Optimización multiobjetivo intervalo-valuada

Juan Pablo Fernández Gutiérrez

ESCUELA DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE CIENCIAS BÁSICAS

MAESTRÍA EN MATEMÁTICAS APLICADAS

MEDELLÍN

Marzo 2008

Optimización multiobjetivo intervalo-valuada

Trabajo de investigación presentado como requisito para optar al título de Magíster en Matemáticas Aplicadas

Juan Pablo Fernández Gutiérrez

Director

María Eugenia Puerta Yepes

Doctor en Ciencias Matemáticas

ESCUELA DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE CIENCIAS BÁSICAS

MAESTRÍA EN MATEMÁTICAS APLICADAS

MEDELLÍN

Marzo 2008

Nota de aceptación

Coordinador de la Maestría

Director del proyecto

Ciudad y fecha (día, mes, año):

Agradecimientos

Lámpara es a mis pies tu palabra y lumbrera en mi camino (salmos 119:105)

Quiero expresar mi gratitud con todas las personas que contribuyeron a la culminación de esta tesis, principalmente a su directora la Dr. Maria Eugenia Puerta Yepes por su generosa ayuda y consejos y a mi madre, mi esposa y mi hija, las mujeres principales en mi vida y todas las personas quienes me hayan ofrecido un invaluable apoyo.

Resumen

En 1990, Ishibuchi y Tanaka presentaron una nueva metodología para los problemas de optimización con *función objetivo lineal intervalo valuada* [4], dando una visión multiobjetivo al concepto de solución, diferenciadose de otros enfoques, como el dado por la optimización inexacta [17]. En 2007 Hsien-Chung Wu hizo desarrollos teóricos para el problema propuesto por Ishibuchi y Tanaka, extiendiendo la visión a funciones no lineales, construyendo una estructura analítica y formulando un teorema similar al teorema de Karush-Kuhn-Tucker para problemas de optimización monó-objetivo de valor real [1] [2]. Posteriormente, en otro artículo [1], Wu considera también funciones restricción intervalo valuadas en el problema de optimización intervalo valuado, pero este se puede reducir al problema de optimización intervalo valuado, con restriciones real valuadas.

Aquí presentamos una visión inspirada en la propuesta hecha por el profesor Hsien-Chung Wu, para el caso de funciones multi-intervalo valuadas, construyendo unas definiciones similares al conjunto de soluciones de Pareto y el conjunto eficiente, para un problema de optimización de funciones multi-intervalo valuadas y retricciones real valuadas.

Se enuncian dos conceptos de solución de Pareto para este problema, al considerar dos ordenes parciales sobre el conjunto de todos los arreglos $m \times 1$ de intervalos reales cerrados, acotados y convexos (multi-intervalos).

Se generar un método para resolver el problema de optimización de funciones multi-intervalo valuadas, apoyados en el uso de las funciones scalarizing lo que permite hallar algunas de las soluciones de Pareto y el conjunto eficiente o frontera de Pareto [8]. Esto es posible a través de un teorema de concidión suficiente y necesaria de difereciación de funciones multi-intervalo valuadas en términos de sus componentes, el cual se enuncia y se demuestra en este trabajo y que está soportado en la construcción análitica que se hace sobre el conjunto de multi-intervalos, donde se definen dos conceptos de diferenciación, uno denominado débil y basado en el teorema de embebimiento de semigrupos topologicos [5]; otro, denominado fuerte que involucra el concepto de diferencia de Hukuhara para espacios convexos [6]. Por último, presentamos algunos ejemplos numéricos acerca del tipo de problema que estamos planteando en este trabajo de grado.

[1] Aquí las restricciones son funciones de valor real

Índice general

Introducción

Algunos problemas prácticos de optimización se pueden modelar matemáticamente como:

$$\begin{array}{cc} \text{mín}\,(\text{máx}) & f\left(x\right) \\ \text{s.a} & x \in \mathcal{F} \subseteq \mathbb{R}^n \end{array}$$

en el caso en que $\mathcal{F}$ es no vacío y $\mathcal{F} \neq \mathbb{R}^n$ se denomina optimización restringida, además, al conjunto $\mathcal{F}$ también se llama el conjunto factible y a cada punto $x \in \mathcal{F}$, se dice punto factible; la función objetivo $f \colon \mathbb{R}^n \to \mathbb{R}^m$, si $m = 1$ se denomina mono-objetivo o tradicional, y cuando $m \geq 2$ se denomina multiobjetivo.

En los problemas de optimización tradicionales hay una gran gama de teorías o herrramientas para determinar la(s) solucion(es) óptima(s). En la actualidad, cuando la aleatoriedad esta presente en los problemas de optimización estos pueden tener su solución en el marco de la optimización estocástica, y cuando es la imprecisión (fuzziness) la que está presente se enmarcan dentro de la optimización difusa o imprecisa (fuzzy optimization) [12], [9]. En 1990 Ishibuchi y Tanaka [4] presentaron una nueva forma de plantear problemas de optimización matemática con funciones objetivos con coeficientes intervalos, visto como un problema multiobjetivo usando unas relaciones de orden parcial apropiadas. En 2007 Hsein-Chung Wu [2], ha construido una teoría para esta formulación, la cual es a la vez una alternativa para enfrentar la problematica de incertidumbre en los problemas de optimización tradicionales, esto es, problemas de optimización con función objetivo intervalo-valuada. En tal propuesta se han hecho desarrollos análogos a la teoría conocida para los problemas de optimización donde la función objetivo es de valor real, llegando a obtener para este nuevo problema, un resultado análogo al teorema de Karush-Kuhn-Tucker (KKT) bajo el espacio métrico completo

$$\mathcal{I} = \{I \subset \mathbb{R} : I \text{ intervalo cerrado, acotado y convexo}\}$$

el cual es dotado de dos ordenes parciales convenientes, generando dos tipos de soluciones y donde la función objetivo es $f \colon \mathbb{R}^n \to \mathcal{I}$, la cual es llamada función objetivo intervalo-valuada.

Otro problema de optimización que debemos considerar para nuestro trabajo, es el siguiente:

$$\min_{x \in \mathcal{F}}\,(\text{máx}) \quad f\left(x\right) \tag{1}$$

donde $f \colon \mathbb{R}^n \to \mathbb{R}^m$, $n \geq 1$ y $m \geq 2$. A este problema se le llama de optimización multiobjetivo o multicriterio (MOP, multiobjective optimization problems). Muchos problemas prácticos admiten, en forma natural, una formulación donde intervienen multiples objetivos o criterios, debido a esto, la optimización multiobjetivo ha tomado gran importancia en la investigación de operaciones, tales como la planeación de ubicación de centrales telefónicas, eléctricas, problemas de selección de varios criterios o características, entre otros.

Cuando tenemos en cuenta la incertidumbre en las imagenes de la función multiobjetivo en el problema (1), en realidad no tenemos un vector como imagen de la función, sino un arreglo $m \times 1$ de intervalos de valores, esto es, el problema multiobjetivo intervalo valuado, que puede ser descrito como sigue:

$$f\colon \mathbb{R}^n \to \mathcal{I}^m, \quad m \geq 2 \tag{2}$$

donde $\mathcal{I}^m = \{[I_1, \cdots, I_m] \subseteq \mathbb{R}^m : I_i \in \mathcal{I},\ i = 1, \cdots, m\}$ y f se llama función multiobjetivo intervalo-valuada.

En este trabajo, generamos una metodología similar a la de los problemas de optimización denominados multiobjetivo, considerando la incertidumbre en ellos, y haciendo un tratamiento semejante al realizado en [2], a través de los ordenes parciales que definimos para $\mathcal{I}^m$ una condición equivalente a la de Pareto en optimización multiobjetivo. Además, considerar también las soluciones por médio de la técnica por metas (goals attending), cuando algunos objetivos tienen un orden de prioridad de acuerdo a la condiciones del problema a resolver. Esto se desarrolla en los capítulos 3 y 4.

Presentaremos el conjunto $\mathcal{F}$ de restricciones por medio de las inecuaciones:

$$\begin{aligned} g_i(x) &\leq 0,\ i = 1, \cdots, k \\ x &\geq 0 \end{aligned}$$

y como parte fundamental de este trabajo, en el capítulo 4, se muestran los resultados de los algorítmos computacionales para resolver el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s. a.} & \begin{cases} g_i(x) \leq 0,\ i = 1, \cdots, k \\ x \geq 0 \end{cases} \end{array} \tag{3}$$

donde $f\colon \mathbb{R}^n \to \mathcal{I}^m$.

Capítulo 1

Propuesta inicial

1.1. Objetivos

1.1.1. Objetivo General

Construir una metodología para resolver el problema de optimización multiobjetivo intervalo-valuado para $m \geq 2$ y $1 \leq n < \infty$ descrito en la ecuación (3).

1.1.2. Objetivos Específicos

1. Estudiar la teoría existe para la optimización con función objetivo intervalo-valuada y extender las herramientas teoricas al problema de optimización multiobjetivo intervalo-valuado.

2. Desarrollar una primera versión de un algorítmo computacional para resolver el problema de optimización multiobjetivo intervalo-valudo.

1.2. Metodología

Iniciamos esta investigación con una busqueda de información sobre todos los temas relacionados al que se propone en este escrito, además de las definiciones básicas que son necesarias para la comprensión y contextualización de todo el contenido del trabajo; estos preliminares se exponen de una forma muy concisa y resumida en el capítulo 2.

En el capítulo 3, basados en las estructura algebráica de los espacios Convexos y las estructuras analíticas de las multifunciones que tienen rango en espacios convexos particulares, aquí hemos introducido una métrica equivalente a la métrica de Hausdorff que ha sido la métrica que tradicionalmente se había usado, para construir la estructura algebráica y análitica necesaria para el problema de optimización en el conjunto de multi-intervalos, destacando especialmente la teoría sobre la estructura diferencial que toma un rol relevante en el desarrollo de los métodos numéricos. Además, también se revisa algunos apartes de la teoría propuesta por Wu, [2] y [1], y algunos elementos de optimización multiobjetivo. Además, se precisa, a partir de los ordenes propuestos, el significado de la palabra mín en nuestro contexto, conjugando los conceptos de solución de la optimización multi-objetivo y la optimización monó-objetivo intervalo valuada, y las primeras definiciones y métodos para resolver dicho problema.

Por último, realizamos unos ejemplos numéricos para observar y reconocer el comportamiento de las soluciones en los sentidos planteados, teniendo en cuenta que a los ejemplos numéricos se les ha impuesto unas condiciones apropiadas para garantizar la(s) solución(es), al requerir el uso de la estructura diferencial antes mencionada.

Al final, se presenta una sección con los problemas abiertos que subyacen en las ideas introducidas en esta tesis y que son motivo para futuros proyectos de investigacion.

Capítulo 2

Conceptos y resultados preliminares

2.1. Notación

Con el ánimo de adoptar una notación habitual, en este trabajo denotamos $\mathbb{R}$ al cuerpo de los números reales, $\mathbb{R}^n$ al espacio vectorial euclideano de dimension n. A los conjuntos con estructuras matemáticas usuales los denotaremos con letras mayúsculas y haremos claridad sobre la estructura particular en el lugar necesario. A los elementos de tales conjuntos los denotaremos en general con letras minúsculas. Denotamos d_* a la métrica con subíndice $*$ en un espacio métrico. Cuando sea necesario enfatizar la norma de un elemento x de un espacio normado N, se escribirá $\|x\|_N$, y escribiremos simplemente $\|x\|$ si no hay lugar a confusión, aquí la norma vectorial será inducida por la métrica correspondiente. Al interior y la frontera de un subconjunto M de un espacio métrico N, lo denotaremos $intM$ y ∂M, respectivamente.

Se usarán las letras minúsculas f, g y h para denotar funciones; las letras mayúsculas A, B y C subíndizadas o no, para hacer referencia a los elementos de los conjuntos $\mathcal{I}$ ó $\mathcal{I}^m$ los cuales se definen en la siguiente sección.

El simbolo $\preceq$ denotará ordenes parciales en general, en particular se presentarán dos ordenes parciales denotados $\preceq_{LU}$ y $\preceq_{UC}$.

En lo que sigue presentaremos los elementos básicos de la propuesta hecha por Wu [2] y [1] sobre el problema de optimización con función monó-objetivo intervalo valuada y restricciones real valuada e intervalo valuadas, así como también algunos elementos de optimización multi-objetivo y resultados de la topología algebráica necesarios para nuestros resultados. Todo esto es necesario para hacer una extensión de la teoría desarrollada para el problema de optimización intervalo-valuado al problema de optimización multiobjetivo intervalo-valuado. Esto se logra llevando a cabo un derrotero teórico similar al desarrollado en [2], por esta razón es importante retomar algunas definiciones de [2].

2.2. Elementos Preliminares

El conjunto $\mathcal{I}$ y su estructura algebraica y análitica es presentada en [2] para el estudio del problema de optimimización intevalo-valuado.

El conjunto $\mathcal{I}$ se define como:

$$\mathcal{I} = \{I \subseteq \mathbb{R} : I \text{ intervalo cerrado, acotado y convexo}\}$$

Si $A \in \mathcal{I}$ entonces $A = \left[a^L, a^U\right]$, donde $a^L \leq a^U$ y los superíndices L y U provienen de sus correspondientes palabras en inglés Lower y Upper, para simbolizar el extremo inferior y el extremo superior del intervalo, ambos extremos deben ser finitos. Además, si $A, B \in \mathcal{I}$, con $A = \left[a^L, a^U\right]$ y $B = \left[b^L, b^U\right]$,

$$A = B \text{ si y sólo si } a^L = b^L \text{ y } a^U = b^U$$

En $\mathcal{I}$ se pueden definir las siguientes operaciones algebraicas:

Sean $A, B \in \mathcal{I}$ y $\alpha \in \mathbb{R}$, con $A = \left[a^L, a^U\right]$ y $B = \left[b^L, b^U\right]$,

1. $A + B = \left[a^L + b^L, a^U + b^U\right]$
2. $$\alpha A = \begin{cases} \left[\alpha a^L, \alpha a^U\right] & \text{Si } \alpha \geq 0 \\ \left[\alpha a^U, \alpha a^L\right] & \text{Si } \alpha < 0 \end{cases}$$

Con estas operaciones en $\mathcal{I}$ es importante hacer notar que:

- Respecto de la operación $+$, el conjunto $\mathcal{I}$ satisface las propiedades: clausurativa, conmutativa, asociativa y modulativa, además se cumple la ley cancelativa, pero no es un espacio vectorial ya que con la operación algebraíca $+$ los elementos no siempre tienen inverso aditivo.

Adicionalmente, otra operación en el conjunto $\mathcal{I}$, es la diferencia de Hukuhara, que se define como sigue:

Definición 2.2.1. *Sean $\mathcal{K}_c(\mathbb{R}^m) = \{C \subseteq \mathbb{R}^m : C$ es conjunto compacto y convexo$\}$ y $A,\ B \in \mathcal{K}_c(\mathbb{R}^m)$, si existe $C \in \mathcal{K}_c(\mathbb{R}^m)$ tal que $A = B + C$, entonces C se llama la diferencia de Hukuhara, denotada por $C = A \ominus B$.*

Al aplicar la definición 2.2.1 en el conjunto $\mathcal{I}$, se tiene que con $A = \left[a^L, a^U\right]$, $B = \left[b^L, b^U\right]$ en $\mathcal{I}$, la diferencia de Hukuhara entre A y B si existe, es el intervalo $C = \left[c^L, c^U\right]$ el cual está definido por $C = \left[a^L - b^L, a^U - b^U\right]$.

Ahora, si definimos la longitud de un intervalo $A = \left[a^L, a^U\right] \in \mathcal{I}$, denotada por $l(A)$ como $l(A) = a^U - a^L$, se obtiene el siguiente resultado.

Proposición 2.2.1. *Sean $A = \left[a^L, a^U\right]$ y $B = \left[b^L, b^U\right] \in \mathcal{I}$, entonces $A \ominus B$ existe si y sólo si $l(A) \geq l(B)$.*

La prueba de la anterior proposición se puede encontrar en [6].

Por otra parte, al conjunto $\mathcal{I}$ también se le dota de una métrica, llamada la métrica de Hausdorff que lo hace un espacio métrico completo, y es definida como:

$$d_H(A,B) = \text{máx}\left\{\sup_{a\in A}\inf_{b\in B}\|a-b\|, \sup_{b\in B}\inf_{a\in A}\|a-b\|\right\}$$

Las pruebas sobre estas afirmaciones se pueden encontrar en [2], [3], [6], [11].

2.3. Órdenes Parciales

Los problemas de optimización matemática mono-objetivo, están definidos sobre el cuerpo de los reales, $\mathbb{R}$, el cual es totalmente ordenado, por esta razón en las presentaciones sobre problemas de este tipo de optimización, se obvia el estudio sobre ordenes parciales. En contraste, en los problemas de optimización matemática multiobjetivo, este tema toma gran relevancia, puesto que al espacio vectorial $\mathbb{R}^m$ no se le ha dotado de un orden total, lo que obliga a reflexionar sobre el significado de la expresión "minimizar o maximizar" una función $f\colon \mathbb{R}^n \to \mathbb{R}^m$. Esto es fundamental porque el concepto de minimizar o maximizar está ligado a ordenadar y poder decidir si un elemento a de un conjunto M antecede o no a otro elemento b del conjunto M. Dado que nuestro objetivo en este trabajo está en relación directa con problemas de optimización multiobjetivo, se hace necesario desarrollar algunos conceptos preliminares sobre ordenes parciales.

Definición 2.3.1. *Una relación binaria $\sim$ sobre un conjunto $N \subset M$ es un subconjunto de $M \times M$. Una relación binaria $\sim$ se dice:*

1. *Reflexiva si para todo $a \in N$, $a \sim a$.*
2. *Irreflexiva si para todo $a \in N$, $a \nsim a$.*
3. *Simétrica si para todo a, $b \in N$ y $a \sim b$ entonces $b \sim a$.*
4. *Asimétrica si para todo a, $b \in N$ y $a \sim b$ entonces $b \nsim a$.*
5. *Antisimétrica si para todo a, $b \in N$, $a \sim b$ y $b \sim a$ entonces $a = b$.*
6. *Transitiva si para todo a, b y $c \in N$, $a \sim b$ y $b \sim c$ entonces $a \sim c$.*

Definición 2.3.2. *Una relación binaria $\sim$ sobre un conjunto N se llama una relación de equivalencia si es reflexiva, simétrica y transitiva.*

Definición 2.3.3. *Una relación binaria $\preceq$ sobre un conjunto N se llama un preorden si es reflexiva y transitiva. El conjunto $(N, \preceq)$ se llama un conjunto preordenado.*

Con el preorden $\preceq$ existen dos relaciones binarias asociadas:

1. Para x, $y \in N$, $x \prec y$ si y sólo si $x \preceq y$ y $y \npreceq x$.
2. Para x, $y \in N$, $x \sim y$ si y sólo si $x \preceq y$ y $y \preceq x$.

Proposición 2.3.1. *Sea $\preceq$ un preorden sobre N. Entonces la relación $\prec$ descrita en 1. es irreflexiva y transitiva y la relación $\sim$ dada en 2. es una relación de equivalencia.*

Proposición 2.3.2. *Una relación asimétrica es irreflexiva. Una relación transitiva e irreflexiva es asimétrica.*

Definición 2.3.4. *Una relación binaria $\preceq$ se llama:*

I. *Un orden parcial si es un preorden y es antisimétrica, es decir, reflexiva, transitiva y antisimétrica.*

II. *Un orden parcial estricto si es asimétrica y transitiva, o también, irreflexiva y transitiva.*

En [2] se presentan los ordenes parciales $\preceq_{LU}$, $\preceq_{UC}$ y $\preceq_{CW}$ para $\mathcal{I}$ que provienen de [4]. Para sus definiciones es necesario considerar los siguientes elementos: a un intervalo $A = \left[a^L, a^U\right] \in \mathcal{I}$ le asociamos la pareja (a_c, a_w), donde $a_c = \frac{a^U + a^L}{2}$ es el centro del intervalo y $a_w = \frac{a^U - a^L}{2}$ es la longitud media del intervalo. Los ordenes parciales se definen como sigue:

Definición 2.3.5. *Sean $A = \left[a^L, a^U\right]$ y $B = \left[b^L, b^U\right] \in \mathcal{I}$:*

1. *$A \preceq_{LU} B$ si y sólo si $a^L \leq b^L$ y $a^U \leq b^U$.*
2. *$A \preceq_{CW} B$ si y sólo si $a_c \leq b_c$ y $a_w \leq b_w$.*
3. *$A \prec_{LU} B$ si y sólo si $A \preceq_{LU} B$ y $A \neq B$.*
4. *$A \prec_{CW} B$ si y sólo si $A \preceq_{CW} B$ y $A \neq B$.*
5. *$A \preceq_{UC} B$ si y sólo si $a^U \leq b^U$ y $a_c \leq b_c$.*
6. *$A \prec_{UC} B$ si y sólo si $A \preceq_{UC} B$ y $A \neq B$*
7. *$A \ll_{LU} B$ si y sólo si $a^L < b^L$ y $a^U < b^U$, orden parcial estricto $\ll_{LU}$.*
8. *$A \ll_{CW} B$ si y sólo si $a_c < b_c$ y $a_w < b_w$, orden parcial estricto $\ll_{CW}$.*

Ishibuchi y Tanaka en [4] probaron que:

(I) $A \preceq_{UC}$ si y sólo si $A \preceq_{LU} B$ o $A \preceq_{CW} B$.

