritmo ISPPMR.

A continuación, se presenta el algoritmo

```
function ejercicio_2_tesis
    clc; clear;
    warning('off', 'all');
    % CONDICIONES INICIALES DEL PROBLEMA
    % -----
    z1=1/sqrt(2);
    z2=sqrt(1-z1^2);
    lambda = 0.2;
    p0 = [2,3];
    x0 = [1,2];
    op_dis = 2; % {1,2,3,4} No se recomienda la Opción 1 (distancia Euclidiana)
                % ya que trabaja sobre todo Rn
    GRAF = 0; % (0) Muestra gráfico (1) No muestra gráfico
    QUIET = 0; % (0) Muestra iteraciones (1) No muestra iteraciones
    MAX_ITER = 200;
    TOL = 10^-6;
```

```

% DEFINIMOS LAS COMPONENTES DE LA FUNCIÓN OBJETIVO
% -----
syms x1 x2
F1 = -x1^(1/3)*x2^(2/3);
F2 = -x1^(0.25)*x2^(0.75);

F_obj = matlabFunction(z1*F1+z2*F2, 'vars', {x1 x2});
F1 = matlabFunction(F1, 'vars', {x1 x2});
F2 = matlabFunction(F2, 'vars', {x1 x2});

% DEFINIMOS LAS RESTRICCIONES
% -----
A = [-1 1;3 1;3 -1];
b = [2;18;12];
lb = [0 0];

% INICIAMOS EL ALGORITMO
% -----

t_start = tic;

%-----
% CONDICIONES INICIALES PARA EL ALGORITMO
%-----
%options = optimoptions('fmincon','Algorithm','sqp','Display','off');
options =
optimoptions('fmincon','Algorithm','sqp','ConstraintTolerance',1.0000e-
12,'Display','off');
%options = optimset('Display','iter','algorithm','sqp');

fprintf('\n%5s\n', 'ALGORITMO INICIADO:');
fprintf('%5s\n\n', '-----');
fprintf('%5s\n', 'FUNCIÓN OPT (SUBPROBLEMA): "fmincon"');
fprintf('%5s\n', 'ALGORITMO : "sqp"');
format short;
fprintf('%5s\n', 'TOLERANCIA RESTRICCIONES : "1.0000e-18"');
fprintf('%5s\n', 'DISPLAY : "off"');

%-----
% INICIO DEL ALGORITMO
%-----
for k = 1:MAX_ITER

    %lambda = 1/k;

    [x,fval, exitflag, output] = fmincon(@(x) (spreadargs(x,F_obj) +
lambda*distancia(x,x0,op_dis)),...
p0,[],[],[],[],[],[],@x)
Restricciones(x,x0),options);

    % Cargamos el historial
    % -----
    history.num_iter(k) = k;
    history.lambda(k) = lambda;
    history.sub_iter(k) = output.iterations;
    history.exitflag(k) = exitflag;
    history.x1(k) = x(1);
    history.x2(k) = x(2);
    history.residuo(k) = norm(x-x0,2);

```

```

history.F1(k) = F1(x(1),x(2));
history.F2(k) = F2(x(1),x(2));
history.objval_prox(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2)) +
lambda*distancia(x,x0,op_dis);
history.objval(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2));

% Checkeamos el criterio de finalización
% -----
if (history.residuo(k) < TOL)
    fprintf('\n%5s\t%10.7f\n', 'ALGORITMO FINALIZADO: Tolerancia
alcanzada:  $|x - x_0| <$ ', TOL);
    fprintf('%5s\n', '-----')
    break;
end

if(history.exitflag(k) == -2)
    fprintf('\n%5s\n', 'ALGORITMO FINALIZADO: Punto no feasible
encontrado, considere posible punto critico de pareto a la iteración
anterior');
    break;
end

% Actualizamos la variable
% -----
x0 = x;
end

toc(t_start);

% #####
% GENERAMOS LA TABLA DE RESULTADOS
% #####
if ~QUIET
    fprintf('%5s\n', '-*-')
    Num_iter = history.num_iter';
    Sub_iter = history.sub_iter';
    Exit_flag = history.exitflag';
    Lambda = history.lambda';
    Residuo = history.residuo';
    Solution = [history.x1' history.x2'];
    Fun1 = history.F1';
    Fun2 = history.F2';
    Fun_obj = history.objval';
    %Validation = history.eval';
    result =
table(Num_iter,Sub_iter,Exit_flag,Lambda,Residuo,Solution,...
        Fun1,Fun2,Fun_obj);
    format long
    disp(result);
end

% #####
% Encontramos los puntos y la frontera de pareto usando GA (gamultiobj)
% #####
options = optimoptions("gamultiobj", "PopulationSize", 150);
fit = @(x)[-x(1)^(1/3)*x(2)^(2/3), -x(1)^(0.25)*x(2)^(0.75)];
[sol1 fval1] = gamultiobj(fit,2,A,b,[],[],lb,[],[],options);
%