(II) $A \prec_{UC}$ si y sólo si $A \prec_{CW} B$ o $A \prec_{CW} B$.

En el segundo orden parcial definido antes, el intervalo con menor centro y menor longitud media es preferido en los problemas de minimización.

2.4. Resultados Importantes sobre Optimización Intervalo-Valuada

En esta sección presentaremos algunos teoremas de optimización de funciones mono-objetivo intervalo-valuadas, que provienen de [2] y [1] y en los cuales nos hemos apoyado para obtener resultados similares para optimización de funciones multi-objetivo intervalo-valuadas.

2.4.1. Optimización con Funciones Objectivo Intervalo-Valuada

Formulación del Problema de Optimización y Conceptos de Solución

En este trabajo estudiamos solo los problemas de minimización sin perdida de generalidad. Consideremos el siguiente problema de optimización con función objetivo intervalo-valuada.

$$\begin{array}{c} \text{mín}\, f(x) = \left[f^L(x), f^U(x) \right] \\ \text{s.a} \;\; x \in \mathcal{F} \subseteq \mathbb{R}^n \end{array} \tag{2.1}$$

donde el conjunto factible $\mathcal{F}$ es un subconjunto convexo de $\mathbb{R}^n$.
Dado que el valor objetivo $f(x)$ es un intervalo cerrado, acotado y convexo, es necesario interpretar lo que significa minimizar en el problema (2.1).

Con los ordenes parciales "$\preceq_{LU}$" "$\preceq_{CW}$" y "$\preceq_{UC}$" en $\mathcal{I}$, definidos en (2.3.5) podemos usar un concepto de solución similar a la solución optima de Pareto, usada en los problemas de optimización multiobjetivo para interpretar el significado de la minimizar en el problema (2.1).

Definición 2.4.1. *Sea x^* una solución factible, es decir, $x^* \in \mathcal{F}$. Decimos que x^* es una* solución de tipo I *del problema (2.1) si no existe $\overline{x} \in \mathcal{F}$ tal que $f(\overline{x}) \prec_{CW} f(x^*)$.*

Definición 2.4.2. *Sea x^* una solución factible, es decir, $x^* \in \mathcal{F}$. Decimos que x^* es un* solución de tipo II *del problema 2.1 si no existe $\overline{x} \in \mathcal{F}$ tal que $f(\overline{x}) \prec_{LU} f(x^*)$ o $f(\overline{x}) \prec_{CW} f(x^*)$.*

Observación. Sea x^* una solución factible. Vemos que si x^* es un solución de tipo I del problema 2.1 entonces x^* es también solución de tipo II del problema 2.1.

La anterior definición, recordando el orden $\prec_{UC}$, se puede reescribir de la siguiente manera:

Definición 2.4.3. *Sea x^* una solución factible. Decimos que x^* es una* solución de tipo II *del problema 2.1 si no existe $\overline{x} \in \mathcal{F}$ tal que $f(\overline{x}) \prec_{UC} f(x^*)$.*

Ejemplo 2.4.1. *Sea $f\colon [0,2] \to \left[x^2+x+1, x^2+2\right]$ una función mono-objetivo intervalo valuada definida sobre $[0,2]$. Observe que $x^* = 0$ es una solución de tipo I del problema, pues $1 < x^2+x+1$ y $2 < x^2+2$ para todo $x \neq 0$. Por otra parte, tomemos $f_c(x) = \left(f^L(x) + f^U(x)\right)/2 = x^2+x/2+2$ y $f_w(x) = \left(f^L(x) - f^U(x)\right)/2 = 1 - x/2$. Ahora note que $x^* = 1/2$ es una solución tipo II pero no es una solución tipo I. Veamos que es una solución tipo II. Suponga que $x^* = 1/2$ no es solución de tipo II, entonces existe un $\overline{x} \in [0,2]$ tal que $f(\overline{x}) \prec_{UC} f(x^*)$. Esto significa que*

$$\left\{ \begin{array}{l} f_c(\overline{x}) < f_c(1/2) \\ f_w(\overline{x}) \leq f_w(1/2) \end{array} \right. \quad \text{ó} \quad \left\{ \begin{array}{l} f_c(\overline{x}) \leq f_c(1/2) \\ f_w(\overline{x}) < f_w(1/2) \end{array} \right.$$

entonces se sigue que $6f_c(\overline{x}) + 2f_w(\overline{x}) < 6f_c(1/2) + 2f_w(1/2)$, simplificando y factorizando se obtiene que $2(\overline{x} - 1/2)^2 < 0$; contradicción. Esto muestra que $x^* = 1/2$ es una solución de tipo II.

Condiciones de Optimalidad Karush-Kuhn-Tucker

Iniciemos con un teorema básico sobre el problema clásico de optimización, esto es, para funciones de valor real.

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s. a.} & g_i(x) \leq 0,\; i = 1, \cdots, k \end{array} \tag{2.2}$$

Suponga que las funciones restricción g_i son convexas sobre $\mathbb{R}^n$ para cada $i = 1, \cdots, k$. Entonces el conjunto factible es $\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \cdots, k\}$ es un conjunto convexo de $\mathbb{R}^n$. Las conocidas condiciones de Karush-Kuhn-Tucker para el problema (2.2) se establecen en el siguiente resultado.

Teorema 2.4.1. *Asuma que las funciones restricción $g_i \colon \mathbb{R}^n \to \mathbb{R}$ son convexas sobre $\mathbb{R}^n$ para $i = 1, \cdots, k$. Sean $\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \cdots, k\}$ el conjunto factible y $x^* \in \mathcal{F}$ un punto factible. Suponga que las función objetivo $f \colon \mathbb{R}^n \to \mathbb{R}$ es convexa en x^*, y f, g_i, $i = 1; \cdots, k$, son continuamente diferenciables en x^*. Si existen multiplicadores $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \cdots, k$ tales que:*

(I) $\nabla f\left(x^*\right) + \sum_{i}^{k} \mu_i \nabla g_i\left(x\right) = \mathbf{0}$;

(II) $\mu_i \nabla g_i\left(x^*\right) = 0$ *para todo* $i = 1, \cdots, k$,

entonces x^* es una solución óptima del problema (2.2)

Ahora consideremos el siguiente problema de minimización intervalo valuado:

Sea $f \colon \mathbb{R}^n \to \mathcal{I}$ una función intervalo valuada.

$$\begin{array}{ll} \text{mín} & f\left(x\right) = \left[f^L(x), f^U(x)\right] \\ \text{s. a.} & g_i\left(x\right) \leq 0,\ i = 1, \cdots, k \end{array} \tag{2.3}$$

donde las funciones restricciones de valor real $g_i \colon \mathbb{R}^n \to \mathbb{R}$ son funciones convexas sobre $\mathbb{R}^n$ para $i = 1, \cdots, k$. Notemos que el problema (2.3) es similar al problema (2.1) al tomar el conjunto convexo $\mathcal{F}$ como $\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \cdots, k\}$.

Condiciones KKT para la Diferenciabilidad Débil y Fuerte

Antes de pasar a la discusión sobre las condiciones KKT para el problema 2.3, definamos el concepto de convexidad para funciones intervalo valuadas.

Definición 2.4.4. *Sea $f\left(x\right) = \left[f^L(x), f^U(x)\right]$ una función intervalo valuada definida sobre un conjunto convexo $D \in \mathbb{R}^n$. Decimos que f es* LU-convexa en x^* *si:*

$$f\left(\lambda x^* + (1-\lambda) x^*\right) \preceq_{LU} \lambda f\left(x^*\right) + (1-\lambda) f\left(x^*\right) \tag{2.4}$$

para cada $\lambda \in (0,1)$ y cada $x^* \in D$. Similarmente, podemos definir la UC-convexidad en x^* *al usar la relación de orden '$\preceq_{UC}$' en* (2.4).

Proposición 2.4.1. *Sea D un subconjunto convexo de $\mathbb{R}^n$ y f una función intervalo valuada definida sobre D.Entonces para $x^* \in D$ tenemos las siguientes propiedades:*

(I) *f es LU-convexa en x^* si y sólo si f^L y f^U son convexas en x^*.*

(II) *f es UC-convexa en x^* si y sólo si f^U y f_c son convexas en x^*.*

(III) *f es LU-convexa en x^* entonces f también es UC-convexa en x^*.*

Sea $\mathcal{F} = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \cdots, k\}$ la región factible del problema (2.3) y un punto x^* en $\mathcal{F}$. Decimos que las funciones restricciones reales $g_i(x) \leq 0, i = 1, \cdots, k$ satisfacen la hipótesis KKT en x^*, si las g_i son convexas sobre $\mathbb{R}^n$ y continuamente diferenciables en x^* para todo $i = 1, \cdots, k$, con esto se sigue:

Teorema 2.4.2. *Suponga que las funciones restricciones reales $g_i(x) \leq 0, i = 1, \cdots, k$ del problema (2.3) satisfacen la hipótesis KKT en x^* y la función objetivo intervalo valuada $f\colon \mathbb{R}^n \to \mathcal{I}$ es LU-convexo y débilmente continuamente diferenciable en x^*,entonces, si existen multiplicadores $0 < \lambda^L,\ \lambda^U \in \mathbb{R}$ y $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \cdots, k$ tales que:*

(I) $\lambda^L \nabla f^L(x^*) + \lambda^U \nabla f^U(x^*) + \sum_{i=1}^{k} \mu_i \nabla g_i(x) = \mathbf{0}$;

(II) $\mu_i \nabla g_i(x^*) = 0$ *para todo* $i = 1, \cdots, k$,

entonces x^* es una solución óptima de tipo I y de tipo II del problema (2.3)

Análogamente, tenemos:

Teorema 2.4.3. *Bajo las misma hipótesis del teorema 2.4.2. Sea $f_c = \dfrac{1}{2}\left(f^L + f^U\right)$. Si existen multiplicadores $0 < \lambda^U,\ \lambda_c \in \mathbb{R}$ y $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \cdots, k$ tales que:*

(I) $\lambda^U \nabla f^U(x^*) + \lambda_c \nabla f_c(x^*) + \sum_{i=1}^{k} \mu_i \nabla g_i(x) = \mathbf{0}$;

(II) $\mu_i \nabla g_i(x^*) = 0$ *para todo* $i = 1, \cdots, k$,

entonces x^* es una solución óptima de tipo I y de tipo II del problema (2.3)

y para la H-diferenciabilidad se tiene:

Teorema 2.4.4. *Suponga que las funciones restricciones reales $g_i(x) \leq 0, i = 1, \cdots, k$ del problema (2.3) satisfacen la hipótesis KKT en x^* y la función objetivo intervalo valuada $f\colon \mathbb{R}^n \to \mathcal{I}$ es LU-convexo y continuamente H-diferenciable en x^*, si existen multiplicadores $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \cdots, k$ tales que:*

(I) $\nabla f(x^*) + \sum_{i=1}^{k} \mu_i \nabla g_i(x) = \mathbf{0}$;

(II) $\mu_i \nabla g_i(x^*) = 0$ *para todo* $i = 1, \cdots, k$,

entonces x^* es una solución óptima de tipo I y de tipo II del problema (2.3)

2.4.2. Otros resultados sobre las restricciones en el problema 2.3

Hasta el momento hemos tratado algunos de los temas de la optimización con función monoobjetivo intervalo valuada descritos en [2], en donde las funciones resctricción son funciones real valuadas. El profesor Hsien-Chung Wu ha seguido avanzando en este tema y ha considerado a las funciones restricción también como funciones intervalo valuadas en el articulo [1], lo que implica que se debe tener en cuenta no sólo un orden parcial apropiado, $\prec$, en la función mono-objetivo sino también, otro orden parcial apropiado, $\widehat{\preceq}$, en las funciones restricción, de tal forma que el problema de optimización intervalo valuado se describe como sigue:

$$\begin{array}{ll} \text{mín} & f(x) = \left[f^L(x), f^U(x)\right] \\ \text{s. a.} & \begin{array}{l} G_i(x) \,\widehat{\preceq}\, \left[b_i^L, b_i^U\right],\ i = 1, \cdots, k \\ \boldsymbol{x} \geq \mathbf{0} \end{array} \end{array} \tag{2.5}$$

donde $G_i(x) = \left[g_i^L(x), g_i^U(x)\right]$ son funciones restricciones intervalo valuadas para $i = 1, \cdots, k$, y "mín", según el orden parcial estricto $\prec$.

De los problemas de optimización dados por (2.5), nos interesan los que se plantean a partir de los ordenes parciales $\preceq_{LU}$ y el orden parcial estricto asociado, y en este caso el problema se reesscribe como:

$$\begin{array}{ll} \text{mín} & f(x) = \left[f^L(x), f^U(x)\right] \\ \text{s. a.} & \begin{array}{l} G_i(x) \preceq_{LU} \left[b_i^L, b_i^U\right],\ i = 1, \cdots, k \\ \boldsymbol{x} \geq \mathbf{0} \end{array} \end{array} \tag{2.6}$$

De otro lado, si se define el siguiente problema auxiliar, de optimización con función mono-objetivo intervalo valuada:

$$\begin{array}{ll} \text{mín} & f(x) = \left[f^L(x), f^U(x)\right] \\ \text{s. a.} & \begin{array}{l} g_i^L(x) \leq b_i^L,\ i = 1, \cdots, k \\ g_i^U(x) \leq b_i^U,\ i = 1, \cdots, k \\ \boldsymbol{x} \geq \mathbf{0} \end{array} \end{array} \tag{2.7}$$

entonces, en [1] se prueba que:

Proposición 2.4.2. *Suponga que los problemas* (2.6) *y* (2.7) *usan el mismo concepto de solución, entonces* (2.6) *y* (2.7) *tienen las mismas soluciones óptimas.*

De la anterior proposición, se concluye que los problemas de optimización con función objetivo intevalo valuada y funciones restricciones intervalo valuadas, puede resolverse a través del problema auxiliar 2.7, donde las funciones restricciones son funciones real valuadas, y por tanto se puede estudiar a partir de las condiciones dadas en el teorema 2.4.4 para la solución.

2.5. Optimización Multiobjetivo

2.5.1. Definiciones generales.

Un problema de decisión que involucra multiples criterios, consiste en determinar el mejor compromiso de solución, y podría establecerse de la siguiente manera:

$$\begin{array}{ll} \text{mín} & \{f_1(x), \cdots, f_m(x)\} \\ \text{s. a.} & x \in \mathcal{F} \subseteq \mathbb{R}^n \end{array} \tag{2.8}$$

donde:

- $\mathcal{F}$ denota el conjunto de alternativas potenciales o soluciones factibles.
- $\{f_1(x), \cdots, f_m(x)\}$ representa un conjunto de m ($m \geq 2$) funciones de valor real, llamadas criterios.

La expresión (2.8) no es realista, normalmente, no existe una alternativa x que optimice todos los criterios simultaneamente. La notación mín indica entonces que estamos buscando el mejor compromiso de solución, de acuerdo a una estructura de preferencia del decisor tomando en cuenta cada uno de los m criterios. La optimización multiobjetivo (MOP, siglás en inglés) se relaciona con los problemas de decisión con multiples criterios en los siguientes aspectos:

- Cada alternativa se caracteriza por un vector $(x_1, \cdots, x_n)^\top = x$ $(x \in \mathbb{R}^n)$ de variables decisión $x_1, \cdots, x_n$.
- El conjunto $\mathcal{F}$ se define en términos de las variables de decisión:

  $$\mathcal{F} = \{x \in \mathbb{R}^n / g_i(x) \leq 0,\ i = 1, \cdots, k \text{ y } \ x \in S\}$$

  donde $g_i(x)$, $i = 1, \cdots, k$ son funciones restricción y $S \subset \mathbb{R}^n$, es usado para representar restricciones adicionales, las cuales no pueden ser expresadas a tráves de funciones, por ejemplo, $S = \mathbb{Z}^n$ ó $S = \{0, 1\}^n$.
- Cada criterio f_i puede ser expresado como una función de las variables de decisión, y se le llama la i-ésima función objetivo.

En general, la expresión de un problema MOP es:

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s. a.} & \begin{array}{l} g_i(x) \leq 0,\ i = 1, \cdots, k \\ x \in S \end{array} \end{array} \tag{2.9}$$

donde $f\colon \mathbb{R}^n \to \mathbb{R}^m$ es una función con $f(x) = (f_1(x), \cdots, f_m(x))^\top$.

Este campo de la optimización ha recibido una atención considerable, como una extensión natural de la optimización matemática clásica (mono-objetivo).

La anterior formulación esta presentada en **el espacio de variables de decisión**, esta es la representación clásica en optimización matemática. Por otra parte, los conceptos básicos y definiciones pueden ser introducidos independientemente usando **el espacio criterio** $\mathcal{Z}_\mathcal{F} = f(\mathcal{F}) = \{z \in \mathcal{Z} : z = f(x) \text{ para algún } x \in \mathcal{F}\}$. En este espacio la imagen de cada punto factible x o alternativa es representada por un arreglo $m \times 1$ de los valores de las funciones criterio, f_j, así, $f(x) = (f_1(x), \cdots, f_m(x))^\top$. Debemos notar que el espacio criterio es más significativo que el espacio de decisión, en el contexto de optmimización multi-objetivo, ya que el interés de cada alternativa debe ser apreciada solamente respecto a sus valores sobre los criterios.

Sean $Z \subset \mathbb{R}^m$ el espacio de criterios y $Z_\mathcal{F}$ la imagen de $\mathcal{F}$, es decir, la imagen del conjunto factible. Entonces tenemos:

$$Z_\mathcal{F} = \{z \in Z / z_j = f_j(x) \quad \text{donde} \quad x \in \mathcal{F}\}. \tag{2.10}$$

Al generar un concepto de orden parcial en el espacio criterio para decidir cual es el mejor compromiso de vector de criterio solución $\boldsymbol{z}$ de $Z_\mathcal{F}$, podemos plantear el problema en la forma:

$$\begin{array}{ll} \text{mín} & (z_1, \cdots, z_m)^\top = \boldsymbol{z} \\ \text{s. a.} & \boldsymbol{z} \in Z_\mathcal{F} \end{array} \tag{2.11}$$

Sin más información acerca de la estructura de preferencias del decisor, la comparación entre vectores criterios, z, puede ser basada sobre el orden parcial natural [1] definido sobre Z:

Definición 2.5.1. *Para todo* $\boldsymbol{z}$ *y* $\boldsymbol{z\prime} \in Z$ *tenemos:*

- $\boldsymbol{z} \leq \boldsymbol{z'}$ *si y sólo si* $\boldsymbol{z_j} \leq \boldsymbol{z'_j}$ *para todo* $j = 1, \cdots, m$.
- $\boldsymbol{z} < \boldsymbol{z'}$ *si y sólo si* $\boldsymbol{z} \leq \boldsymbol{z'}$ *y* $\boldsymbol{z} \neq \boldsymbol{z'}$.
- $\boldsymbol{z} \ll \boldsymbol{z'}$ *si y sólo si* $\boldsymbol{z_j} < \boldsymbol{z'_j}$ *para todo* $j = 1, \cdots, m$.

Definición 2.5.2. *Sea* $\boldsymbol{z} \in Z_{\mathcal{F}}$ *se llama* no dominado *si y sólo si no existe* $\boldsymbol{z'} \in Z_{\mathcal{F}}$ *tal que* $\boldsymbol{z'} < \boldsymbol{z}$.

Definición 2.5.3. *Sea* $\boldsymbol{z} \in Z_{\mathcal{F}}$ *se llama* débilmente no dominado *si y sólo si no existe* $\boldsymbol{z'} \in Z_{\mathcal{F}}$ *tal que* $\boldsymbol{z'} \ll \boldsymbol{z}$.

Las soluciones correspondientes a puntos (débilmente) no dominados son llamadas soluciones (débilmente) eficientes. Las soluciones eficientes $\boldsymbol{z}$ también se conocen como óptimos de Pareto y a los puntos $x \in \mathcal{F}$ tales que $\boldsymbol{z} = f(x)$ puntos no dominados. Estos conjuntos generalmente no tienen una caracterización, sin embargo, algunos conceptos son útiles para guiar la construcción de los mismos.

2.5.2. Funciones Scalarizing

Una función scalarizing es un función s: $Z_{\mathcal{F}} \times \Omega \rightarrow \mathbb{R}$, donde Ω es un conjunto de parametros dados por la estructura de preferencia del decisor.

Debemos enfatizar que tal función debe ser considerada como una herramienta usada para generar soluciones. Este proceso de generación es simplemente ejecutado al seleccionar un parámetro específico $\omega' \in \Omega$, y optimizando la función sobre $Z_{\mathcal{F}}$ o algún subconjunto, este problema se puede formular como:

$$\begin{array}{ll} \text{mín} & s(\boldsymbol{z}, \omega') \\ \text{s. a.} & \boldsymbol{z} \in Z_{\mathcal{F}} \end{array} \tag{2.12}$$

un ejemplo clásico de función scalarizing es *weighted sum*, dada por:

$$s(\boldsymbol{z}, \boldsymbol{\lambda}) = \sum_{j=1}^{m} \lambda_j z_j \tag{2.13}$$

donde $\lambda \in \mathbb{R}^m = \Omega$. Hay otros tipos de funciones scalarizing y la elección depende de la estructura de preferencia del decisor.