```

```

%         figure(1);
%
%         plot(sol1(:,1),sol1(:,2),'ms')
%         hold on
%         plot(history.x1(k),history.x2(k),'ro')
%         text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
%         xlabel('x(1)')
%         ylabel('x(2)')
%         title('Puntos de pareto (con Gamultiobj)')
%         grid on
%
%         figure(2);
%         plot(fval1(:,1),fval1(:,2),'ms')
%         hold on
%         plot(history.F1(k),history.F2(k),'ro')
%         text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
%         xlabel('F(1)')
%         ylabel('F(2)')
%         title('Frontera de pareto (con Gamultiobj)')
%         grid on
%
%         %
#####
%         % Encontramos los puntos y la frontera de pareto usando
(paretosearch)
%         %
#####
%         fun = @(x)[-x(1)^(1/3)*x(2)^(2/3),-x(1)^(0.25)*x(2)^(0.75)];
%         options = optimoptions('paretosearch','ParetoSetChangeTolerance',1e-
8,'ParetoSetSize',200);
%         [sol2 fval2] = paretosearch(fun,2,A,b,[],[],lb,[],[],options);
%
%         figure(3);
%         plot(sol2(:,1),sol2(:,2),'b*')
%         hold on
%         plot(history.x1(k),history.x2(k),'ro')
%         text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
%         xlabel('x(1)')
%         ylabel('x(2)')
%         title('Puntos de pareto (con Paretosearch)')
%         grid on
%
%         figure(4);
%         plot(fval2(:,1),fval2(:,2),'b*')
%         hold on
%         plot(history.F1(k),history.F2(k),'ro')
%         text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
%         xlabel('F(1)')
%         ylabel('F(2)')
%         title('Frontera de pareto (con Paretosearch)')
%         grid on

% #####
% GENERAMOS LA REGIÓN FEASIBLE
% #####

```

```

h1 = 2; h2 = 1;
dist = @(x1,x2) x1*log(x1/h1)-x1+h1 + ...
              x2*log(x2/h2)-x2+h2 + ...
              (x1 - x2 + 2)*log((x1 - x2 + 2)/(h1 - h2 + 2))-(x1 - x2
+ 2)+(h1 - h2 + 2) + ...
              (-3*x1 - x2 + 18)*log((-3*x1 - x2 + 18)/(-3*h1 - h2 +
18))-(-3*x1 - x2 + 18)+(-3*h1 - h2 + 18) + ...
              (-3*x1 + x2 + 12)*log((-3*x1 + x2 + 12)/(-3*h1 + h2 +
12))-(-3*x1 + x2 + 12)+(-3*h1 + h2 + 12);
figure(5);
fsurf(dist,[0 6 0 6]);
title('Espacio de restricciones')
xlabel('x(1)');
ylabel('x(2)');

% #####
% GENERAMOS LOS GRÁFICOS DE LAS FUNCIONES
% #####
figure(6);
fsurf(F1,[0 6 0 6]);
hold on
fsurf(F2,'ShowContours','on');
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
zlabel('Z');

% #####
% GENERAMOS LAS CURVAS DE NIVEL DE LAS FUNCIONES COMPONENTES
% #####
figure(7);
fcontour(F1,[0 6 0 6]);
hold on
fcontour(F2,[0 6 0 6]);
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
%xt = [0 1];
%yt = [0 1];
%str = {'$\leftarrow\min [-2e^{(-2x_{1}^{2}-
x_{2}^{2})}]$','$\leftarrow\min [2(x_{1}-1)^{2}+(x_{2}-1)^{2}]$'};
%t = text(xt,yt,str,'Color','black');
%set(t,'Interpreter','latex','fontsize',12)
%plot([0 1],[0 1],'r*')

% #####
% GENERAMOS LOS GRÁFICOS
% #####
if ~GRAF
figure(8);
plot(history.num_iter,history.residuo,'b-o');
title('Convergencia "RESIDUO"')
hold on
yline(0,'r-.','Umbral','LineWidth',3);
xlabel('Número iteraciones');
ylabel('Residuo');
%ylim([-0.5 4]);
xlim([1 history.num_iter(k)]);
legend('||x^{k+1}-x^{k}||','CERO')