Para las funciones scalarizing hay dos requerimientos básicos:

1. s genera solamente puntos no dominados.
2. cualquier punto no dominado podría ser generado por s

y como requerimiento práctico, s no debería involucrar proceso computacionales pesados.

Considerando la función *weighted sum*, tenemos:

[1] Si $z, z' \in \mathbb{R}^m$, con $z = (z_1, z_2, \cdots, z_m)$ y $z' = (z'_1, z'_2, \cdots, z'_m)$, $z \leq z'$ si sólo si $z_j \leq z'_j$, para todo $j = 1, \cdots, m$

Teorema 2.5.1. *Considerando el problema:*

$$\begin{array}{ll} \text{mín} & \sum_{j=1}^{m} \lambda_j z_j \\ s.\ a. & \boldsymbol{z} \in Z_{\mathcal{F}} \end{array} \tag{2.14}$$

1. a) *Si* $\boldsymbol{z\prime}$ *es una solución óptima de* (2.14) *con* $\boldsymbol{\lambda} \gg \mathbf{0}$, *entonces* $\boldsymbol{z'}$ *es no dominado.*
   b) *Si* $\boldsymbol{z'}$ *es la única solución óptima de* (2.14) *con* $\boldsymbol{\lambda} > \mathbf{0}$, *entonces* $\boldsymbol{z'}$ *es no dominado.*
   c) *Si* $\boldsymbol{z'}$ *es una solución óptima de* (2.14) *con* $\boldsymbol{\lambda} > \mathbf{0}$, *entonces* $\boldsymbol{z'}$ *es débilmente no dominado.*

2. *Si* $\boldsymbol{z'}$ *es no dominado y* $Z_{\mathcal{F}}$ *es* convexo, *entonces existe* $\boldsymbol{\lambda} > \mathbf{0}$ *tal que* $\boldsymbol{z'}$ *es una solución óptima de* (2.14).

2.6. Teorema de Rådström.

Los teoremas de embebimiento son una herramienta muy útil en el estudio del álgebra y la topología algebráica, el sentido práctico nació con el álgebra, pero su uso se extendió por todas las áreas de la matemática. Hans Rådström publicó un artículo en la revista *Proceedings of the American Mathematical Society*, Vol. 3, $N°$.1 (Feb., 1952), pp 165-169, titulado **An Embedding Theorem for Spaces of Convex Sets** [5] del cual se basa esta sección. El tema de este artículo fue planteado para otras aplicaciones, pero Banks y Jacobs [6] usaron este resultado para crear un cálculo diferencial para lo que denominaron multifuncines, además de otras dos formas de generar la diferencial. Este proyecto fue apoyado por la NASA y por la fuerza aerea de los Estados Unidos de America, y su aplicación principal fue hacia problemas de control óptimo con multifunciones.

Desde la teoría del análisis convexo se tiene que para un espacio topológico lineal real M, si S y R son subconjuntos convexos en M y si λ es un número real, los subconjuntos $S+R$ y λS están bien definidos y son convexos en M. Aquí $S + R = \{z|z = x + y,\ x \in S,\ y\ \in R\}$, $\lambda S = \{z|z = \lambda x,\ x \in S\}$. Estas operaciones satisfacen para $S,\ R,\ Z \in M$ y $\lambda,\ \mu \in \mathbb{R}$:

1. $(S + R) + Z = S + (R + Z)$
2. $S + R = R + S$
3. $\lambda (S + R) = \lambda S + \lambda R$
4. $\lambda (\mu S) = (\lambda \mu S)$
5. $1S = S$

Así, el conjunto de todos los subconjuntos convexos de M es un semigrupo conmutativo bajo la suma. Si la situación hubiera sido tal que no sólo fuera un semigrupo sino también un grupo, y además se cumpliera que $(\lambda + \mu) S = \lambda S + \mu S$, entonces el conjunto de todos los subconjuntos convexos podría haber sido un espacio vectorial. Sin embargo, esto es cierto sólo si λ y μ tienen el mismo signo, en particular si ambos son positivos. Esto nos lleva a una pregunta natural, ¿Cuales son las condiciones bajo las cuales un semigrupo conmutativo bajo la operación suma, puede ser embebido en un grupo y bajo cuales condiciones la multiplicación por escalar puede

ser extendida en este grupo de tal manera que el sistema resultante sea un espacio vectorial?, la pregunta anterior es resuelta en el artículo [5], usando un método clásico para extender semigrupos conmutativos, el cual es usado en la construcción de los números negativos.
Enunciamos el teorema que muestra las condiciones bajo las cuales se realiza la extensión antes mencionada.

Teorema 2.6.1. *Teorema de Extension de Semigrupo*

- *Sea M un semigrupo conmutativo en el cual la ley cancelativa se cumple, esto es:*
  Para S, R, $Z \in M$, tenemos:

  1. $(S+R)+Z = S+(R+Z)$,
  2. $S+R = R+S$,
  3. *si* $S+Z = R+Z$ *entonces* $S = R$.

  Entonces M puede ser embebido en un grupo N. Además N puede ser elegido minimal en el siguiente sentido: Si G es cualquier grupo en el cual M es embebido, entonces N es isomorfico a un subgrupo de G que contiene a M.

- *Si existe una multiplicación por escalar real no negativo en M y satisface:*
  4. $\lambda(S+R) = \lambda S + \lambda R$,
  5. $(\lambda+\mu)S = \lambda S + \mu S$,
  6. $\lambda(\mu S) = \lambda\mu S$,
  7. $1S = S$,
  entonces una multiplicación por escalar real puede ser definida en N tal que convierte a N en un espacio vectorial y así para $\lambda \geq 0$ y $S \in M$ el producto λS coincide con el dado en M.

- *Si además una métrica $d(S,R)$ esta dada en M y satisface que:*
  8. $d(S+Z, R+Z) = d(S,R)$,
  9. $d(\lambda S, \lambda R) = \lambda d(S,R)$,
  10. $S+R$ *y* λS *son operaciones continuas en la topología inducida por d en M,*
  entonces una métrica puede ser definida en N y así convierte a N en un espacio vectorial normado y es tal que si S, $R \in M$, la distancia entre S y R es igual a $d(S,R)$.

Los siguientes resultados conocidos, son fundamentales tanto para el teorema de embebimiento de Rådström, enunciado con la métrica d_H, como para su aplicación en nuestro trabajo, con un métrica definida en el conjunto $\mathcal{I}^m$ y que cumple con las mismas propiedades de la métrica d_H en su enunciado, además de ser equivalente a la métrica d_H.

Lema 2.6.1. *Sean S, Z y R conjuntos dados en un espacio lineal normado real. Suponga que Z es cerrado y convexo, R es acotado, y que $S+R \subset Z+R$, entonces $S \subset Z$.*

La invarianza de la metrica de Hausdorff es dada a través del siguiente lema.

Lema 2.6.2. *Sean S y Z conjuntos convexos en un espacio lineal normado M. También suponga que $S+\lambda E$ y $Z+\lambda E$ son cerrados para todo $\lambda \geq 0$, donde E es la esfera unidad. Sea R cualquier conjunto cerrado en M. Entonces $d_H(S,Z) = d_H(S+R, Z+R)$.*

Combinando los resultados anteriores, Rådström probó en [5], lo que ahora se denomina el teorema de embebimiento de Rådström y que enunciamos a continuación.

Teorema 2.6.2 (Teorema de Extension de Rådström)**.** *Sea M un espacio lineal normado real, si L cualquier espacio de puntos los cuales son conjuntos cerrados, acotados y convexos en M, el cual tiene las siguientes propiedades:*

1. *L es cerrado bajo la adicción y multiplicación por escalar no negativo,*

2. *Si $R \in L$ y S es la esfera unitaria de M, entonces $R + S$ es cerrado,*

3. *L es métrizado por la métrica de Hausdorff,*

entonces, L puede ser embebido como un cono convexo en un espacio normado real N de tal manera que:

a. *El embebimiento es isométrico.*

b. *La suma en L induce la suma en N.*

c. *La multiplicación por escalar no negativo en L induce la correspondiente operación en N.*

Además, N puede ser elegido minimal en el siguiente sentido: Si H es cualquier espacio lineal normado real, en el cual L es embebido en el sentido anterior, entonces H contiene un subespacio que contiene a L y es isomorfo a N.
Los siguiente dos ejemplos satisfacen las condiciones impuestas sobre L.

α **.** *El conjunto de todos los conjuntos convexos compactos de dimensión finita.*

β **.** *El conjunto de todos los conjuntos convexos y compactos.*

Capítulo 3

Optimización de Funciones Multi-Intervalo Valuadas.

En este capítulo desarrollaremos la teoría necesaria para generar conceptos de solución, y métodos numéricos que permiten aproximar soluciones del problema de optimización multi-intervalo valuado, haciendo un derrotero teórico similar al presentado por Wu en [2].

3.1. Algebra y Ordenes en el Conjunto $\mathcal{I}^m$

En este trabajo se desarrollará una estructura similar a la del conjunto $\mathcal{I}$ para el conjunto $\mathcal{I}^m$, necesario para el estudio de los problemas de optimización multiobjetivo intervalo valuados.

Definimos $\mathcal{I}^m = \{I_1 \times \cdots \times I_m : I_j \in \mathcal{I}\}$ para todo $j = 1, \cdots, m$ y por abuso de lenguaje, lo representaremos como un arreglo $m \times 1$ de intervalos, esto es,

$$\mathcal{I}^m = \left\{ \begin{bmatrix} I_1 \\ \vdots \\ I_m \end{bmatrix} \subseteq \mathbb{R}^m : I_j \in \mathcal{I}, \text{ para todo } j = 1, \cdots, m \right\}$$

sobre este conjunto, cada elemento lo llamaremos un hiperrectángulo o multi-intervalo. Observe que si $A \in \mathcal{I}^m$, entonces

$$\begin{aligned} A &= \begin{bmatrix} [a_1^L, a_1^U] \\ \vdots \\ [a_m^L, a_m^U] \end{bmatrix} \\ &= \left\{ x \in \mathbb{R}^m : x = (x_1, x_2, \cdots, x_j, \cdots, x_m)^t, \ \ a_j^L \leq x_j \leq a_j^U, \text{ para todo } j = 1, \cdots, m \right\} \subseteq \mathbb{R}^m. \end{aligned}$$

Decimos también que si $A, B \in \mathcal{I}^m$ con $A = \begin{bmatrix} [a_1^L, a_1^U] \\ \vdots \\ [a_m^L, a_m^U] \end{bmatrix}$ y $B = \begin{bmatrix} [b_1^L, b_1^U] \\ \vdots \\ [b_m^L, b_m^U] \end{bmatrix}$

$$A = B \text{ si y sólo si } a_j^L = b_j^L \text{ y } a_j^U = b_j^U \text{ para todo } j = 1, \cdots, m$$

En $\mathcal{I}^m$ definimos las operaciones suma y producto por escalar como sigue:

Sean A, $B \in \mathcal{I}^m$ y $\alpha \in \mathbb{R}$

1. $A+B=\begin{bmatrix} A_1+B_1 \\ \vdots \\ A_m+B_m \end{bmatrix}$ donde $A_j=\left[a_j^L,a_j^U\right]$ y $B_j=\left[b_j^L,b_j^U\right]$.
   dado que cada $A_j+B_j\in\mathcal{I}$, para todo $j=1,\cdots,m$ entonces $A+B\in\mathcal{I}^m$.

2. $\alpha A=\begin{bmatrix} \alpha A_1 \\ \vdots \\ \alpha A_m \end{bmatrix}$
   nuevamente, $\alpha A_j\in\mathcal{I}$, para todo $j=1,\cdots,m$ luego, $\alpha A\in\mathcal{I}^m$.

Con lo anterior, las operaciones suma y producto por escalar son clausurativas sobre $\mathcal{I}^m$. Adicionalmente se tiene el siguiente resultado:

Proposición 3.1.1. *Sean A,B y $C\in\mathcal{I}^m$ y $\alpha,\lambda\in\mathbb{R}$ no negativos, con* $A=\begin{bmatrix} [a_1^L,a_1^U] \\ \vdots \\ [a_m^L,a_m^U] \end{bmatrix}$,

$B=\begin{bmatrix} [b_1^L,b_1^U] \\ \vdots \\ [b_m^L,b_m^U] \end{bmatrix}$ *y* $C=\begin{bmatrix} [c_1^L,c_1^U] \\ \vdots \\ [{}_m^L,c_m^U] \end{bmatrix}$

La operación + satisface las propiedades:

P.-1 *Asociatividad.* $(A+B)+C=A+(B+C)$

P.-2 *Conmutatividad.* $A+B=B+A$

P.-3 *Elemento neutro.* $\mathbf{0}=\begin{bmatrix} [0,0] \\ \vdots \\ [0,0] \end{bmatrix}$ *tal que* $A+\mathbf{0}=\mathbf{0}+A=A$

La operación producto por escalar satisface las propiedades:

P-4. *Asociatividad.* $\alpha(\lambda A)=(\alpha\lambda)A$

P-5. *Elemento neutro.* $1\in\mathbb{R}$, $1A=A$

y por último, las leyes distributivas de la suma y el producto por escalar.

P-6. *Distributividad con la suma.* $\alpha(A+B)=\alpha A+\alpha B$

P-7. *Distributividad con la suma escalar.* $(\alpha+\lambda)A=\alpha A+\lambda A$

Demostración. Supongamos A,B y $C\in\mathcal{I}^m$ y $\alpha,\lambda\in\mathbb{R}$, no negativos, entonces:

P.-1 $(A+B)+C=\left(\begin{bmatrix} A_1+B_1 \\ \vdots \\ A_m+B_m \end{bmatrix}\right)+\begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix}=\begin{bmatrix} (A_1+B_1)+C_1 \\ \vdots \\ (A_m+B_m)+C_m \end{bmatrix}$ pero dado que $(A_j+B_j)+C_j\in\mathcal{I}$ para todo $j=1,\cdots,m$ se cumple que

$$(A_j+B_j)+C_j=\left[\left(a_j^L+b_j^L\right)+c_j^L,\left(a_j^U+b_j^U\right)+c_j^U\right] \text{ para todo } j=1,\cdots,m.$$

pero como las componentes satisfacen la propiedad asociativa, entonces podemos reescribir

$$(A_j + B_j) + C_j = A_j + (B_j + C_j) \text{ para todo } j = 1, \cdots, m.$$

así se concluye la asociatividad.

P.-2 $A + B = \begin{bmatrix} A_1 + B_1 \\ \vdots \\ A_m + B_m \end{bmatrix}$ pero dado que por definición $A_j + B_j \in \mathcal{I}$ para todo $j = 1, \cdots, m$ se cumple que

$$A_j + B_j = \left[a_j^L + b_j^L, a_j^U + b_j^U\right] \text{ para todo } j = 1, \cdots, m.$$

pero como las componentes satisfacen la propiedad conmutativa, entonces podemos reescribir

$$A_j + B_j = B_j + A_j \text{ para todo } j = 1, \cdots, m.$$

así se concluye la conmutatividad.

P.-3 $A + \mathbf{0} = \begin{bmatrix} A_1 + \mathbf{0}_1 \\ \vdots \\ A_m + \mathbf{0}_m \end{bmatrix}$ y para cada $j = 1, \cdots, m$ se cumple

$$A_j + \mathbf{0}_j = \left[a_j^L + 0, a_j^U + 0\right] = A_j$$

con lo cual se obtiene el resultado.

P.-4 $\alpha\,(\lambda A) = \alpha \left(\lambda \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}\right) = \alpha \left(\begin{bmatrix} \lambda\left[a_1^L, a_1^U\right] \\ \vdots \\ \lambda\left[a_m^L, a_m^U\right] \end{bmatrix}\right)$ pero como λ y α son no negativos se sigue que

$$\alpha\,(\lambda A) = \begin{bmatrix} \left[(\alpha\lambda)\,a_1^L, (\alpha\lambda)\,a_1^U\right] \\ \vdots \\ \left[(\alpha\lambda)\,a_m^L, (\alpha\lambda)\,a_m^U\right] \end{bmatrix} = (\alpha\lambda)A$$

P.-5 Inmediato de la definición.

P.-6 $\alpha\,(A + B) = \alpha \left(\begin{bmatrix} A_1 + B_1 \\ \vdots \\ A_m + B_m \end{bmatrix}\right) = \begin{bmatrix} \alpha\,(A_1 + B_1) \\ \vdots \\ \alpha\,(A_m + B_m) \end{bmatrix}$ pero

$$\begin{aligned} \alpha\,(A_j + B_j) &= \alpha\left[\left(a_j^L + b_j^L\right), \left(a_j^U + b_j^U\right)\right] \\ &= \left[\alpha\left(a_j^L + b_j^L\right), \alpha\left(a_j^U + b_j^U\right)\right] = \left[\alpha a_j^L + \alpha b_j^L, \alpha a_j^U + \alpha b_j^U\right] \end{aligned}$$

para todo $j = 1, \cdots, m$, por lo tanto, $\alpha\,(A + B) = \alpha A + \alpha B$

P.-7 Por definición, $(\alpha + \lambda)\,A = \begin{bmatrix} (\alpha + \lambda)\,A_1 \\ \vdots \\ (\alpha + \lambda)\,A_m \end{bmatrix}$ y por no negatividad de α y λ se sigue que

$$(\alpha + \lambda)\left[a_1^L, a_1^U\right] = \left[(\alpha + \lambda)\,a_1^L, (\alpha + \lambda)\,a_1^U\right]$$

□

Esta es la aritmética para el conjunto $\mathcal{I}^m$ que nos interesa, aúnque $\mathcal{I}^m$ no es espacio vectorial, pues ningún hiperrectángulo A posee inverso aditivo.

Aunque las propiedades $\boldsymbol{P.-4}$ y $\boldsymbol{P.-6}$ también se cumplen para escalares negativos, no es interesante desde el punto de vista teórico, por otro lado la propiedad $\boldsymbol{P.-7}$ no se cumple en general si tomamos $\lambda=-\alpha$, con α positivo, ya que el lado izquierdo es igual a $\mathbf{0}$ pero el lado derecho puede ser distinto de $\mathbf{0}$

$$0=(\alpha+(-\alpha))A\neq\alpha A+(-\alpha)\,A=\begin{bmatrix}\left[\alpha a_1^L,\alpha a_1^U\right]+\left[-\alpha a_1^U,-\alpha a_1^L\right]\\ \vdots\\ \left[\alpha a_m^L,\alpha a_m^U\right]+\left[-\alpha a_m^U,-\alpha a_m^L\right]\end{bmatrix}$$

Adicionalmente, aplicando el lema 2 de [5], que dice:

Lema 3.1.1. *Sean S y Z son conjuntos cerrados y convexos en un espacio vectorial normado real M y R un subconjunto acotado de M, si $S+R=Z+R$ entonces $S=Z$.*

se obtiene que $\mathcal{I}^m$ es un semigrupo conmutativo que satisface la ley cancelativa, ya que cada elemento A, B y $C\in\mathcal{I}^m$ son conjuntos cerrados, acotados y convexos de $\mathbb{R}^m$, con lo cual estamos frente a un conjunto con una estructura muy cercana a la de los espacios vectoriales.

Esta estructura es la necesaria para poder aplicar el teorema de embebimiento en un espacio vectorial, a un semigrupo conmutativo que satisface la ley cancelativa y producto por escalar no negativo. Este teorema se presenta en [5] y posteriormente lo enunciamos en este trabajo para ser aplicado al conjunto $\mathcal{I}^m$ y poder definir un concepto de diferenciación para estudiar el problema de optimización multi-intervalo valuado.

Ahora, si consideremos $A=\begin{bmatrix}A_1\\ \vdots\\ A_m\end{bmatrix}$ y $B=\begin{bmatrix}B_1\\ \vdots\\ B_m\end{bmatrix}\in\mathcal{I}^m$, entonces $A\ominus B=\begin{bmatrix}A_1\ominus B_1\\ \vdots\\ A_m\ominus B_m\end{bmatrix}$, donde $A_j\ominus B_j$ es la diferencia de Hukuhara aplicada en $\mathcal{I}$, de lo cual se tiene el siguiente resultado.

Proposición 3.1.2. *Sean $A=\begin{bmatrix}A_1\\ \vdots\\ A_m\end{bmatrix}$ y $B=\begin{bmatrix}B_1\\ \vdots\\ B_m\end{bmatrix}\in\mathcal{I}^m$, entonces $A\ominus B$ existe si y sólo si $l\left(A_j\right)\geq l\left(B_j\right)$, para todo $j=1,\cdots,m$.*

Demostración. Sean $A=\begin{bmatrix}A_1\\ \vdots\\ A_m\end{bmatrix}$ y $B=\begin{bmatrix}B_1\\ \vdots\\ B_m\end{bmatrix}\in\mathcal{I}^m$. Suponga que $A\ominus B$ existe, entonces existe $C\in\mathcal{I}^m$, tal que $A=B+C$, pero $C=\begin{bmatrix}C_1\\ \vdots\\ C_m\end{bmatrix}$ así que $A_j=B_j+C_j$, para $j=1,\cdots,m$ pero $A_j=\left[a_j^L,a_j^U\right]$, $B_j=\left[b_j^L,b_j^U\right]$ y $C_j=\left[c_j^L,c_j^U\right]$, para $j=1,\cdots,m$, de lo cual se sigue que:

$$a_j^L = b_j^L+c_j^L \tag{3.1}$$
$$a_j^U = b_j^U+c_j^U \tag{3.2}$$

para $j = 1, \cdots, m$, de (3.1) y (3.2) se sigue que

$$\begin{aligned} c_j^L &= a_j^L - b_j^L \quad (3.3) \\ c_j^U &= a_j^U - b_j^U \quad (3.4) \end{aligned}$$

para $j = 1, \cdots, m$, pero $C_j \in \mathcal{I}$, junto con (3.3) y (3.4) se sigue que

$$a_j^L - b_j^L \leq a_j^U - b_j^U \quad (3.5)$$

y al reescribir (3.5) se obtiene

$$a_j^U - a_j^L \geq b_j^U - b_j^L$$

para $j = 1, \cdots, m$, obteniendo el resultado.

Suponiendo $l(A_j) \geq l(B_j)$, para $j = 1, \cdots, m$. Basta seguir (3.1), (3.2), (3.3), (3.4) y (3.5) en el sentido contrario. □

Como consecuencia de lo anterior, si definimos $v(A)$ con $A \in \mathcal{I}^m$, como $v(A) = \prod_{j=1}^{m} l(A_j)$, entonces se tiene la condición necesaria para el conjunto $\mathcal{I}^m$ similar a la del resultado en 2.2.1 para el conjunto $\mathcal{I}$.

Proposición 3.1.3. *Sean A y $B \in \mathcal{I}^m$, entonces, si $A \ominus B$ existe entonces $v(A) \geq v(B)$.*

Demostración. Si $A \ominus B$ existe entonces por la proposición anterior, se tiene que $l(A_j) \geq l(B_j)$, para $j = 1, \cdots, m$ y por definición de v, se obtiene el resultado. □

En la proposición anterior, el reciproco no es cierto puesto que, si $m = 2$, tomamos $A = \begin{bmatrix} [0, 0.5] \\ [1, 5] \end{bmatrix}$ y $B = \begin{bmatrix} [0, 1] \\ [1, 2] \end{bmatrix}$, observe que $v(A) = 2 \geq 1 = v(B)$, pero no existe $C \in \mathcal{I}^2$, tal que $A = B + C$, puesto que no existe $C_1 \in \mathcal{I}$, tal que $[0, 0.5] = [0, 1] + C_1$.

Los ordenes parciales y ordenes parciales estrictos en $\mathcal{I}$, definidos antes, generan los siguientes ordenes parciales en $\mathcal{I}^m$.

Teorema 3.1.1. *Sean $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$ y $B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix} \in \mathcal{I}^m$. Las relaciones binarias definidas como:*

1.
$$A \preceq_\eta B \quad \text{si y sólo si} \quad A_j \preceq_\eta B_j \quad \text{para todo} \quad j = 1, \cdots, m \quad (3.6)$$

2.
$$A \prec_\eta B \quad \text{si y sólo si} \quad A \preceq_\eta B \quad \text{y} \quad A \neq B, \quad (3.7)$$

3.
$$A \ll_\eta B \quad \text{si y sólo si} \quad A_j \ll_\eta B_j \quad \text{para todo} \quad j = 1, \cdots, m \quad (3.8)$$

para $\eta = LU$ ó CW, definen en $\mathcal{I}^m$ una relación de orden parcial con 3.6 y ordenes parciales estrictos con 3.7 y 3.8

Demostración.

1. Como $\preceq_\eta$ son ordenes parciales en $\mathcal{I}$ para $\eta = LU$ ó CW, se sigue que $A_j \preceq_\eta A_j$ para todo $j = 1, \cdots, m$, así $\preceq_\eta$ es reflexiva en $\mathcal{I}^m$. También si $A_j \preceq_\eta B_j$ y $B_j \preceq_\eta A_j$ para todo $j = 1, \cdots, m$, por antisimetría, se sigue que $A_j = B_j$ para todo $j = 1, \cdots, m$ y $\preceq_\eta$ es antisimétrica. Por último, si $A_j \preceq_\eta B_j$ y $B_j \preceq_\eta C_j$ para todo $j = 1, \cdots, m$, por transitividad se tiene que $A_j \preceq_\eta C_j$ para todo $j = 1, \cdots, m$, de lo cual se concluye que $\preceq_\eta$ son ordenes parciales en $\mathcal{I}^m$ para $\eta = LU$ ó CW.

2. Por proposicion 2.3.2 y la definición 2.3.4 se sigue que es orden parcial estricto.

3. Como $\ll_\eta$ son ordenes parciales estrictos en $\mathcal{I}$ para $\eta = LU$ ó CW, se sigue que para todo $A_j \in \mathcal{I}$, se tiene que no se cumple $A_j \ll_\eta A_j$ en $\mathcal{I}$ para todo $j = 1, \cdots, m$, y así se cumple que $\ll_\eta$ es irreflexiva, además hereda la transitividad de $\ll_\eta$ en $\mathcal{I}$ para $\ll_\eta$ en $\mathcal{I}^m$ y de la definición 2.3.4 se obtiene que $\ll_\eta$ son ordenes parciales estrictos en $\mathcal{I}^m$ para $\eta = LU$ ó CW.

□

Los ordenes parciales definidos anteriormente, son compatibles con la suma y el producto por escalar positivo, esto es, si A, B y $C \in \mathcal{I}^m$ entonces $A + C \preceq_{LU} B + C$ es equivalente a $A_j + C_j \preceq_{LU} B_j + C_j$ para todo $j = 1, \cdots, m$, lo cual es equivalente a $a_j^L + c_j^L \leq b_j^L + c_j^L$ y $a_j^U + c_j^U \leq b_j^U + c_j^U$ para $j = 1, \cdots, m$, y como son reales y el orden en $\mathbb{R}$ es compatible con la suma se obtiene el resultado. Análogamente si $\lambda > 0$, entonces $\lambda A \preceq_{LU} \lambda B$, entonces $\lambda A_j \preceq_{LU} \lambda B_j$ para $j = 1, \cdots, m$, si y sólo si $\lambda a_j^L \leq \lambda b_j^L$ y $\lambda a_j^U \leq \lambda b_j^U$ para $j = 1, \cdots, m$ y como el orden en $\mathbb{R}$ también es compatible con la multiplicación por escalar positivo, se concluye que $A \preceq_{LU} B$.

A partir de estos ordenes y sus propiedades en $\mathcal{I}^m$, definimos el conjunto

$$K_{\preceq_{LU},\mathbf{0}} = \{A \in \mathcal{I}^m : \mathbf{0} \preceq_{LU} A\}$$

o equivalentemente,

$$K_{\preceq_{LU},\mathbf{0}} = \{A \in \mathcal{I}^m : 0 \preceq_{LU} A_j, \text{ para } j = 1, \cdots, m\}$$

el cual es no vacio pues que $\mathbf{0} \in K_{\preceq_{LU},\mathbf{0}}$; este conjunto satisface las siguientes propiedades:

- Si $A \in K_{\preceq_{LU},\mathbf{0}}$, entonces $\lambda A \in K_{\preceq_{LU},\mathbf{0}}$ para todo λ no negativo, ya que la relación de orden parcial es compatible con la multiplicación por escalar positivo.

- Si A, $B \in K_{\preceq_{LU},\mathbf{0}}$, entonces para todo $\lambda \in (0, 1)$ se obtiene que
$$\lambda A + (1-\lambda) B = \begin{bmatrix} \left[\lambda a_1^L + (1-\lambda) b_1^L, \lambda a_1^U + (1-\lambda) b_1^U\right] \\ \vdots \\ \left[\lambda a_m^L + (1-\lambda) b_m^L, \lambda a_m^U + (1-\lambda) b_m^U\right] \end{bmatrix}$$
y como para todo $j = 1, \cdots, m$ se satisface que $0 \leq \lambda a_j^L + (1-\lambda) b_j^L$, por la compatibilidad del orden en $\mathbb{R}$ con la suma, $\lambda A + (1-\lambda) B \in K_{\preceq_{LU},\mathbf{0}}$, para todo $\lambda \in (0, 1)$.

Observe que las definiciones del conjunto, que clásicamente se conoce como Cono, y la propiedad de convexidad de un conjunto, se cumplen en $K_{\preceq_{LU},\mathbf{0}}$ y son las mismas que se definen en espacios vectoriales.

Por otra parte, si $A \preceq_{LU} B$ y $B \ominus A$ existe de la definición 2.2.1 en la página 6, entonces se sigue que $a_j^L \leq b_j^L$, $a_j^U \leq b_j^U$ y $b_j^L - a_j^L \leq b_j^U - a_j^U$ para todo $j = 1, \cdots, m$, de lo cual se concluye $0 \leq b_j^L - a_j^L \leq b_j^U - a_j^U$ y por lo tanto $B \ominus A \in K_{\preceq_{LU},\mathbf{0}}$. Observe que $A \preceq_{LU} B$ no implica que la diferencia de Hukuhara exista, como ejemplo tomemos $m = 1$, $A = [1, 2]$ y $B = [1.9, 2.1]$, se tiene que $A \preceq_{LU} B$ pero $B \ominus A$ no existe pues, $l\,(B) = 0.2$ y $l\,(A) = 1$ por proposición 3.1.2 en la página 22.

3.2. Funciones Multi-Intervalo Valuadas

Iniciamos esta sección definiendo la métrica en el conjunto $\mathcal{I}^m$ y con ella poder construir las definiciones de límite, continuidad y dos clases de diferenciación de funciones multi-intervalo valudadas.

3.2.1. Límite y Continuidad de Funciones Multi-Intervalo Valuadas

Definición 3.2.1. *Sean A, $B \in \mathcal{I}^m$, definimos la función Hausdorff infinito por:*

$$d_{H_\infty}(A, B) = \text{máx}_{j \in \{1,\cdots,m\}} \{d_H(A_j, B_j)\} \tag{3.9}$$

donde $d_H(A, B) = \text{máx}\left\{\sup\limits_{a \in A} \inf\limits_{b \in B} \|a - b\|, \sup\limits_{b \in B} \inf\limits_{a \in A} \|a - b\|\right\}$ es la métrica de Hausdorff en $\mathbb{R}$, visto como espacio euclideo.

Proposición 3.2.1. *La función definida en 3.9 es una métrica en $\mathcal{I}^m$.*

Demostración. Sean A, B y $C \in \mathcal{I}^m$, entonces

I. $d_{H_\infty}(A, B) \geq 0$, ya que $d_H(A_j, B_j) \geq 0$ para todo $j = 1, \cdots, m$.

II. Si $A = B$, entonces $A_j = B_j$ para todo $j = 1, \cdots, m$ y como d_H es métrica, se tiene que $d_H(A_j, B_j) = 0$ para todo $j = 1, \cdots, m$ y así se obtiene $d_{H_\infty}(A, B) = 0$. Ahora, si $d_{H_\infty}(A, B) = 0$, entonces $d_H(A_j, B_j) = 0$ para todo $j = 1, \cdots, m$, pero como d_H es métrica, se concluye que $A_j = B_j$ para todo $j = 1, \cdots, m$. Por lo tanto, $d_{H_\infty}(A, B) = 0$ si y sólo si $A = B$

III. Dado que $d_H(A_j, B_j) = d_H(B_j, A_j)$ para todo $j = 1, \cdots, m$, se sigue que $d_{H_\infty}(A, B) = d_{H_\infty}(B, A)$.

IV. Para $A, B, C \in \mathcal{I}^m$, $d_{H_\infty}(A, C) \leq d_{H_\infty}(A, B) + d_{H_\infty}(B, C)$ puesto que

$$d_H(A_j, B_j) \leq d_H(A_j, C_j) + d_H(C_j, B_j) \tag{3.10}$$

para todo $j = 1, \cdots, m$ ya que d_H es una métrica, y tomando *max* a ambos lados de la desigualda 3.22 y por propiedad de *max* se prueba la afirmación.

□

Observe que $d_{H_\infty}(A,B) = \max_{j\in\{1,\cdots,m\}} \left\{ \left|a_j^L - b_j^L\right|, \left|a_j^U - b_j^U\right| \right\}$, para efectos prácticos.

Definición 3.2.2. *Sea $\{A^k\}$ una sucesión en $\mathcal{I}^m$ y $A \in \mathcal{I}^m$. Decimos que la sucesión de hiper-rectángulos $\{A^k\}_{k=0}^{\infty}$ converge a A, y escribimos*

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

si para todo $\epsilon > 0$, existe un $N(\epsilon) > 0$, $N(\epsilon) \in \mathbb{N}$ tal que, para todo $k > N(\epsilon)$, $k \in \mathbb{N}$, tenemos que $d_{H_\infty}(A^k, A) < \epsilon$.

En la anterior definición, para el caso $m = 1$ estudiado en [2], al revisar las componentes A_j^k se obtiene que $\lim_{k\to\infty} A_j^k = A_j$ para todo $j = 1,\cdots,m$, si y sólo si $\lim_{k\to\infty}\left(a_j^L\right)^k = a_j^L$ y $\lim_{k\to\infty}\left(a_j^U\right)^k = a_j^U$ en $\mathbb{R}$, donde $A_j = \left[a_j^L, a_j^U\right]$, para $j = 1,\cdots,m$.

Así, obtenemos un resultado análogo al del cálculo en varias variables para funciones multi-intervalos valuadas.

Proposición 3.2.2. *Sean $\left\{ A^k = \begin{bmatrix} A_1^k \\ \vdots \\ A_m^k \end{bmatrix} \right\}$ una sucesión en $\mathcal{I}^m$ y $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathcal{I}^m$. $\lim_{k\to\infty} A^k = A$ si y sólo si $\lim_{k\to\infty} A_j^k = A_j$, para $j = 1,\cdots,m$.*

Demostración. Sean $\left\{ A^k = \begin{bmatrix} A_1^k \\ \vdots \\ A_m^k \end{bmatrix} \right\}$ una sucesión en $\mathcal{I}^m$ y $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathcal{I}^m$. Supongamos que $\lim_{k\to\infty} A^k = A$, luego para $\epsilon > 0$, existe $N(\epsilon) \in \mathbb{N}$, tal que $d_{H_\infty}(A^k, A) < \epsilon$ para todo $k > N(\epsilon)$, por lo tanto para todo $j = 1,\cdots,m$, y para $k > N(\epsilon)$, se cumple que

$$d_H\left(A_j^k, A_j\right) < \epsilon \text{ ya que } d_{H_\infty}\left(A^k, A\right) \geq d_H\left(A_j^k, A_j\right)$$

y de lo cual se obtiene que $\lim_{k\to\infty} A_j^k = A_j$, para $j = 1,\cdots,m$.

Del otro lado, sean $\left\{A_j^k\right\}_{k=0}^{\infty}$, $j = 1,\cdots,m$, sucesiones en $\mathcal{I}$ y $A_j \in \mathcal{I}$, tales que $\lim_{k\to\infty} A_j^k = A_j$, para $j = 1,\cdots,m$, por lo tanto dado $\epsilon > 0$, existen $N_j(\epsilon) \in \mathbb{N}$, tal que para todo $k > N_j(\epsilon)$ y todo $j = 1,\cdots,m$,

$$d_H\left(A_j^k, A_j\right) < \epsilon$$

tomando $N(\epsilon) = \max\{N_1(\epsilon),\cdots,N_m(\epsilon)\}$, se sigue que para $k > N(\epsilon)$, que

$$\max_{j=1,\cdots,m}\left\{d_H\left(A_j^k, A_j\right)\right\} < \epsilon$$

luego

$$d_{H_\infty}\left(A^k, A\right) < \epsilon$$

y así se obtiene el resultado. □

Definición 3.2.3.

$$f\colon \mathbb{R}^n \to \mathcal{I}^m$$

es llamada una función multi-intervalo valuada, es decir, $f(x) = f(x_1, \cdots, x_n)$ *es un arreglo* $m \times 1$ *donde cada componente es un intervalo cerrado, acotado y convexo.*

La función multi-intervalo valuada f puede ser escrita como $f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$ donde cada $f_j\colon \mathbb{R}^n \to \mathcal{I}$, $j = 1, \cdots, m$ son funciones intervalo-valuadas (ver [2])

Definición 3.2.4. *Sean* f *una función multi-intervalo valuada sobre* $\mathbb{R}^n$*,* $c \in \mathbb{R}^n$ *y* $A \in \mathcal{I}^m$*. Definimos*

$$\lim_{x \to c} f(x) = A$$

si para todo $\epsilon > 0$*, existe un* $\delta(\epsilon)$ *tal que si* $\|x - c\| < \delta(\epsilon)$ *entonces* $d_{H_\infty}(f(x), A) < \epsilon$*.*

Con esta definición obtenemos los siguientes resultados.

Proposición 3.2.3. *Sean* f *una función multi-intervalo valuada sobre* $\mathbb{R}^n$*,* $c \in \mathbb{R}^n$ *y* $A \in \mathcal{I}^m$*.* $\lim_{x \to c} f(x) = A$ *si y sólo si* $\lim_{x \to c} f_j(x) = A_j$ *para todo* $j = 1, \cdots, m$*.*

Demostración. Sean f una función multi-intervalo valuada sobre $\mathbb{R}^n$, $c \in \mathbb{R}^n$, $A \in \mathcal{I}^m$ y supongamos que $\lim_{x \to c} f(x) = A$, así para $\epsilon > 0$, tenemos que existe $\delta(\epsilon)$, tal que para $\|x - c\| < \delta(\epsilon)$ implica que $d_{H_\infty}(f(x), A) < \epsilon$, por lo tanto, por la definición de la métrica Hausdorff infinito, para todo $j = 1, \cdots, m$, y para $\|x - c\| < \delta(\epsilon)$, se cumple que

$$d_H(f_j(x), A_j) \leq d_{H_\infty}(f(x), A) < \epsilon$$

de lo cual se obtiene que $\lim_{x \to c} f_j(x) = A_j$, para $j = 1, \cdots, m$.

En el otro sentido, sean f_j funciones intervalo-valuadas para $j = 1, \cdots, m$ las componentes de f, y sean $A_j \in \mathcal{I}$ tal que $\lim_{x \to c} f_j(x) = A_j$, para $j = 1, \cdots, m$, por lo tanto, dado $\epsilon > 0$, existen $\delta_j(\epsilon) \in \mathbb{N}$ tal que para $\|x - c\| < \delta_j(\epsilon)$ y

$$d_H(f_j(x), A_j) < \epsilon$$

para $j = 1, \cdots, m$. Tomando $\delta(\epsilon) = \min\{\delta_1(\epsilon), \cdots, \delta_m(\epsilon)\}$, se sigue que para $\|x - c\| < \delta(\epsilon)$, se cumple

$$d_{H_\infty}(f(x), A) < \epsilon$$

donde $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$ y así se sigue el resultado. □

Nuevamente, en analogía con el cálculo, si $A = f(c)$, entonces decimos que la función multi-intervalo valuada f es continua.

Similarmente a los resultados del cálculo vectorial tenemos:

Proposición 3.2.4. f *es una función continua en* $\mathcal{I}^m$ *si y sólo si las funciones componentes* f_j *de* f *son funciones continuas en* $\mathcal{I}$*.*

Demostración. Sean f una función multi-intervalo valuada continua y $x_0 \in \mathbb{R}^n$, entonces $\lim\limits_{x \to x_0} f(x) = f(x_0)$, por proposición 3.2.3 tomando $A = f(x_0)$ es equivalente a que $\lim\limits_{x \to x_0} f_j(x) = f_j(x_0)$ para todo $j = 1, \cdots, m$, luego las funciones componentes f_j de f son funciones continuas en x_0. □

En el caso $n = 1$ y $m = 1$, podemos definir los límites laterales derechos e izquierdo en forma natural, denotados por

$$\lim_{x \to c^+} f(x) \text{ y } \lim_{x \to c^-} f(x)$$

y en este caso la función f es continua si y sólo si $\lim\limits_{x \to c^+} f(x) = f(c) = \lim\limits_{x \to c^-} f(x)$, (ver [2]).