```

```

        grid on
    end
end

function [C,Ceq] = Restricciones(x,x0)

    % Restricciones de Pareto (desigualdad)
    % -----
    C1 = (-x(1)^(1/3)*x(2)^(2/3)) - (-x0(1)^(1/3)*x0(2)^(2/3));
    C2 = (-x(1)^(0.25)*x(2)^(0.75)) - (-x0(1)^(0.25)*x0(2)^(0.75));

    C = [C1;C2];

    % Restricciones de Igualdad
    % -----
    Ceq = [];
end

function d = distancia(x,x0,op_dis)

    % DEFINIMOS LAS COMPONENTES PARA LAS DISTANCIAS PROXIMALES
    % -----
    x1 = x(1);
    x2 = x(2);
    x3 = x(1) - x(2) + 2;
    x4 = -3*x(1) - x(2) + 18;
    x5 = -3*x(1) + x(2) + 12;

    y1 = x0(1);
    y2 = x0(2);
    y3 = x0(1) - x0(2) + 2;
    y4 = -3*x0(1) - x0(2) + 18;
    y5 = -3*x0(1) + x0(2) + 12;

    % -----

    n = length(x);

    if op_dis == 1 % Euclidean (vale para Rn)
        d01 = norm(x1-y1,2)^2;
        d02 = norm(x2-y2,2)^2;
        d03 = norm(x3-y3,2)^2;
        d04 = norm(x4-y4,2)^2;
        d05 = norm(x5-y5,2)^2;
    end

    if op_dis == 2 % Kullback-Leibler divergence (Vale para Octante positivo)
        d01 = sum(x1.*log(x1./y1)-x1+y1);
        d02 = sum(x2.*log(x2./y2)-x2+y2);
        d03 = sum(x3.*log(x3./y3)-x3+y3);
        d04 = sum(x4.*log(x4./y4)-x4+y4);
        d05 = sum(x5.*log(x5./y5)-x5+y5);
    end

    if op_dis == 3 % Itakura-Saito distance (Vale para Octante positivo)
        d01 = sum(x1./y1 - log(x1./y1) - ones(n,1));
        d02 = sum(x2./y2 - log(x2./y2) - ones(n,1));
        d03 = sum(x3./y3 - log(x3./y3) - ones(n,1));
        d04 = sum(x4./y4 - log(x4./y4) - ones(n,1));
    end
end

```

```

    d05 = sum(x5./y5 - log(x5./y5) - ones(n,1));
end

if op_dis == 4 % Second-order homogeneous (Vale para Octante positivo)
    sigma = 0.001; v = 0.01;
    d01 = sum((v/2)*((x1-y1).*(x1-y1))+sigma*((y1.*y1).*log(y1./x1) +
x1.*y1 - y1.*y1));
    d02 = sum((v/2)*((x2-y2).*(x2-y2))+sigma*((y2.*y2).*log(y2./x2) +
x2.*y2 - y2.*y2));
    d03 = sum((v/2)*((x3-y3).*(x3-y3))+sigma*((y3.*y3).*log(y3./x3) +
x3.*y3 - y3.*y3));
    d04 = sum((v/2)*((x4-y4).*(x4-y4))+sigma*((y4.*y4).*log(y4./x4) +
x4.*y4 - y4.*y4));
    d05 = sum((v/2)*((x5-y5).*(x5-y5))+sigma*((y5.*y5).*log(y5./x5) +
x5.*y5 - y5.*y5));
end

d = d01 + d02 + d03 + d04 + d05;
end

function r = spreadargs(x, fhandle)
    t = num2cell(x);
    r = fhandle(t{:});
end

```

DESCRIPCION DE LA IMPLEMENTACION DEL ALGORITMO EJEMPLO 3

El problema a optimizar es dado por

$$\begin{cases} \min g(x_1, x_2) = (3200x_1^{1/4}x_2^{1/2} - 80x_1 - 40x_2)z_1^k + (80x_1 + 40x_2)z_2^k + \\ \lambda_k d\left(\left((x_1, x_2), (x_1^k, x_2^k)\right)\right) \\ \text{Sa:} \\ 3200x_1^{1/4}x_2^{1/2} - 80x_1 - 40x_2 - 3200(x_1^k)^{1/4}(x_2^k)^{1/2} + 80x_1^k + 40x_2^k \leq 0 \\ 80x_1 + 40x_2 - 80x_1^k - 40x_2^k \leq 0 \end{cases}$$

La distancia proximal $d(\cdot, \cdot)$, distancia Kullback-Leibler Bregman definida por

$$d(x, y) := \sum_{i=1}^2 x_i \ln\left(\frac{x_i}{y_i}\right) + y_i - x_i$$

Los parámetros fijos $z_k = (z_1^k, z_2^k) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\lambda_k = 0.1$, además el punto $x^0 = (100, 10)$ donde se obtiene la iteración 0 y el punto inicial $p^0 = (10, 10)$ para resolver todos los subproblemas y el criterio de parada $\|x^{k+1} - x^k\| < 0.000001$ para finalizar el algoritmo ISPPMR.

A continuación, se presenta el algoritmo

```
function ejercicio_3_tesis
```

```

clc; clear;
warning('off','all');
% CONDICIONES INICIALES DEL PROBLEMA
% -----
z1=1/sqrt(2);
z2=sqrt(1-z1^2);
lambda = 0.1;
p0 = [10,10];
x0 = [100,10];
op_dis = 2; % {1,2,3,4} No se recomienda la Opción 1 (distancia Euclidiana)
           % ya que trabaja sobre todo Rn
GRAF = 0; % (0) Muestra gráfico (1) No muestra gráfico
QUIET = 0; % (0) Muestra iteraciones (1) No muestra iteraciones
MAX_ITER = 200;
TOL = 10^-6;

% DEFINIMOS LAS COMPONENTES DE LA FUNCIÓN OBJETIVO
% -----
syms x1 x2
F1 = -3200*x1^(0.25)*x2^(0.5)+80*x1+40*x2;
F2 = 80*x1+40*x2;

F_obj = matlabFunction(z1*F1+z2*F2, 'vars', {x1 x2});
F1 = matlabFunction(F1, 'vars', {x1 x2});
F2 = matlabFunction(F2, 'vars', {x1 x2});

% DEFINIMOS LAS RESTRICCIONES
% -----
A = [80 40];
b = [10240];
lb = [0 0];

% INICIAMOS EL ALGORITMO
% -----

t_start = tic;

%-----
% CONDICIONES INICIALES PARA EL ALGORITMO
%-----
options = optimoptions('fmincon','Algorithm','sqp','Display','off');
options =
optimoptions('fmincon','Algorithm','sqp','ConstraintTolerance',1.0000e-
12,'Display','off');
options = optimset('Display','iter','algorithm','sqp');

fprintf('\n%5s\n','ALGORITMO INICIADO:');
fprintf('%5s\n\n','-----');
fprintf('%5s\n','FUNCIÓN OPT (SUBPROBLEMA): "fmincon"');
fprintf('%5s\n','ALGORITMO : "sqp"');
format short;
fprintf('%5s\n','TOLERANCIA RESTRICCIONES : "1.0000e-18"');
fprintf('%5s\n','DISPLAY : "off"');

%-----
% INICIO DEL ALGORITMO
%-----
for k = 1:MAX_ITER

```

```

%lambda = 1/k;

[x,fval, exitflag, output] = fmincon(@(x) (spreadargs(x,F_obj) +
lambda*distancia(x,x0,op_dis)),...
p0,[],[],[],[],[],[],@x)
Restricciones(x,x0),options);

% Cargamos el historial
% -----
history.num_iter(k) = k;
history.lambda(k) = lambda;
history.sub_iter(k) = output.iterations;
history.exitflag(k) = exitflag;
history.x1(k) = x(1);
history.x2(k) = x(2);
history.residuo(k) = norm(x-x0,2);
history.F1(k) = F1(x(1),x(2));
history.F2(k) = F2(x(1),x(2));
history.objval_prox(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2)) +
lambda*distancia(x,x0,op_dis);
history.objval(k) = z1*F1(x(1),x(2)) + z2*F2(x(1),x(2));

% Checkeamos el criterio de finalización
% -----
if (history.residuo(k) < TOL)
    fprintf('\n%5s\t%10.7f\n', 'ALGORITMO FINALIZADO: Tolerancia
alcanzada: |x - x0| <',TOL);
    fprintf('%5s\n', '-----')
    break;
end

if(history.exitflag(k) == -2)
    fprintf('\n%5s\n', 'ALGORITMO FINALIZADO: Punto no fealizable
encontrado, considere posible punto critico de pareto a la iteración
anterior');
    break;
end

% Actualizamos la variable
% -----
x0 = x;
end

toc(t_start);

% #####
% GENERAMOS LA TABLA DE RESULTADOS
% #####
if ~QUIET
    fprintf('%5s\n', '-*-')
    Num_iter = history.num_iter';
    Sub_iter = history.sub_iter';
    Exit_flag = history.exitflag';
    Lambda = history.lambda';
    Residuo = history.residuo';
    Solution = [history.x1' history.x2'];
    Fun1 = history.F1';
    Fun2 = history.F2';