Proposición 3.2.5. *Sea $f\colon \mathbb{R}^n \to \mathcal{I}^m$ una función multi-intervalo valuada, $f = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}$, con $f_j\colon \mathbb{R}^n \to \mathcal{I}$, para $j = 1, \cdots, m$, entonces f es continua en c en el sentido multi-intervalo valuada si y sólo si f_j^L y f_j^U son continuas (en los reales) en c, donde $f_j = \left[f_j^L, f_j^U\right]$ y $f_j^L(x) \leq f_j^U(x)$ para $j = 1, \cdots, m$.*

Demostración. Al usar la proposición 3.2.3 y proposición 3.3 en [2], se obtiene el resultado. □

3.2.2. Diferenciación de Funciones Multi-Intervalo Valuadas

En esta sección estudiaremos dos clases de diferenciación de las funciones multi-intervalo valuadas. Iniciamos definiendo estos conceptos para funciones $f\colon \mathbb{R}^n \to \mathcal{I}^m$, con $n \geq 2$ y $m \geq 2$, el caso $m = 1$, (se estudia en [2]). En esta sección hacemos notar que la función f también puede ser descrita como $f = \left[f^L, f^U\right]$, $f^L, f^U\colon \mathbb{R}^n \to \mathbb{R}^m$, donde $f^L(x) = \begin{bmatrix} f_1^L(x) \\ \vdots \\ f_j^L(x) \\ \vdots \\ f_m^L(x) \end{bmatrix}$ y

$$f^U(x) = \begin{bmatrix} f_1^U(x) \\ \vdots \\ f_j^U(x) \\ \vdots \\ f_m^U(x) \end{bmatrix} \text{ y } f_j^L(x) \leq f_j^U(x), \text{ para todo } j = 1, \cdots, m \text{ y todo } x \in \mathbb{R}^n.$$

Antes de seguir con el estudio de diferenciación, veamos las propiedades de la métrica d_H que permiten usar el teorema 2.6.2. La prueba de estas propiedades se realizan en [3]:

Teorema (Propiedades de la métrica d_H)**.** *Sean $A = \left[a^L, a^U\right]$, $B = \left[b^L, b^U\right]$, $C = \left[c^L, c^U\right]$ y $D = \left[d^L, d^U\right] \in \mathcal{I}$.*

1. $d_H(A + B, A + C) = d_H(B, C)$
2. $d_H(A + B, C + D) = d_H(A, C) + d_H(B, D)$
3. $d_H(\alpha B, \alpha C) = |\alpha|\, d_H(B, C)$, $\alpha \in \mathbb{R}$

Dado que la métrica d_{H_∞} se define a partir de la metrica de Hausdorff, d_{H_∞} tiene las siguientes propiedades:

Teorema 3.2.1. *La métrica d_{H_∞} satisface que para A, B y $C \in \mathcal{I}^m$ y $\alpha \geq 0$:*

1. $d_{H_\infty}(A+C, B+C) = d_{H_\infty}(A, B)$
2. $d_{H_\infty}(\alpha A, \alpha B) = \alpha d_{H_\infty}(A, B)$
3. *La suma y el producto por escalar son funciones continuas con respecto a la métrica Hausdorff infinito.*

Demostración.

1. $d_{H_\infty}(A+C, B+C) = \max\limits_{1\leq j\leq m}\{d_H(A_j+C_j, B_j+C_j)\} = \max\limits_{1\leq j\leq m}\{d_H(A_j, B_j)\}$ por el numeral 1 del teorema anterior.

2. Se tiene que, $d_{H_\infty}(\alpha A, \alpha B) = \max\limits_{1\leq j\leq m}\{d_H(\alpha A_j, \alpha B_j)\}$ y por propiedades del máximo se sigue el resultado.

3. Para la continuidad, sea $\left\{A^k\right\}$ y $\left\{B^k\right\}$ sucesiones convergentes en $\mathcal{I}^m$, tal que $A^k \to A$ y $B^k \to B$ cuando $k \to \infty$, es decir, dado $\epsilon > 0$, $\exists N_1(\epsilon)$ tal que $\forall k \geq N_1(\epsilon)$, $d_{H_\infty}\left(A^k, A\right) < \epsilon/2$. También $\exists N_2(\epsilon)$ tal que $\forall k \geq N_2(\epsilon)$, $d_{H_\infty}\left(B^k, B\right) < \epsilon/2$.

   Por otra parte,
   $$d_{H_\infty}\left(A^k+B^k, A+B\right) = \max_{1\leq j\leq m}\left\{d_H\left(A_j^k+B_j^k, A_j+B_j\right)\right\} \tag{3.11}$$
   y por la propiedad 2 de la métrica d_H se sigue que
   $$d_H\left(A_j^k+B_j^k, A_j+B_j\right) \leq d_H\left(A_j^k, A_j\right) + d_H\left(B_j^k, B_j\right) \tag{3.12}$$
   y por propiedades del máximo y reemplazando (3.12) en (3.11) se obtiene
   $$d_{H_\infty}\left(A^k+B^k, A+B\right) \leq d_{H_\infty}\left(A^k, A\right) + d_{H_\infty}\left(B^k, B\right) < \epsilon$$
   para todo $k > \max\{N_1, N_2\}$, con lo cual se obtiene la convergencia de la suma, así la suma es función continua con respecto a d_{H_∞}.
   Por otra parte, sean $\alpha_k \to \alpha$ en $\mathbb{R}$, entonces
   $$d_{H_\infty}\left(\alpha_k A^k, \alpha A\right) \leq d_{H_\infty}\left(\alpha_k A^k, \alpha_k A\right) + d_{H_\infty}(\alpha_k A, \alpha A)$$
   por desigualdad triángular, y dado que $d_{H_\infty}(\alpha A, \alpha B) = |\alpha|\, d_{H_\infty}(A, B)$ y $d_{H_\infty}(\alpha_k A, \alpha A) = \max\limits_{j\in\{1,\cdots,m\}}\left\{\left|\alpha_k a_j^L - \alpha a_j^L\right|, \left|\alpha_k a_j^U - \alpha a_j^U\right|\right\} = |\alpha_k - \alpha| \max\limits_{j\in\{1,\cdots,m\}}\left\{\left|a_j^L\right|, \left|a_j^U\right|\right\} = |\alpha_k - \alpha|$ $d_{H_\infty}(A, \mathbf{0})$, se sigue
   $$d_{H_\infty}(\alpha_k A_k, \alpha A) \leq |\alpha_k|\, d_{H_\infty}\left(A^k, A\right) + |\alpha_k - \alpha|\, d_{H_\infty}(A, \mathbf{0})$$
   y dado que $\{\alpha_k\}$ es convergente entonces es acotado y se sigue la convergencia de $\alpha_k A_k \to \alpha A$ y por tanto la continuidad de la función producto por escalar, con respecto a d_{H_∞} es probada.

$\square$

Con los resultados anteriores, tenemos que el conjunto $\mathcal{I}^m$ satisface todas las condiciones del teorema de embebimiento 2.6.1. El teorema 3.2.1 asegura que la métrica que se define en él es invariante bajo traslaciones, homogenea con el producto por escalar no negativo y las operaciones suma y producto por escalar son funciones continuas. Ahora, la manera en que se construye el espacio vectorial normado en el cual se embebe el conjunto $\mathcal{I}^m$, se logra definiendo primero una relación de equivalencia "$\sim$ " en $\mathcal{I}^m \times \mathcal{I}^m$ de la siguiente forma: para $(A,B),(C,D) \in \mathcal{I}^m \times \mathcal{I}^m$ $(A,B) \sim (C,D)$ si y sólo si $A+D = B+C$. Dada la relación de equivalencia en $\mathcal{I}^m \times \mathcal{I}^m$, el conjunto cociente $(\mathcal{I}^m \times \mathcal{I}^m / \sim)$ resulta ser un espacio vectorial (ver[5]), de esta forma existe una función $\pi : \mathcal{I}^m \to (\mathcal{I}^m \times \mathcal{I}^m / \sim)$ tal que para cada $A \in \mathcal{I}^m$, $\pi(A)$ es la clase asociada por π en $(\mathcal{I}^m \times \mathcal{I}^m / \sim)$ y esta dada por $\widetilde{A} = \langle A, 0 \rangle = \{(C,D) \in \mathcal{I}^m \times \mathcal{I}^m : A + D = C\}$, como este conjunto cociente es un espacio vectorial, el inverso adictivo de la clase $\langle A, B\rangle$ es $\langle B, A\rangle$ y el producto por escalar está definido por

$$\lambda \langle A, B\rangle = \begin{cases} \langle \lambda A, \lambda B\rangle & \text{si } \lambda \geq 0 \\ \langle |\lambda| B, |\lambda| A\rangle & \text{si } \lambda < 0 \end{cases}$$

donde se tiene que el inverso adictivo $\langle B, A\rangle = -\langle A, B\rangle$ y la métrica inducida por d_{H_∞} esta definida por $\eta_{H_\infty}(\langle A, B\rangle, \langle C, D\rangle) = d_{H_\infty}(A+D, B+C)$ la cual induce la norma en $(\mathcal{I}^m \times \mathcal{I}^m / \sim)$ definida por $\|\langle A, B\rangle - \langle C, D\rangle\|_{(\mathcal{I}^m \times \mathcal{I}^m/\sim)_\infty} = \eta_{H_\infty}(\langle A, B\rangle, \langle C, D\rangle)$.

La relevancia de la extensión del conjunto $\mathcal{I}^m$ a un espacio vectorial normado en nuestro trabajo, es que ya podemos implementar la definición de derivada Frechet sobre el espacio normado $(\mathcal{I}^m \times \mathcal{I}^m / \sim)$, además de que su restricción al cono convexo isomorfo a $\mathcal{I}^m$ define un concepto de diferenciación en él, al usar el embebimiento π para obtener de una función $f\colon \mathbb{R}^n \to \mathcal{I}^m$, la función $\widetilde{f}\colon \mathbb{R}^n \to (\mathcal{I}^m \times \mathcal{I}^m / \sim)$ donde $\widetilde{f} = \pi \circ f$ y su relación por el embebimiento. En adelante denotaremos por N al espacio vectorial normado $(\mathcal{I}^m \times \mathcal{I}^m / \sim)$. Así se genera la siguiente definición.

Definición 3.2.5. *Sea $X \subseteq \mathbb{R}^n$ abierto. Una función multi-intervalo valuada $f\colon X \to \mathcal{I}^m$, se llama debílmente diferenciable en $x_0 \in X$ si la función $\widetilde{f}\colon X \to N$ es diferenciable total (ó Frechet) en x_0 ([7], pag 172.) donde $\widetilde{f}(x) = \langle f(x), 0\rangle$, esto es, para cada $h \in \mathbb{R}^n$, existe $\delta\widetilde{f}(x_0; h) \in N$ el cual es lineal y continuo con respecto a h tal que*

$$\lim_{\|h\|\to 0} \frac{\left\|\widetilde{f}(x_0+h) - \widetilde{f}(x_0) - \delta\widetilde{f}(x_0;h)\right\|_{N_\infty}}{\|h\|} = 0. \tag{3.13}$$

La anterior definición es equivalente a

$$\widetilde{f}(x_0+h) - \widetilde{f}(x_0) - \delta\widetilde{f}(x;h) = o(\|h\|)$$

y como para todo x_0 y h, tenemos que $\delta\widetilde{f}(x_0;h) \in N$ podemos escribir $\delta\widetilde{f}(x_0;h) = \left\langle A^h(x_0), B^h(x_0)\right\rangle$, $h \in \mathbb{R}^n$, y $A^h(x_0)$, $B^h(x_0) \in \mathcal{I}^m$, por abuso de notación simplemente A^h y B^h, entonces en términos de la métrica de Hausdorff infinito (3.13) significa:

$$d_{H_\infty}\left(f(x_0+h) + B^h, f(x_0) + A^h\right) = o(\|h\|)$$

o su equivalente

$$\lim_{\|h\|\to 0} \frac{d_{H_\infty}\left(f(x_0+h) + B^h, f(x_0) + A^h\right)}{\|h\|} = 0.$$

Al tomar una base $\xi^1, \cdots, \xi^n$ de $\mathbb{R}^n$ entonces $h = \sum_{i=1}^{n} \overline{h_i}\xi^i$, $\overline{h_i} \in \mathbb{R}$ para todo $i = 1, \cdots, n$. Si $\delta\widetilde{f}\left(x_0; \xi^i\right) = \left\langle A^{\xi^i}, \mathbf{0}\right\rangle$, $i = 1, \cdots, n$, entonces decimos que f **es conicamente diferenciable en** $x_0 \in X$ y tenemos que

$$\delta\widetilde{f}\left(x_0; h\right) = \sum_{i=1}^{n} \overline{h_i} \left\langle A^{\xi^i}, \mathbf{0}\right\rangle$$

debido a la linealidad con respecto a h. Tomando la base canónica de $\mathbb{R}^n$, $e_1, \cdots, e_n$, entonces si f es debílmente diferenciable en $x_0 \in X$, entonces existen $A^{e_i}\left(x_0\right)$, $B^{e_i}\left(x_0\right) \in \mathcal{I}^m$, $i = 1, \cdots, n$ tal que

$$\delta\widetilde{f}\left(x_0; h\right) = \sum_{i=1}^{n} h_i \left\langle A^{e_i}, B^{e_i}\right\rangle \tag{3.14}$$

para todo $h = \begin{pmatrix} h_1, & h_2, & \cdots & h_n \end{pmatrix}^{\top} \in \mathbb{R}^n$. Ahora podemos asumir además que

$$A^{e_i} = \begin{bmatrix} A_1^{e_i} \\ \vdots \\ A_m^{e_i} \end{bmatrix} \text{ y } B^{e_i} = \begin{bmatrix} B_1^{e_i} \\ \vdots \\ B_m^{e_i} \end{bmatrix} \tag{3.15}$$

donde $A_j^{e_i} = \left[\left(a_j^L\right)^{e_i}, \left(a_j^U\right)^{e_i}\right]$ y $B_j^{e_i} = \left[\left(b_j^L\right)^{e_i}, \left(b_j^U\right)^{e_i}\right]$ para todo $j = 1, \cdots, m$ y $i = 1, \cdots, n$ y que

$$f\left(x\right) = \begin{bmatrix} f_1\left(x\right) \\ \vdots \\ f_m\left(x\right) \end{bmatrix}$$

donde $f_j\left(x\right) = \left[f_j^L\left(x\right), f_j^U\left(x\right)\right]$ para todo $j = 1, \cdots, m$. Por un cálculo directo obtenemos la relación

$$\begin{aligned}
& \left\|\widetilde{f}\left(x_0 + h\right) - \widetilde{f}\left(x_0\right) - \delta\widetilde{f}\left(x_0; h\right)\right\|_{N_\infty} \\
= & \left\|\left\langle f\left(x_0 + h\right), \mathbf{0}\right\rangle - \left\langle f\left(x_0\right), \mathbf{0}\right\rangle - \sum_{i=1}^{n} h_i \left\langle A^{e^i}, B^{e^i}\right\rangle\right\|_{N_\infty} \\
= & \left\|\left\langle f\left(x_0 + h\right), \mathbf{0}\right\rangle + \left\langle \mathbf{0}, f\left(x_0\right)\right\rangle + \sum_{i=1}^{n} h_i \left\langle B^{e^i}, A^{e^i}\right\rangle\right\|_{N_\infty} \\
= & \left\|\left\langle f\left(x_0 + h\right) + \sum_{h_i \geq 0} h_i B^{e_i} - \sum_{h_i < 0} h_i A^{e_i},\ f\left(x_0\right) - \sum_{h_i < 0} h_i B^{e_i} + \sum_{h_i \geq 0} h_i A^{e_i}\right\rangle\right\|_{N_\infty} \\
= & d_{H_\infty}\left(W, V\right)
\end{aligned} \tag{3.16}$$

donde

$$W = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix} \text{ y } V = \begin{bmatrix} V_1 \\ \vdots \\ V_m \end{bmatrix}$$

y los extremos de cada W_j y V_j, para $j = 1, \cdots, m$, son:

$$\begin{aligned}
w_j^L &= f_j^L(x_0 + h) + \sum_{h_i \geq 0} h_i \left(b_j^L\right)^{e^i} - \sum_{h_i < 0} h_i \left(a_j^L\right)^{e^i} \\
w_j^U &= f_j^U(x_0 + h) + \sum_{h_i \geq 0} h_i \left(b_j^U\right)^{e^i} - \sum_{h_i < 0} h_i \left(a_j^U\right)^{e^i} \\
v_j^L &= f_j^L(x_0) + \sum_{h_i \geq 0} h_i \left(a_j^L\right)^{e^i} - \sum_{h_i < 0} h_i \left(b_j^L\right)^{e^i} \\
v_j^U &= f_j^U(x_0) + \sum_{h_i \geq 0} h_i \left(a_j^U\right)^{e^i} - \sum_{h_i < 0} h_i \left(b_j^U\right)^{e^i}
\end{aligned}$$

Ahora observemos que

$$\left|w_j^L - u_j^L\right| = \left|f_j^L(x_0 + h) - f_j^L(x_0) - \sum_{i=1}^{n} h_i \left(\left(a_j^L\right)^{e^i} - \left(b_j^L\right)^{e^i}\right)\right| \tag{3.17}$$

$$\left|w_j^U - u_j^U\right| = \left|f_j^U(x_0 + h) - f_j^U(x_0) - \sum_{i=1}^{n} h_i \left(\left(a_j^U\right)^{e^i} - \left(b_j^U\right)^{e^i}\right)\right| \tag{3.18}$$

para $j = 1, \cdots, m$, y como f es debílmente diferenciable, entonces las expresiones (3.17) y (3.18) son $o(\|h\|)$ y así el siguiente teorema es probado.

Teorema 3.2.2. *Sea X un subconjunto abierto de $\mathbb{R}^n$. Si la función multi-intervalo valuada $f\colon X \to \mathcal{I}^m$, donde $f = \left[f^L, f^U\right]$, f^L, $f^U\colon \mathbb{R}^n \to \mathbb{R}^m$, es debílmente diferenciable en $x_0 \in X$ y si $\delta\widetilde{f}(x_0; h)$ satisface (3.14) y (3.15), entonces las trasformaciones f_j^L y $f_j^U\colon \mathbb{R}^n \to \mathbb{R}$ son diferenciables en x_0, $j = 1, \cdots, m$ y sus derivadas están dadas por*

$$\left(f_j^L\right)'_{x_0}(h) = \sum_{i=1}^{n} h_i \left(\left(a_j^L\right)^{e^i} - \left(b_j^L\right)^{e^i}\right)$$

$$\left(f_j^U\right)'_{x_0}(h) = \sum_{i=1}^{n} h_i \left(\left(a_j^U\right)^{e^i} - \left(b_j^U\right)^{e^i}\right)$$

para $j = 1, \cdots, n$ y $h = \begin{pmatrix} h_1, & h_2, & \cdots & h_n \end{pmatrix}^\top$.

Afortunadamente, el reciproco del teorema anterior, también se puede demostrar.

Teorema 3.2.3. *Sea X un subconjunto abierto de $\mathbb{R}^n$, y las funciones f_j^L, $f_j^U\colon X \to \mathbb{R}$ satisfacen $f_j^L(x) \leq f_j^U(x)$, $j = 1, \cdots, m$. Si las funciones f_j^L y f_j^U, $j = 1, \cdots, m$ son diferenciables en $x_0 \in X$, entonces la función multi-intervalo valuada $f(x) = \begin{bmatrix} \left[f_1^L(x), f_1^U(x)\right] \\ \vdots \\ \left[f_m^L(x), f_m^U(x)\right] \end{bmatrix}$ es debílmente diferenciable en $x_0 \in X$ y*

$$\delta\widetilde{f}(x_0; h) = \sum_{i=1}^{n} h_i \left\langle \prod_{j=1}^{m} \left[\alpha_i^j(x_0) + \frac{\partial f_j^L}{\partial x_i}(x_0), \beta_i^j(x_0) + \frac{\partial f_j^U}{\partial x_i}(x_0)\right], \prod_{j=1}^{m} \left[\alpha_i^j(x_0), \beta_i^j(x_0)\right] \right\rangle \tag{3.19}$$

donde los $\alpha_i^j(x_0), \beta_i^j(x_0)$ *son números arbitrarios que satisfacen:*

$$\begin{gathered}\alpha_i^j(x_0) \leq \beta_i^j(x_0) \\ \alpha_i^j(x_0) + \frac{\partial f_j^L}{\partial x_i}(x_0) \leq \beta_i^j(x_0) + \frac{\partial f_j^U}{\partial x_i}(x_0)\end{gathered} \tag{3.20}$$

para $i = 1, \cdots, n$ *y* $j = 1, \cdots, m$.

Demostración. Sean $\alpha_i^j(x_0)$, $\beta_i^j(x_0)$ números que satisfacen (3.20) y $F(x_0; h)$ denotando el lado derecho de (3.19). Entonces $\left\|\widetilde{f}(x_0+h) - \widetilde{f}(x_0) - F(x_0; h)\right\|_{N_\infty}$ es la cantidad exacta sobre el lado izquierdo de (3.16) y si en (3.15) tomamos

$$\begin{array}{cc}\left(a_j^L\right)^{e^i} = \alpha_i^j(x_0) + \frac{\partial f_j^L}{\partial x_i}(x_0) & \left(a_j^U\right)^{e^i} = \beta_i^j(x_0) + \frac{\partial f_j^U}{\partial x_i}(x_0) \\ \left(b_j^L\right)^{e^i} = \alpha_i^j(x_0) & \left(b_j^U\right)^{e^i} = \beta_i^j(x_0)\end{array}$$

para $i = 1, \cdots, n$ y $j = 1, \cdots, m$ y de la hipótesis que las funciones f_j^L, f_j^U son diferenciables en x_0 se obtiene

$$\begin{aligned} & \left\|\widetilde{f}(x_0+h) - \widetilde{f}(x_0) - F(x_0; h)\right\|_{N_\infty} \\ \leq\ & \sum_{j=1}^{m}\left|f_j^L(x_0+h) - f_j^L(x_0) - \sum_{i=1}^{n} h_i\left(\frac{\partial f_j^L}{\partial x_i}(x_0)\right)\right| \\ & +\left|f_j^U(x_0+h) - f_j^U(x_0) - \sum_{i=1}^{n} h_i\left(\frac{\partial f_j^U}{\partial x_i}(x_0)\right)\right| \\ =\ & o(\|h\|)\end{aligned}$$

□

Observemos que en el teorema anterior podriamos elegir $\alpha_i^j(x_0) = \beta_i^j(x_0) = 0$, si

$$\frac{\partial f_j^L}{\partial x_i}(x_0) \leq \frac{\partial f_j^U}{\partial x_i}(x_0)$$

y si para algún i y j,

$$\frac{\partial f_j^L}{\partial x_i}(x_0) > \frac{\partial f_j^U}{\partial x_i}(x_0)$$

podriamos tomar

$$\alpha_i^j(x_0) = -\frac{\partial f_j^U}{\partial x_i}(x_0)$$

y

$$\beta_i^j(x_0) = -\frac{\partial f_j^L}{\partial x_i}(x_0)$$

para $i = 1, \cdots, n$ y $j = 1, \cdots, m$. Además si para todo $i = 1, \cdots, n$ y $j = 1, \cdots, m$ se cumple que $\frac{\partial f_j^L}{\partial x_i}(x_0) \leq \frac{\partial f_j^U}{\partial x_i}(x_0)$, entonces tenemos que

$$\delta\widetilde{f}(x_0; h) = \sum_{i=1}^{n} h_i\left\langle \prod_{j=1}^{m}\left[\frac{\partial f_j^L}{\partial x_i}(x_0), \frac{\partial f_j^U}{\partial x_i}(x_0)\right], \prod_{j=1}^{m}[0,0]\right\rangle = \sum_{i=1}^{n} h_i \prod_{j=1}^{m}\left[\frac{\partial f_j^L}{\partial x_i}\widetilde{(x_0), \frac{\partial f_j^U}{\partial x_i}}(x_0)\right]$$

y así f sería conicamente diferenciable.