```

```

        Fun_obj = history.objval';
        %Validation = history.eval';
        result =
table(Num_iter,Sub_iter,Exit_flag,Lambda,Residuo,Solution,...
        Fun1,Fun2,Fun_obj);
        format long
        disp(result);
end

% #####
% Encontramos los puntos y la frontera de pareto usando GA (gamultiobj)
% #####
options = optimoptions("gamultiobj","PopulationSize",150);
fit = @(x)[-
3200*x(1)^(0.25)*x(2)^(0.5)+80*x(1)+40*x(2),80*x(1)+40*x(2)];
[sol1 fval1] = gamultiobj(fit,2,A,b,[],[],lb,[],[],options);

figure(1);

plot(sol1(:,1),sol1(:,2),'ms')
hold on
plot(history.x1(k),history.x2(k),'ro')
text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
xlabel('x(1)')
ylabel('x(2)')
title('Puntos de pareto (con Gamultiobj)')
grid on

figure(2);
plot(fval1(:,1),fval1(:,2),'ms')
hold on
plot(history.F1(k),history.F2(k),'ro')
text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPD','Color','red','FontSize',12)
xlabel('F(1)')
ylabel('F(2)')
title('Frontera de pareto (con Gamultiobj)')
grid on

% #####
% Encontramos los puntos y la frontera de pareto usando (paretosearch)
% #####
fun = @(x)[-
3200*x(1)^(0.25)*x(2)^(0.5)+80*x(1)+40*x(2),80*x(1)+40*x(2)];
options = optimoptions('paretosearch','ParetoSetChangeTolerance',1e-
8,'ParetoSetSize',200);
[sol2 fval2] = paretosearch(fun,2,A,b,[],[],lb,[],[],options);

figure(3);
plot(sol2(:,1),sol2(:,2),'b*')
hold on
plot(history.x1(k),history.x2(k),'ro')
text(history.x1(k),history.x2(k),'\leftarrow Solución
ISPPMR','Color','red','FontSize',12)
xlabel('x(1)')
ylabel('x(2)')
title('Puntos de pareto (con Paretosearch)')
grid on

```

```

figure(4);
plot(fval2(:,1),fval2(:,2),'b*')
hold on
plot(history.F1(k),history.F2(k),'ro')
text(history.F1(k),history.F2(k),'\leftarrow Solución
ISPPMR', 'Color','red', 'FontSize',12)
xlabel('F(1)')
ylabel('F(2)')
title('Frontera de pareto (con Paretosearch)')
grid on

% #####
% GENERAMOS LA REGIÓN FEASIBLE
% #####
h1 = 2; h2 = 1;
dist = @(x1,x2) x1*log(x1/h1)-x1+h1 + ...
          x2*log(x2/h2)-x2+h2 + ...
          (-80*x1 -40*x2 + 10240)*log((-80*x1 -40*x2 + 10240)/(-
80*h1 -40*h2 + 10240))-(-80*x1 -40*x2 + 10240)+(-80*h1 -40*h2 + 10240);
figure(5);
fsurf(dist,[0 150 0 300]);
title('Espacio de restricciones')
xlabel('x(1)');
ylabel('x(2)');

% #####
% GENERAMOS LOS GRÁFICOS DE LAS FUNCIONES
% #####
figure(6);
fsurf(F1,[0 150 0 300]);
hold on
fsurf(F2,'ShowContours','on');
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
zlabel('Z');

% #####
% GENERAMOS LAS CURVAS DE NIVEL DE LAS FUNCIONES COMPONENTES
% #####
figure(7);
fcontour(F1,[0 150 0 300]);
hold on
fcontour(F2,[0 150 0 300]);
title('Funciones componentes')
xlabel('x(1)');
ylabel('x(2)');
%xt = [0 1];
%yt = [0 1];
%str = {'$\leftarrow\min [-2e^{(-2x_{1}^{2}-
x_{2}^{2})}]$','$\leftarrow\min [2(x_{1}-1)^{2}+(x_{2}-1)^{2}]$'};
%t = text(xt,yt,str,'Color','black');
%set(t,'Interpreter','latex','fontsize',12)
%plot([0 1],[0 1],'r*')

% #####
% GENERAMOS LOS GRÁFICOS
% #####

```

```

    if ~GRAF
        figure(8);
        plot(history.num_iter,history.residuo,'b-o');
        title('Convergencia "RESIDUO"')
        hold on
        yline(0,'r-.','Umbral','LineWidth',3);
        xlabel('Número iteraciones');
        ylabel('Residuo');
        %ylim([-0.5 4]);
        xlim([1 history.num_iter(k)]);
        legend('||x^{k+1}-x^{k}||','CERO')
        grid on
    end
end

function [C,Ceq] = Restricciones(x,x0)

    % Restricciones de Pareto (desigualdad)
    % -----
    C1 = (-3200*x(1)^(0.25)*x(2)^(0.5)+80*x(1)+40*x(2)) - (-
3200*x0(1)^(0.25)*x0(2)^(0.5)+80*x0(1)+40*x0(2));
    C2 = (80*x(1)+40*x(2)) - (80*x0(1)+40*x0(2));

    C = [C1;C2];

    % Restricciones de Igualdad
    % -----
    Ceq = [];
end

function d = distancia(x,x0,op_dis)

    % DEFINIMOS LAS COMPONENTES PARA LAS DISTANCIAS PROXIMALES
    % -----
    x1 = x(1);
    x2 = x(2);
    x3 = -80*x(1) -40*x(2) + 10240;

    y1 = x0(1);
    y2 = x0(2);
    y3 = -80*x0(1) -40*x0(2) + 10240;

    % -----

    n = length(x);

    if op_dis == 1 % Euclidean (vale para Rn)
        d01 = norm(x1-y1,2)^2;
        d02 = norm(x2-y2,2)^2;
        d03 = norm(x3-y3,2)^2;
    end

    if op_dis == 2 % Kullback-Leibler divergence (Vale para Octante positivo)
        d01 = sum(x1.*log(x1./y1)-x1+y1);
        d02 = sum(x2.*log(x2./y2)-x2+y2);
        d03 = sum(x3.*log(x3./y3)-x3+y3);
    end

    if op_dis == 3 % Itakura-Saito distance (Vale para Octante positivo)

```

```

    d01 = sum(x1./y1 - log(x1./y1) - ones(n,1));
    d02 = sum(x2./y2 - log(x2./y2) - ones(n,1));
    d03 = sum(x3./y3 - log(x3./y3) - ones(n,1));
end

if op_dis == 4 % Second-order homogeneous (Vale para Octante positivo)
    sigma = 0.001; v = 0.01;
    d01 = sum((v/2)*((x1-y1).*(x1-y1))+sigma*((y1.*y1).*log(y1./x1) +
x1.*y1 - y1.*y1));
    d02 = sum((v/2)*((x2-y2).*(x2-y2))+sigma*((y2.*y2).*log(y2./x2) +
x2.*y2 - y2.*y2));
    d03 = sum((v/2)*((x3-y3).*(x3-y3))+sigma*((y3.*y3).*log(y3./x3) +
x3.*y3 - y3.*y3));
end

    d = d01 + d02 + d03;
end

function r = spreadargs(x, fhandle)
    t = num2cell(x);
    r = fhandle(t{:});
end

```

Un Método Proximal para Optimización Multiobjetivo Cuasi-convexa y su aplicación a la economía

por Osmar Bermejo

Fecha de entrega: 14-sep-2023 06:48p.m. (UTC-0500)

Identificador de la entrega: 2166400205

Nombre del archivo: TESIS_DOCTORAL_OSMAR_BERMEJO_UNI_2022_R.docx (3.31M)

Total de palabras: 25353

Total de caracteres: 159383

8

www.math.uwaterloo.ca

Fuente de Internet

<1 %

9

Submitted to Universidad Nacional del Santa

Trabajo del estudiante

<1 %

10

schorsch.efi.fh-nuernberg.de

Fuente de Internet

<1 %

11

1library.co

Fuente de Internet

<1 %

12

dspace.utb.edu.ec

Fuente de Internet

<1 %

Excluir citas

Activo

Excluir coincidencias < 15 words

Excluir bibliografía

Activo