Corolario 3.2.1. *Sea $X \subseteq \mathbb{R}^n$ abierto. Una función multi-intervalo valuada $f\colon X \to \mathcal{I}^m$, con $f = \left[f^L, f^U\right]$, f^L, $f^U\colon X \to \mathbb{R}^m$, es debílmente diferenciable en x_0 si y sólo si f^L y f^U, son diferenciables (en el sentido Frechet) en x_0 ([15], definición 12.2).*

Demostración. Resultado inmediato de los dos teoremas precedentes. □

En el numeral 3.2.1, obtuvimos que $\mathcal{I}^m$ es un espacio métrico, y por medio de su métrica podemos generar una función no negativa $|\cdot|_{H_\infty}\colon \mathcal{I}^m \to \mathbb{R}$ tal que para $A \in \mathcal{I}^m$

$$\begin{aligned} |A|_{H_\infty} &= d_{H_\infty}(A, \mathbf{0}) \\ &= \max_{j\in\{1,\cdots,m\}} \left\{\left|a_j^L\right|, \left|a_j^U\right|\right\} \end{aligned} \tag{3.21}$$

donde $\mathbf{0} = \begin{bmatrix}0 & \cdots & 0\end{bmatrix}^\top \in \mathcal{I}^m$, a esta función se le denominará función valor absoluto, además es continua y satisface el siguiente teorema.

Teorema 3.2.4. *Sea $|\cdot|_{H_\infty}\colon \mathcal{I}^m \to \mathbb{R}$, definida por (3.21), entonces $|\cdot|_{H_\infty}$ es función no negativa tal que:*

1. $|\cdot|_{H_\infty}$ *es continua.*
2. $|A|_{H_\infty} = \mathbf{0}$ *si y sólo si* $A = \mathbf{0}$
3. $|\lambda A|_{H_\infty} = |\lambda|\,|A|_{H_\infty}$ *para todo* $\lambda \in \mathbb{R}$.
4. $|A+B|_{H_\infty} \leq |A|_{H_\infty} + |B|_{H_\infty}$

Demostración. $|\cdot|_{H_\infty}$ es función no negativa se sigue de la definición de d_{H_∞} por ser métrica.

1. Sea $\left\{A^k\right\}$ una sucesión convergente en $\mathcal{I}^m$ y $A \in \mathcal{I}^m$ tal que $\lim\limits_{k\to\infty} A^k = A$. Sea $\epsilon > 0$ dado, entonces como d_{H_∞} satisface la desigualdad triángular, se sigue que

$$d_{H_\infty}\left(A^k, \mathbf{0}\right) \leq d_{H_\infty}\left(A^k, A\right) + d_{H_\infty}(A, \mathbf{0})$$

o sea,

$$d_{H_\infty}\left(A^k, \mathbf{0}\right) - d_{H_\infty}(A, \mathbf{0}) \leq d_{H_\infty}\left(A^k, A\right) \tag{3.22}$$

análogamente se obtiene

$$d_{H_\infty}(A, \mathbf{0}) - d_{H_\infty}\left(A^k, \mathbf{0}\right) \leq d_{H_\infty}\left(A^k, A\right) \tag{3.23}$$

así, de (3.22) y (3.23)

$$\left|d_{H_\infty}(A, \mathbf{0}) - d_{H_\infty}\left(A^k, \mathbf{0}\right)\right| \leq d_{H_\infty}\left(A^k, A\right) \tag{3.24}$$

y como $\left\{A^k\right\}$ es convergente en $\mathcal{I}^m$, existe (ϵ) tal que para todo $k \geq N(\epsilon)$

$$d_{H_\infty}\left(A^k, A\right) < \epsilon \tag{3.25}$$

al aplicar (3.25) en (3.24) se obtiene 1.

2. Se sigue de la definición de métrica.

3. Se tiene que

$$\begin{aligned} |\lambda A|_{H_\infty} &= d_{H_\infty}(\lambda A, \mathbf{0}) \\ &= \max_{j\in\{1,\cdots,m\}} \left\{ \left|\lambda a_j^L\right|, \left|\lambda a_j^U\right| \right\} \\ &= |\lambda| \max_{j\in\{1,\cdots,m\}} \left\{ \left|a_j^L\right|, \left|a_j^U\right| \right\} \end{aligned}$$

   por propiedades del máximo en los números reales.

4. Nuevamente, como d_{H_∞} satisface la desigualdad triángular, se tiene que

$$d_{H_\infty}(A+B, \mathbf{0}) \leq d_{H_\infty}(A+B, B) + d_{H_\infty}(B, \mathbf{0})$$

   pero

$$d_{H_\infty}(A+B, B) = d_{H_\infty}(A, \mathbf{0})$$

   por ser una métrica invariante bajo traslaciones.

□

La función anterior, tiene una aplicación similar a la de la norma en los espacios vectoriales normados, salvo que esta no puede definir una métrica sobre $\mathcal{I}^m$, puesto que este no es un espacio vectorial y por lo tanto la igual usual $d(x,y) = \|x-y\|$ no es cierta, porque implica el uso del inverso aditivo de y el cual no existe en $\mathcal{I}^m$. Por último, aquí se ha probado que la métrica Haussdorf infinito es compatible con el producto por escalar no negativo, y con respecto a la suma es subaditiva.

Los siguientes resultados basados en la diferencia de Hukuhara, $\ominus$, nos permiten hacer cálculos y demostraciones sobre la diferenciabilidad de Hukuhara, que se define mas adelante.

Proposición 3.2.6. *Sean A, B y $C \in \mathcal{I}^m$, entonces $(A \ominus B) \ominus C = A \ominus (B + C)$, siempre que exista la diferencia de Hukuhara.*

Demostración. Sean A, B y $C \in \mathcal{I}^m$, entonces

donde

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix} \text{ y } C = \begin{bmatrix} C_1 \\ \vdots \\ C_m \end{bmatrix}$$

además suponemos que $(A \ominus B) \ominus C$ existe, entonces existe $D \in \mathcal{I}^m$ tal que $A \ominus B = D$, así $A_j \ominus B_j = D_j$ para $j = 1, \cdots, m$, esto es, $d_j^L = a_j^L - b_j^L$ y $d_j^U = a_j^U - b_j^U$ para $j = 1, \cdots, m$, además existe $E \in \mathcal{I}^m$ tal que $(A \ominus B) \ominus C = D \ominus C = E$, análogamente se sigue que

$$e_j^L = d_j^L - c_j^L = a_j^L - (b_j^L + c_j^L)$$

y

$$e_j^U = d_J^U - c_j^U = a_j^U - (b_j^U + c_j^U)$$

para $j = 1, \cdots, m$, luego $E_j = A_j \ominus (B_j + C_j)$, para $j = 1, \cdots, m$, de lo cual se sigue el resultado. □

Para definir una otra forma de diferenciación de funciones multi-intervalo valuadas, debemos fijar primero una dirección, y así poder usar la diferencia de Hukuhara en el mismo sentido que se utiliza para funciones mono-objetivo intervalo valuadas en [2].

Definición 3.2.6. *Sean $f\colon \mathbb{R}^n \to \mathcal{I}^m$, x_0 y $u \in \mathbb{R}^n$. La derivada direccional de f en el punto x_0 y en la dirección u, denotada por medio del símbolo $f'(x_0;u)$, se define por la ecuación*

$$f'(x_0;u) = \lim_{\lambda\to 0^+} \frac{f(c+\lambda u)\ominus f(c)}{\lambda} = \lim_{\lambda\to 0^+} \frac{f(c)\ominus f(c-\lambda u)}{\lambda} \tag{3.26}$$

siempre que los límites de la derecha existan y sean iguales.

Los limites anteriores en término de la métrica Hausdorff infinito significan:

$$\lim_{\lambda\to 0^+} d_{H_\infty}\left(\frac{f(c+\lambda u)\ominus f(c)}{\lambda}, f'(x_0;u)\right) = 0$$

y

$$\lim_{\lambda\to 0^+} d_{H_\infty}\left(\frac{f(c)\ominus f(c-\lambda u)}{\lambda}, f'(x_0;u)\right) = 0$$

Cuando elegimos a u como uno de los vectores e_k de la base canónica de $\mathbb{R}^n$, la derivada direccional recibe el nombre de derivada parcial y la denotaremos por $\partial_k f(x_o) = f'(x_0;e_k)$ donde el subíndice se refiere al k-ésimo vector de la base canónica. Análogamente al cálculo, si $f\colon \mathbb{R}^n \to \mathcal{I}^m$ es una función multi-intervalo valuada, $f = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}$, con $f_j\colon \mathbb{R}^n \to \mathcal{I}$, para $j = 1,\cdots,m$, y una de las derivadas parciales $\partial_1 f,\cdots,\partial_n f$ existe en x_0 y las restantes $n-1$ derivadas parciales existen en una cierta bola $B_r(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$ [1] y son continuas en x_0, entonces se puede construir la derivada total $\delta f(x_0;h)$, con $h \in \mathbb{R}^n$, representada por medio de una matriz, con intervalos como coeficientes, en la siguiente forma:

$$\delta f(x_0;h) = \begin{bmatrix} \partial_1 f_1(x_0) & \partial_2 f_1(x_0) & \cdots & \partial_n f_1(x_0) \\ \partial_1 f_2(x_0) & \partial_2 f_2(x_0) & & \partial_n f_2(x_0) \\ & & \ddots & \\ \partial_1 f_m(x_0) & \partial_2 f_m(x_0) & \cdots & \partial_n f_m(x_0) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \tag{3.27}$$

Se puede observar que esto se ha construido a partir del teorema de condición suficiente para la diferenciación total en funciones vectoriales reales, el cual enunciamos a continuación.

Proposición 3.2.7 (Apostol, 2^a edición, Teorema 12.2)**.** *Sean $f\colon \mathbb{R}^n \to \mathbb{R}^m$ una función y $x_0 \in \mathbb{R}^n$. Supongamos que una de las derivadas parciales $\partial_1 f,\cdots,\partial_n f$ existe en x_0 y que las restantes $n-1$ derivadas parciales existen en una cierta bola $B_r(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$ y son continuas en x_0. Entonces f es diferenciable total en x_0.*

Ahora, lo que estamos diciendo es que, si una función multi-intervalo valuada f esta definida sobre un dominio abierto X y es diferenciable total en todo X, entonces en un punto fijo $x_0 \in X$, la diferencial total tiene la forma $\delta f(x_0;h) = A_{x_0}h$, donde A_{x_0} tiene representación matricial en la base canónica dada por (3.27). Así, cuando x_0 varia sobre X, la correspondencia $x_0 \to A_{x_0}$

[1] la norma en $\mathbb{R}^n$ puede ser cualquiera, por esta razón no hacemos ninguna aclaración

define una trasformación de X sobre el conjunto de funciones de $\mathbb{R}^n$ a $\mathcal{I}^m$. Cuando $\delta f(x_0; h)$ existe se dice que f **es H-diferenciable en** x_0

Además, tenemos el siguiente resultado, que nos relaciona las dos clases de diferenciación que hemos estudiado hasta ahora, la diferenciación débil y la H-diferenciación o diferenciación fuerte.

Proposición 3.2.8. *Sea $f\colon \mathbb{R}^n \to \mathcal{I}^m$, una función multi-intervalo valuada, si f es conicamente debílmente diferenciable en x_0 entonces f es H-diferenciable en x_0.*

Demostración. Como f es conicamente debílmente diferenciable, entonces

$$\delta \widetilde{f}(x_0; h) = \sum_{i=1}^{n} h_i \prod_{j=1}^{m} \left[\frac{\partial f_j^L}{\partial x_i}(x_0), \widetilde{\frac{\partial f_j^U}{\partial x_i}(x_0)} \right]$$

y esto satisface que

$$\lim_{\|h\| \to 0} \frac{d_{H_\infty}\left(f(x_0 + h) - \sum_{h_i < 0} h_i \prod_{j=1}^{m} \left[\frac{\partial f_j^L}{\partial x_i}(x_0), \frac{\partial f_j^U}{\partial x_i}(x_0) \right], f(x_0) + \sum_{h_i \geq 0} h_i \prod_{j=1}^{m} \left[\frac{\partial f_j^L}{\partial x_i}(x_0), \frac{\partial f_j^U}{\partial x_i}(x_0) \right] \right)}{\|h\|} = 0$$

en particular para $h = \lambda e_k$, se tiene

$$\lim_{\lambda \to 0^+} \frac{d_{H_\infty}\left(f(x_0 + \lambda e_k), f(x_0) + \lambda A^{e_k} \right)}{\lambda} = 0$$

y

$$\lim_{\lambda \to 0^-} \frac{d_{H_\infty}\left(f(x_0 + \lambda e_k) - \lambda A^{e_k}, f(x_0) \right)}{-\lambda} = 0$$

donde $A^{e_k} = \prod_{j=1}^{m} \left[\frac{\partial f_j^L}{\partial x_k}(x_0), \frac{\partial f_j^U}{\partial x_k}(x_0) \right]$, de lo cual se obtiene:

$$\lim_{\lambda \to 0^+} d_{H_\infty}\left(\frac{f(x_0 + \lambda e_k) \ominus f(x_0)}{\lambda}, A^{e_k} \right) = 0$$

y

$$\lim_{\lambda \to 0^+} d_{H_\infty}\left(A^{e_k}, \frac{f(x_0) \ominus f(x_0 - \lambda e_k)}{\lambda}, \right) = 0$$

se concluye el resultado. □

Ejemplo 3.2.1. *Sea $f(x) = \begin{bmatrix} [x^2 - 1, x^4 + x + 0.5] \\ [2x^2, 3x^2] \end{bmatrix}$, las gráficas de las componentes está en la gráfica 3.1.*

Observe que $\delta f^L(x_0; h) = h \begin{bmatrix} 2x_0 \\ 4x_0 \end{bmatrix}$ y $\delta f^U(x_0; h) = h \begin{bmatrix} 4x_0^3 + 1 \\ 6x_0 \end{bmatrix}$, puesto que

$$\lim_{h \to 0^+} \frac{f_1(x_0 + h) \ominus f_1(x_0)}{h} = \lim_{h \to 0^+} \frac{f_1(x_0) \ominus f_1(x_0 - h)}{h} = \left[2x_0, 4x_0^3 + 1 \right]$$

solo cuando $x_0 \geq -0.88465$ aproximadamente, y

$$\lim_{h \to 0^+} \frac{f_2(x_0 + h) \ominus f_2(x_0)}{h} = \lim_{h \to 0^+} \frac{f_2(x_0) \ominus f_2(x_0 - h)}{h} = [4x_0, 6x_0]$$

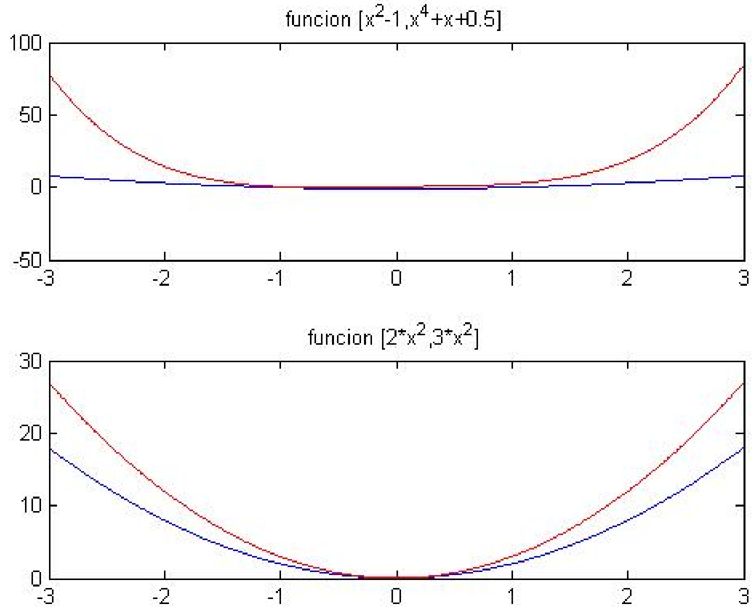


Figura 3.1: Gráficas de las componentes.

solo cuando $x_0 \geq 0$*, por último ambas derivadas parciales son continuas para* $x_0 \geq 0$ *entonces* f *es H-diferenciable para* $x_0 \geq 0$ *y su H-derivada esta dada por:*

$$\delta f\left(x_0; h\right) = \begin{bmatrix} \left[2x_0, 4x_0^3 + 1\right] \\ \left[4x_0, 6x_0\right] \end{bmatrix} h$$

Pasamos decribir el problema central de este trabajo.

3.3. El Problema Central y su Solución

Sea $f\colon \mathbb{R}^n \to \mathcal{I}^m$ una función multi-intervalo valuada,

$f\left(x\right) = \begin{bmatrix} f_1\left(x\right) \\ \vdots \\ f_m\left(x\right) \end{bmatrix}$ donde cada $f_j\colon \mathbb{R}^n \to \mathcal{I}$, $j = 1, \cdots, m$ son funciones intervalo-valuadas.

El problema

$$\begin{array}{cc} \text{mín} & f\left(x\right) \\ \text{s.a} & \begin{array}{c} g_i\left(x\right) \leq 0,\ i = 1, \cdots, k \\ x \in S \end{array} \end{array} \tag{3.28}$$

contiene algunos elementos que matemáticamente no han sido precisados.

Al conjunto de todos los puntos x que satisfacen las ecuaciones

$$\begin{array}{c} g_i\left(x\right) \leq 0,\ i = 1, \cdots, k \\ x \in S \end{array}$$

se le llama conjunto o región factible, denotada por $\mathcal{F}$. Cada criterio f_j puede ser expresado como una función de las variables decisión y así f_j es llamada una función objetivo, con lo cual f es una función multi-objetivo.

En este escrito se usará el criterio de Pareto descrito en la sección 2.7 para optimización multi-objetivo real, así el problema 3.28 se escribe como:

$$\begin{array}{cc} \text{mín}_{Par} & f(x) \\ \text{s.a} & g_i(x) \leq 0,\, i = 1, \cdots, k \\ & x \in S \end{array} \tag{3.29}$$

Algunos de los resultados necesarios para resolver el problema 3.28, han sido desarrollados en el capítulo 2 y resaltamos que los conceptos y técnicas que se usarán en este problema, se basan en las técnicas de la optimización multiobjetivo y la optimización de funciones mono-objetivo intervalo valuadas.

Pasamos ahora a definir la siguiente terminología:

Definición 3.3.1. *$z^* \in \mathcal{Z}_{\mathcal{F}}$ es fuertemente no dominado de tipo I si y sólo si no existe $z \in \mathcal{Z}_{\mathcal{F}}$ tal que $z \prec_{LU} z^*$.*

En otras palabras, un punto no dominado de tipo I es tal que, cualquier otro punto de $\mathcal{Z}_{\mathcal{F}}$ el cual "disminuye" el valor de uno de los criterios, también "aumenta" al menos uno de los otros criterios en el orden parcial $\prec_{LU}$. Esto corresponde al mismo razonamiento de minimizar en el sentido de Vilfredo Pareto de MOP. Otras definiciones equivalentes:

1. No existe $x \in \mathcal{F}$ tal que $f_j(x) \preceq_{LU} f_j(x^*)$ para todo $j = 1, \cdots, m$ y $f_j(x) \neq f_j(x^*)$ para algún $j \in \{1, \cdots, m\}$.
2. Si $f(x) \preceq_{LU} f(x^*)$ para algún $x \in \mathcal{F}$ entonces $f(x) = f(x^*)$.
3. No existe $x \in \mathcal{F}$ tal que $f_j^{\eta}(x) \leq f_j^{\eta}(x^*)$ para todo $j = 1, \cdots, m,\quad \overline{\eta} \in \{L, U\}$ y $f_j^{\overline{\eta}}(x) \neq f_j^{\overline{\eta}}(x^*)$ para algún $j \in \{1, \cdots, m\}$, $\overline{\eta} \in \{L, U\}$.

Además, tenemos una relación importante.

1. Si $z^* \in \mathcal{Z}_{\mathcal{F}}$ es no dominado de tipo I entonces no existen $d \in K_{\preceq_{LU}, \mathbf{0}}$ y $x \in \mathcal{F}$ tal que $f(x) + d = z^*$.

Al punto x^* tal que $z^* = f(x^*)$, se le llama óptimo de pareto de tipo I si z^* es no dominado de tipo I, y a z^* se le llama un punto eficiente de tipo I. Si x^1 y $x^2 \in \mathcal{F}$ y $f\left(x^1\right) \prec_{LU} f\left(x^2\right)$, decimos que x^1 domina de tipo I a x^2 y que $f\left(x^1\right)$ domina de tipo I a $f\left(x^2\right)$. Al conjunto de todos los puntos óptimos de pareto de tipo I, $x^* \in \mathcal{F}$, se le denota $\mathcal{F}_{Par-I}$, y llamamos al conjunto $\mathcal{Z}_{eff-I} = f(\mathcal{F}_{Par-I})$ el conjunto eficiente (frontera de Pareto) de tipo I, quien es el conjunto de todas las imagenes de los óptimos de Pareto de tipo I.

Definición 3.3.2. *$z^* \in \mathcal{Z}_{\mathcal{F}}$ es debílmente no dominado de tipo I si y sólo si no existe $z \in \mathcal{Z}_{\mathcal{F}}$ tal que $z \ll_{LU} z^*$.*

Sobre el conjunto $\mathcal{Z}_{eff-I}$ se tiene que.

Proposición 3.3.1. *$\mathcal{Z}_{eff-I} \subset \partial \mathcal{Z}_{\mathcal{F}}$. .*

Demostración. Sea $z^* \in \mathcal{Z}_{eff-I}$ tal que $z^* \notin \partial \mathcal{Z}_{\mathcal{F}}$. Por lo tanto $z^* \in int\left(\mathcal{Z}_{\mathcal{F}}\right)$, luego existe $\varepsilon > 0$, tal que $B\left(z^*, \varepsilon\right) \subset \mathcal{Z}_{\mathcal{F}}$, y dado que d_{H_∞} es invariante bajo traslaciones, se tiene que, $B\left(z^*, \varepsilon\right) = z^* + B\left(\mathbf{0}, \varepsilon\right)$, luego, como todo $\overline{z} \in B\left(\mathbf{0}, \varepsilon\right)$, satisface que $\text{máx}\left\{d_H\left(\overline{z}_j, [0,0]\right)\right\} < \varepsilon$, para $j = 1, \cdots, m$, esto es equivalente a que $\text{máx}\left\{\left|\overline{z}_j^L\right|, \left|\overline{z}_j^U\right|\right\} < \varepsilon$, para $j = 1, \cdots, m$, así, podemos elegir $0 < d_j^L < \alpha$ y $d_j^L \leq d_j^U < \alpha$, donde $\alpha = \text{mín}\left\{\frac{\varepsilon}{2}, \frac{\left(z_j^U\right)^* - \left(z_j^L\right)^*}{2}\right\}$ para $j = 1, \cdots, m$, y se obtiene que $z = z^* \ominus d \in B\left(z^*, \varepsilon\right) \subset \mathcal{Z}_{\mathcal{F}}$, donde $d = \begin{bmatrix} \left[d_1^L, d_1^U\right] \\ \vdots \\ \left[d_m^L, d_m^U\right] \end{bmatrix}$, además, como $d \in B\left(\mathbf{0}, \varepsilon\right)$ y $d \in K_{\preceq_{LU}, \mathbf{0}}$, se tiene que $z + d = z^* \in \mathcal{Z}_{\mathcal{F}}$, o sea, $z^* \notin \mathcal{Z}_{eff-I}$. □

Como consecuencia de lo anterior, se sigue el siguiente resultado.

Corolario 3.3.1. *Si $\mathcal{Z}_{\mathcal{F}}$ es abierto entonces $\mathcal{Z}_{eff-I} = \phi$.*

Demostración. Si $\mathcal{Z}_{\mathcal{F}}$ es abierto, $\partial \mathcal{Z}_{\mathcal{F}} = \phi$, y por la proposición anterior, se sigue el resultado. □

En lo que sigue, mostraremos un **teorema de condición necesaria** para el cálculo de los elementos del conjunto $\mathcal{Z}_{eff-I}$.

Consideremos el siguiente problema de optimización(weighted sum):

$$\text{mín}_{x \in \mathcal{F}} \sum_{j=1}^{m} \lambda_j f_j\left(x\right) \tag{3.30}$$

según el orden parcial $\prec_{LU}$, donde λ_j son escalares no negativos tales que $\sum_{j=1}^{m} \lambda_j = 1$. Notemos que este problema es una problema de optimización con función mono-objetivo intervalo valuada.

Entonces se obtiene el siguiente resultado.

Teorema 3.3.1. *Sean $\lambda_j \geq 0$, $j = 1, \cdots, m$ tal que $\sum_{j=1}^{m} \lambda_j = 1$. Entonces si x^* es solución óptima de tipo I de (3.30) entonces x^* es solución óptima de tipo I de (3.29).*

Demostración. Sea x^* una solución óptima de tipo I de (3.30) y que x^* no es solución de tipo I de (3.29), entonces existe un $x \in \mathcal{F}$ que domina de tipo I a x^*, esto es, $f\left(x\right) \preceq_{LU} f\left(x^*\right)$ así, $f_j\left(x\right) \preceq_{LU} f_j\left(x^*\right)$ para todo $j = 1, \cdots, m$, y por la no negatividad de los λ_j y la homogeneidad del orden parcial por la multiplicación por escalar no negativo, se sigue que $\lambda_j f_j\left(x\right) \preceq_{LU} \lambda_j f_j\left(x^*\right)$, y como también $\preceq_{LU}$ es compatible con la suma se sigue $\sum_{j=1}^{m} \lambda_j f_j\left(x\right) \preceq_{LU} \sum_{j=1}^{m} \lambda_j f_j\left(x^*\right)$, así, x^* no es solución no dominada de tipo I de (3.29), lo cual es contradictorio con el supuesto. De esta forma se obtiene el resultado. □

Además introducimos un segundo concepto de solución siguiendo Ishibuchi and Tanaka [4] en el cual genera el orden parcial $\preceq_{UC}$, así el problema de optimización multi-intervalo valuado (3.29) se le asocia otro tipo de solución:

Definición 3.3.3. *$z^* \in \mathcal{Z}_{\mathcal{F}}$ es fuertemente no dominado de tipo II si y sólo si no existe $z \in \mathcal{Z}_{\mathcal{F}}$ tal que $z \prec_{LU} z^*$ o $z \prec_{CW} z^*$.*

En otras palabras, un punto no dominado tipo II es tal que cualquier otro punto de $\mathcal{Z}_{\mathcal{F}}$, el cual "disminuye" el valor de uno de los criterios, también "aumenta" al menos uno de los otros criterios en el orden $\prec_{LU}$ o $\prec_{CW}$. Como está probado en ([2]), a cada punto $z = \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$ y $y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$, se puede asociar el orden parcial $\prec_{UC}$, definido por

$$y \preceq_{UC} z \text{ si y sólo si } y_j^U \leq z_j^U \text{ y } y_{cj} \leq z_{cj} \text{ para todo } j = 1, \cdots, m$$

y

$$y \prec_{UC} z \text{ si y sólo si } y \preceq_{UC} z \text{ y } z \neq y$$

cumple que si $y_j \preceq_{UC} z_j$ es equivalente a $y_j \preceq_{LU} z_j$ o $y_j \preceq_{CW} z_j$ demostrado en [2], y como es para todo j, esto es equivalente a $z \preceq_{LU} z^*$ o $z \preceq_{CW} z^*$, análogamente se sigue $y \prec_{UC} z$ es equivalente a $z \prec_{LU} z^*$ o $z \prec_{CW} z^*$, además como los ordenes $\preceq_{LU}$ y $\preceq_{CW}$ son compatibles con la suma y el producto por escalar positivo, este orden también lo es.

Otras definiciones equivalentes:

1. No existe $x \in \mathcal{F}$ tal que $f_j(x) \preceq_{UC} f_j(x^*)$ para todo $j = 1, \cdots, m$ y $f_i(x) \neq f_i(x^*)$ para algún $i \in \{1, \cdots, m\}$.
2. Si $f(x) \prec_{UC} f(x^*)$ para algún $x \in \mathcal{F}$ entonces $f(x) = f(x^*)$.
3. No existe $x \in \mathcal{F}$ tal que $f_j^{\eta}(x) \leq f_j^{\eta}(x^*)$ para todo $j = 1, \cdots, m$,
$$\overline{\eta} \in \{U, c\}$$
y $f_i^{\overline{\eta}}(x) \neq f_i^{\overline{\eta}}(x^*)$ para algún $i \in \{1, \cdots, m\}$, $\overline{\eta} \in \{U, c\}$.

Al punto x^* tal que $z^* = f(x^*)$, se llama óptimo de pareto tipo II si z^* es no dominado de tipo II, y a z^* se le llama un punto eficiente de tipo II. Si x^1 y $x^2 \in \mathcal{F}$ y $f(x^1) \prec_{UC} f(x^2)$ decimos que x^1 domina de tipo II a x^2, y que $f(x^1)$ domina de tipo II a $f(x^2)$. Al conjunto de todos los puntos óptimos de pareto de tipo II, $x^* \in \mathcal{F}$, se le denota como $\mathcal{F}_{Par-II}$ y llamamos al conjunto $\mathcal{Z}_{eff-II} = f(\mathcal{F}_{Par-II})$ el conjunto eficiente de tipo II, quuien es el conjunto de todas las imagenes de los óptimos de Pareto de tipo II.

Definición 3.3.4. *$z^* \in \mathcal{Z}_{\mathcal{F}}$ es debîlmente no dominado de tipo II si y sólo si no existe $z \in \mathcal{Z}_{\mathcal{F}}$ tal que $z \prec_{UC} z^*$.*

Sobre $\mathcal{Z}_{eff-II}$ se tiene que:

Proposición 3.3.2. *$\mathcal{Z}_{eff-II} \subset \partial\mathcal{Z}_{\mathcal{F}}$, donde $\partial\mathcal{Z}_{\mathcal{F}}$.*

Demostración. Sea $z^* \in \mathcal{Z}_{eff-II}$ tal que $z^* \notin \partial\mathcal{Z}_{\mathcal{F}}$. Por lo tanto $z^* \in int(\mathcal{Z}_{\mathcal{F}})$, luego existe $\varepsilon > 0$, tal que $B(z^*, \varepsilon) \subset \mathcal{Z}_{\mathcal{F}}$, y dado que d_{H_∞} es invariante bajo traslaciones, se tiene que $B(z^*, \varepsilon) = z^* + B(\mathbf{0}, \varepsilon)$. Ahora, como todo $\overline{z} \in B(\mathbf{0}, \varepsilon)$, si se considera el orden $\preceq_{LU}$, el resultado se obtiene de la proposición 3.3.1 ; además si se asocia a $\overline{z}$ la pareja $(\overline{z}_c, \overline{z}_w)$, se cumple

$$|\overline{z_{cj}}| = \left| \frac{\overline{z_j}^L + \overline{z_j}^U}{2} \right| \leq \frac{\left|\overline{z_j}^L\right| + \left|\overline{z_j}^U\right|}{2}$$

y

$$|\overline{z_{wj}}| = \left|\frac{\overline{z_j}^U - \overline{z_j}^L}{2}\right| = \frac{\overline{z_j}^U - \overline{z_j}^L}{2}$$

para todo $j = 1, \cdots, m$, pero como $\overline{z} \in B(\mathbf{0}, \varepsilon)$, se tiene que $\left|\overline{z_j}^L\right|, \left|\overline{z_j}^U\right| < \varepsilon$, de lo cual es claro que $|\overline{z_{cj}}| < \varepsilon$, para todo $j = 1, \cdots, m$; para la segunda observe que $-\varepsilon < -\overline{z_j}^L < \varepsilon$ y $-\varepsilon < \overline{z_j}^U < \varepsilon$ de lo cual se sigue, $0 \leq \overline{z_j}^U - \overline{z_j}^L < 2\varepsilon$ y así se concluye $|\overline{z_{cj}}| < \varepsilon$ para todo $j = 1, \cdots, m$, con esto podemos asociar un punto $d = \{d_c, d_w\}$ donde d_c y $d_w \in \mathbb{R}^m_+$ con $d_c, d_w < \varepsilon$, el cual pertenece al conjunto $B(\mathbf{0}, \varepsilon)$, y por lo tanto el punto dado por $z = \{(z_c^*) - d_c, (z_w^*) - d_w\} \in B(z^*, \varepsilon)$ y cumple que $z_c \leq (z_c^*)$ y $z_w \leq (z_w^*)$, o sea, $z^* \notin \mathcal{Z}_{eff-II}$. □

Como consecuencia de esto se sigue.

Corolario 3.3.2. *Si $\mathcal{Z}_{\mathcal{F}}$ es abierto entonces $\mathcal{Z}_{eff-II} = \phi$.*

Demostración. Si $\mathcal{Z}_{\mathcal{F}}$ es abierto, $\partial\mathcal{Z}_{\mathcal{F}} = \phi$ y por proposición anterior se sigue el resultado. □

En lo que sigue, mostraremos otro **teorema de condición necesaria** para el cálculo de los elementos del conjunto $\mathcal{Z}_{eff-II}$.

Consideremos el siguiente problema de optimización(weighted sum)

$$\text{mín}_{x\in\mathcal{F}} \sum_{j=1}^{m} \lambda_j f_j(x) \tag{3.31}$$

según el orden $\prec_{UC}$, donde λ_j son escalares no negativos tales que $\sum\limits_{j=1}^{m} \lambda_j = 1$, entonces se obtiene el siguiente resultado.

Teorema 3.3.2. *Sean $\lambda_j \geq 0$, $j = 1, \cdots, m$ tal que $\sum\limits_{j=1}^{m} \lambda_j = 1$. Entonces si x^* es solución óptima de tipo II de (3.31) entonces x^* es solución óptima de tipo II de (3.29).*

Demostración. Sea x^* una solución óptima de tipo II de (3.30) y que x^* no es solución de tipo II de (3.29), entonces existe un $x \in \mathcal{F}$ que domina a x^*, esto es, $f(x) \preceq_{UC} f(x^*)$, así, $f_j(x) \preceq_{UC} f_j(x^*)$ para todo $j = 1, \cdots, m$, y por la no negatividad de los λ_j y la homogeniedad del orden parcial $\prec_{UC}$, se sigue que $\lambda_j f_j(x) \preceq_{UC} \lambda_j f_j(x^*)$ y como $\preceq_{UC}$ es compatible con la suma, se sigue $\sum_{j=1}^{m} \lambda_j f_j(x) \preceq_{UC} \sum_{j=1}^{m} \lambda_j f_j(x^*)$, luego x^* no es solución no dominada de tipo II, contradición, con lo cual se obtiene el resultado. □

En este punto, podemos concluir que el problema 3.28 puede ser resuelto a través de los problemas descritos en (3.30) o (3.31), según el orden parcial, los cuales son problemas de optimización con función objetivo intervalo valuada, y por lo tanto, la teoría descrita en el articulo [2], cuyos resultado central se enuncia en 2.4.4, puede ser aplicada para resolver el problema (3.28) de optimización de funciones multi-objetivo intervalo valudas, si la función f es diferenciable, ya sea H-diferenciable o debílmente diferenciable para las soluciones de tipo I o soluciones de tipo II. Además, por definición las soluciones de tipo I también son soluciones de tipo II.

Capítulo 4

Métodos y Ejemplos Numéricos.

En los capítulos anteriores describimos el problema

$$\begin{array}{ll} \text{mín} & f(x) \\ \text{s.a} & \left\{ \begin{array}{l} g_i(x) \leq 0,\, i = 1, \cdots, k \\ x \geq 0 \end{array} \right. \end{array} \tag{4.1}$$

donde $f\colon \mathbb{R}^n \to \mathcal{I}^m$ y $g_i\colon \mathbb{R}^n \to \mathbb{R}$ para todo $i = 1, \cdots, k$; además se mostró que con el problema auxiliar de optimización con función mono-objectivo intervalo valuada

$$\begin{array}{ll} \text{mín} & \sum_{i=1}^{m} \alpha_i f_i(x) \\ \text{s.a} & \left\{ \begin{array}{l} g_i(x) \leq 0,\, i = 1, \cdots, k \\ x \geq 0 \end{array} \right. \end{array} \tag{4.2}$$

con $f\colon \mathbb{R}^n \to \mathcal{I}$, puede generar puntos no dominados del problema original, por tal motivo es necesario iniciar con un método para resolver el problema anterior. Recordemos que el problema auxiliar también tiene un concepto de solución multiobjetivo con los ordenes propuesto $\prec_{LU}$ y $\prec_{UC}$ con dos criterios, las funciones componentes f^L y f^U ó f^U y f_c, respetivamente.

Nosotros implementamos unos algoritmos basandonos en el toolbox de optimización de MATLAB®, no realizaremos un estudio de convergencia de los métodos puesto que no ha sido nuestro objetivo. Proponer nuevos métodos y algoritmos para aproximar soluciones no dominadas para este problema y hacer un estudio detallado sobre convergencia de(l) (los) método(s) propuesto(s) sería parte de otro trabajo.

En lo que sigue suponemos que la función multi-intervalo valuada es débilmente diferenciable y LU-convexa. Por la teoría desarrollada en los capítulos 2 y 3, se concluye que el problema auxiliar de optimización con función mono-objetivo intervalo valuada es débilmente diferenciable y LU-convexa en $\mathcal{I}$, dado que los coeficientes en la función weighted sum son positivos.

4.1. Ejemplos de problemas intervalo-valuados

Para el problema 4.2, presentamos dos programas fminintval y fminintvalga, los cuales hacen un llamado de los archivos fmincon y ga de los toolbox de optimización y algoritmos genéticos de MATLAB, usando nuevamente una función weighted sum, la cual es evaluada por medio una función anómina en MATLAB para poder ejecutar los archivos antes mencionados.

Ejemplo 4.1.1. *Sea* $f\colon \mathbb{R}^2 \to \mathcal{I}$ *con* $f(x) = \left[f^L, f^U\right] = \left[x_1^2 + x_2^2 + 1, x_1^2 + 2x_2^2 + 2\right]$, *observe que las funciones* f^L *y* f^U *son convexas y diferenciables, por lo tanto,* f *es una función débilmente diferenciable y LU-convexa. Las restricciones son:*

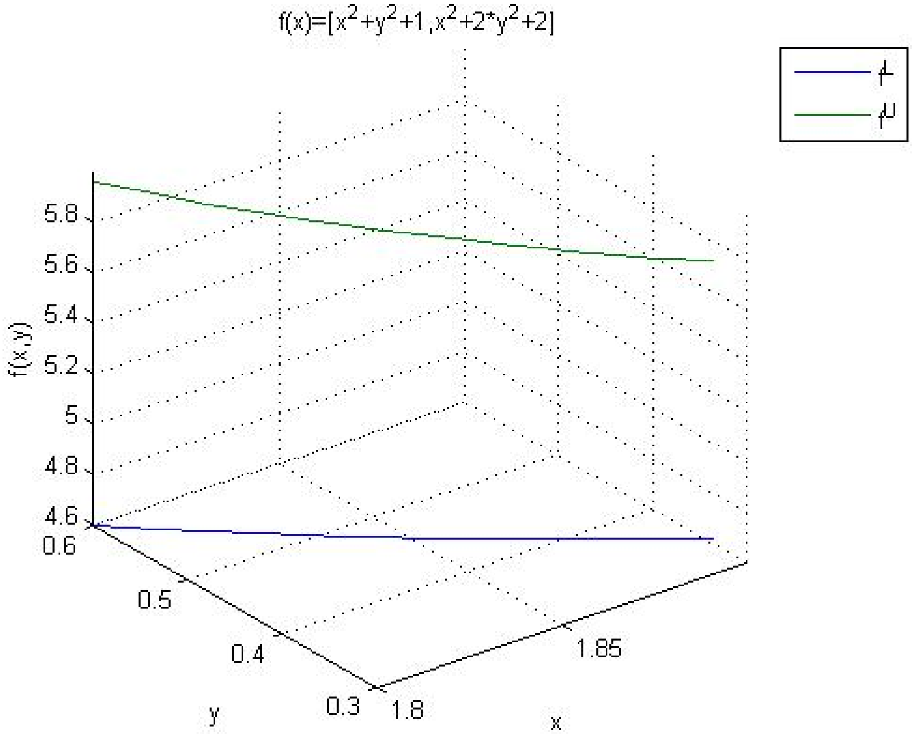
$$\begin{bmatrix} -1 & -1 \\ -6 & -2 \end{bmatrix} x \leq \begin{pmatrix} -1 \\ -12 \end{pmatrix}$$

$$\boldsymbol{x} \geq \boldsymbol{0}$$

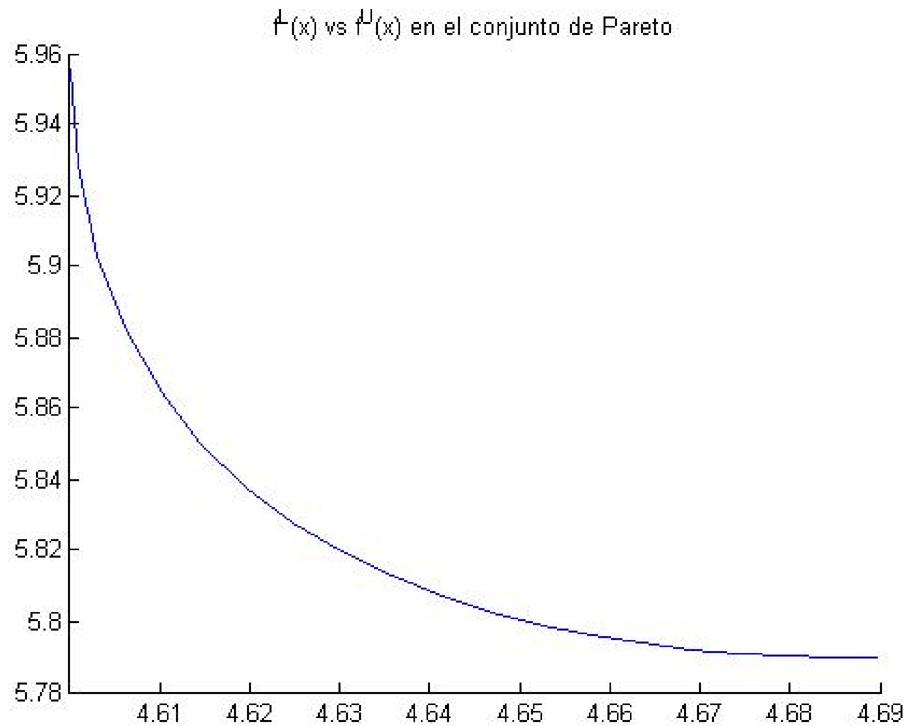
Al usar el archivo fminintval, *con punto inicial* $x_0 = (1,\ 1)^\top$, *realiza* 20 *puntos, obtenemos:*
» [x fval]=fminintval(F1,[1 1]',A,b,[0 0]',[],[],20)

x_1	x_2	$f^L(x)$	$f^U(x)$
1.8947	0.3159	4.6897	5.7895
1.8920	0.3240	4.6846	5.7896
1.8892	0.3325	4.6795	5.7901
1.8862	0.3414	4.6743	5.7909
1.8830	0.3509	4.6690	5.7921
1.8797	0.3609	4.6635	5.7938
1.8762	0.3714	4.6581	5.7960
1.8725	0.3826	4.6525	5.7989
1.8685	0.3945	4.6469	5.8026
1.8643	0.4072	4.6413	5.8071
1.8598	0.4207	4.6357	5.8127
1.8550	0.4351	4.6302	5.8195
1.8498	0.4505	4.6248	5.8278
1.8443	0.4671	4.6196	5.8378
1.8384	0.4849	4.6147	5.8499
1.8319	0.5042	4.6102	5.8644
1.8250	0.5251	4.6062	5.8819
1.8174	0.5477	4.6030	5.9030
1.8092	0.5724	4.6008	5.9285
1.8002	0.5995	4.6000	5.9594

Su gráfica en 3 dimensiones es:



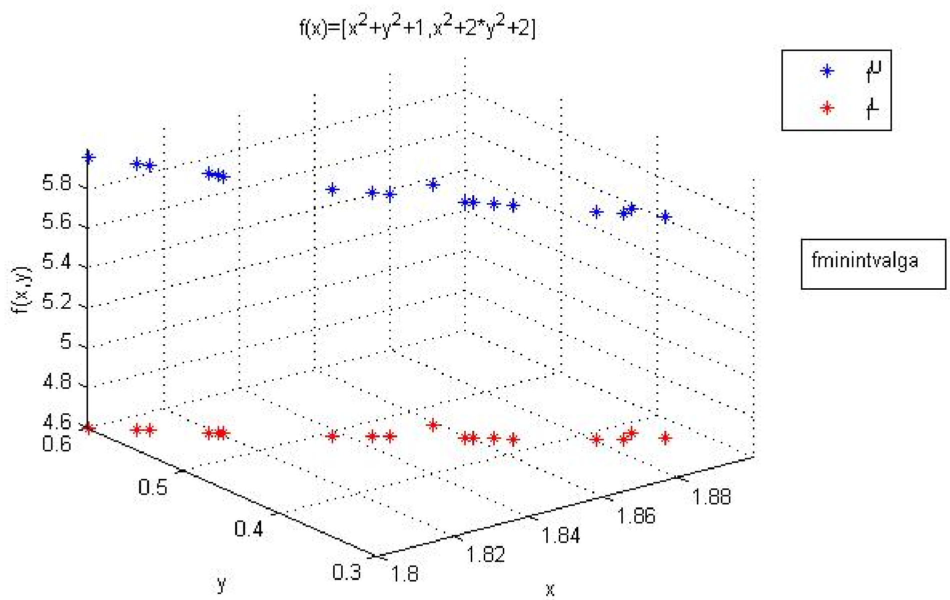
y al comparar las funciones $f^L(x)$ y $f^U(x)$ restringidas al conjunto de Pareto, se tiene:



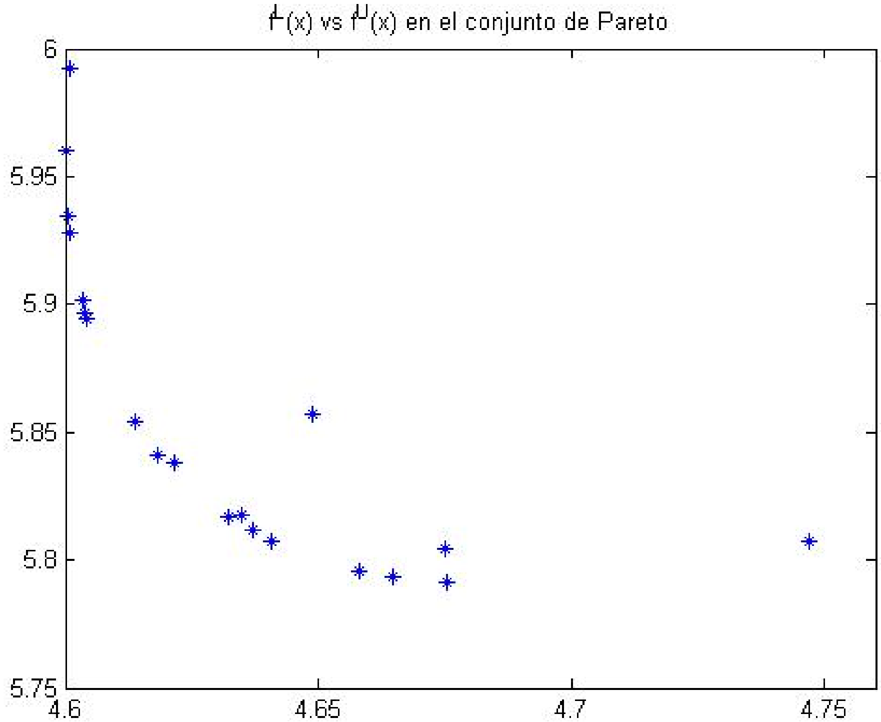
Ejemplo 4.1.2. *Al repetir el ejemplo anterior con el programa* fminintvalga *se tiene:*
» [x1 fval1]=fminintvalga(F1,2,A,b,[0 0]',[],[],20)

x_1	x_2	$f^L(x)$	$f^U(x)$
1.8548	*0.4567*	*4.6488*	*5.8574*
1.8762	*0.3713*	*4.6581*	*5.7960*
1.8580	*0.4275*	*4.6348*	*5.8175*
1.8867	*0.3402*	*4.6753*	*5.7910*
1.8803	*0.3590*	*4.6646*	*5.7934*
1.7916	*0.6255*	*4.6011*	*5.9923*
1.9201	*0.2459*	*4.7471*	*5.8076*
1.8454	*0.4649*	*4.6216*	*5.8378*
1.8639	*0.4083*	*4.6408*	*5.8076*
1.8608	*0.4176*	*4.6371*	*5.8114*
1.8567	*0.4299*	*4.6321*	*5.8170*
1.8094	*0.5719*	*4.6009*	*5.9280*
1.8426	*0.4722*	*4.6181*	*5.8411*
1.8828	*0.3604*	*4.6749*	*5.8048*
1.8073	*0.5782*	*4.6005*	*5.9349*
1.8368	*0.4900*	*4.6139*	*5.8540*
1.8196	*0.5413*	*4.6039*	*5.8969*
1.8204	*0.5389*	*4.6041*	*5.8946*
1.8181	*0.5459*	*4.6035*	*5.9015*
1.8000	*0.6000*	*4.6000*	*5.9600*

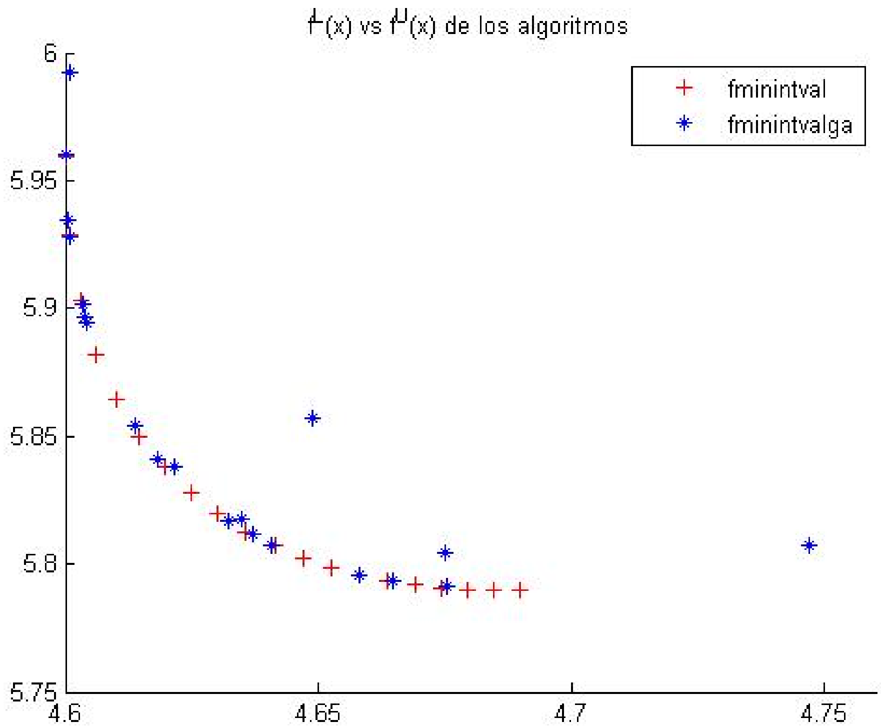
Su gráfica en 3 dimensiones es:



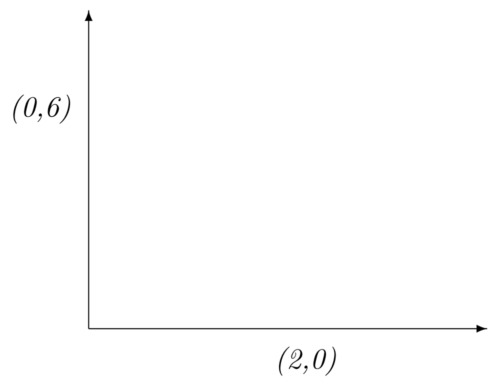
y al comparar las funciones $f^L(x)$ y $f^U(x)$ restringidas al conjunto de Pareto, se tiene:



y al comparar los resultados se observa mejor al comparar las aproximaciones de $f^L(x)$ *y* $f^U(x)$ *restringidas al conjunto de Pareto dada por los dos algoritmo.*



En este problema, sabemos que el conjunto de Pareto es la recta que une los puntos $(2,0)^\top$ *y* $(0,6)^\top$*, pues la región factible:*

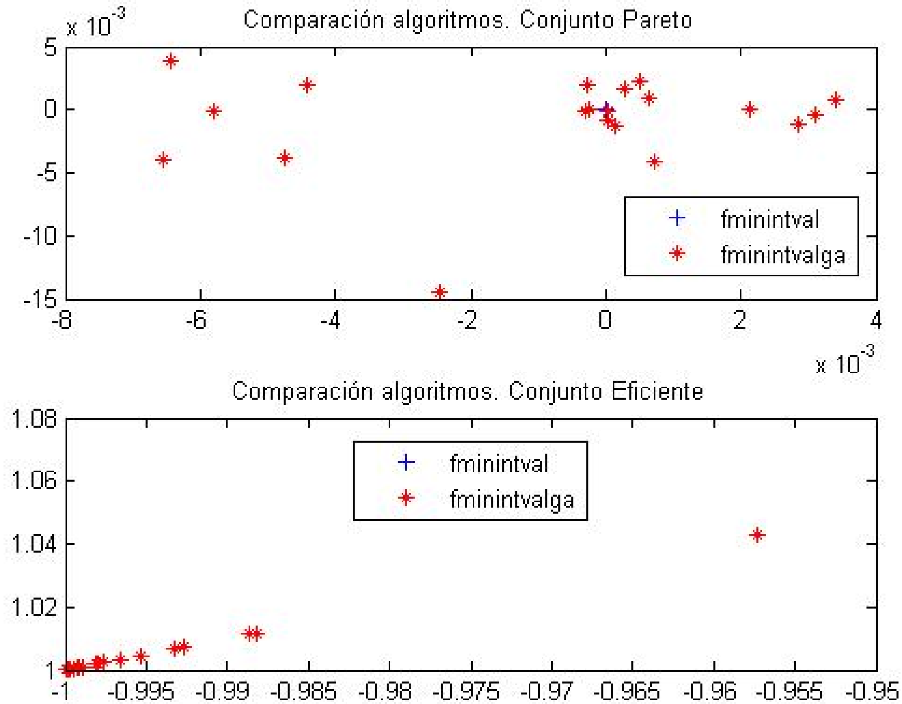


Por último, realizamos un ejemplo el cual no es convexo, con una función conocida, la función de Rastrigin, $Ras(x) = 20 + x_1^2 + x_2^2 - 10\left(\cos(2\pi x_1) + \cos(2\pi x_1)\right)$ la cual tiene un mínimo global en $(0,0)^\top$, la cual es muy usada como función de prueba para algoritmos genéticos y otras clases de algoritmos.

Ejemplo 4.1.3. *Sea*

$$f(x) = \left[19 + x_1^2 + x_2^2 - 10\left(\cos(2\pi x_1) + \cos(2\pi x_1)\right), 21 + x_1^2 + 2x_2^2 - 10\left(\cos(2\pi x_1) + \cos(2\pi x_1)\right)\right]$$

sujeto a $-5 \leq x_1 \leq 5$ *y* $-5 \leq x_2 \leq 5$. *Iniciamos con el programa* fminintval *ejecutando el comando:*
*»[x fval]=fminintval(F2,[-4 5]',[],[],-5*ones(2,1),5*ones(2,1),[],20);*
y
*»[xr1 fvalr1]=fminintvalga(F2,2,[],[],-5*ones(2,1),5*ones(2,1),[],20)*
al comparar los resultados se obtiene:



Este ejemplo es muy particular, la escogencia de las funciones f^L y f^U, hacen que coincidan los mínimos globales y está es la razón por la que ambas metodolgías aproximan el punto $(0,0)^\top$ en el espacio decisión y el intervalo $[-1,1]$ en el espacio criterio.

4.2. Ejemplos de problemas multi-intervalo-valuados

En la sección anterior, observamos que el problema auxiliar ((4.2)) al que se llega, no tiene necesariamente solución única, esto nos lleva a una pregunta muy interesante, ¿de todas las soluciones asociadas al vector $(\alpha_1, \cdots, \alpha_m)^\top$, cual escogemos?, nosotros elegimos la solución de la función f_c, al tomar, en el caso que se pida una sola solución a fminintval, la combinación convexa con $\alpha = 1/2$.

Ahora veamos el siguiente ejemplo:

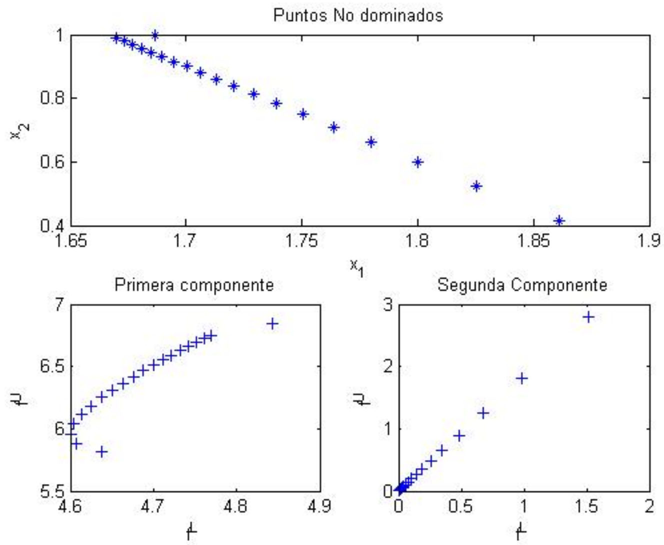
Ejemplo 4.2.1. *Sea*

$$f(x) = \begin{bmatrix} [x_1^2 + x_2^2 + 1, x_1^2 + 2x_2^2 + 2] \\ [(x_1^2+1)(x_2-1)^2, (x_1+1)^2(x_2-1)^2] \end{bmatrix}$$

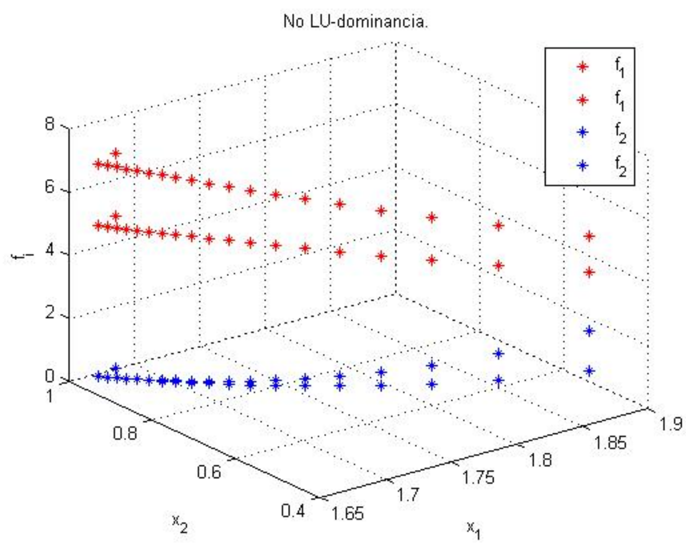
usamos las misma restricción que en el ejemplo 4.1.1 y ejecutando el programa:
» [x fval]=fminmulintval(Funmul,2,[4 5]',A,b,zeros(2,1),[],[],20)
obtenemos los resultados en la tabla de la página siguiente.

i	x_1^i	x_2^i	$f(x^i)$
1	*1.6863*	*0.9994*	$\begin{bmatrix}[4.8426, 6.8415]\\ [0.0000, 0.0000]\end{bmatrix}$
2	*1.6699*	*0.9904*	$\begin{bmatrix}[4.7693, 6.7502]\\ [0.0004, 0.0007]\end{bmatrix}$
3	*1.6732*	*0.9803*	$\begin{bmatrix}[4.7607, 6.7216]\\ [0.0015, 0.0028]\end{bmatrix}$
4	*1.6769*	*0.9693*	$\begin{bmatrix}[4.7515, 6.6910]\\ [0.0036, 0.0068]\end{bmatrix}$
5	*1.6808*	*0.9576*	$\begin{bmatrix}[4.7421, 6.6592]\\ [0.0069, 0.0129]\end{bmatrix}$
6	*1.6850*	*0.9449*	$\begin{bmatrix}[4.7322, 6.6251]\\ [0.0116, 0.0219]\end{bmatrix}$
7	*1.6896*	*0.9311*	$\begin{bmatrix}[4.7218, 6.5887]\\ [0.0183, 0.0344]\end{bmatrix}$
8	*1.6947*	*0.9160*	$\begin{bmatrix}[4.7110, 6.5501]\\ [0.0273, 0.0512]\end{bmatrix}$
9	*1.7002*	*0.8995*	$\begin{bmatrix}[4.6996, 6.5087]\\ [0.0393, 0.0737]\end{bmatrix}$
10	*1.7063*	*0.8812*	$\begin{bmatrix}[4.6878, 6.4643]\\ [0.0552, 0.1034]\end{bmatrix}$
11	*1.7130*	*0.8609*	$\begin{bmatrix}[4.6756, 6.4167]\\ [0.0762, 0.1425]\end{bmatrix}$
12	*1.7206*	*0.8381*	$\begin{bmatrix}[4.6630, 6.3654]\\ [0.1038, 0.1941]\end{bmatrix}$
13	*1.7292*	*0.8123*	$\begin{bmatrix}[4.6501, 6.3100]\\ [0.1405, 0.2623]\end{bmatrix}$
14	*1.7390*	*0.7829*	$\begin{bmatrix}[4.6372, 6.2501]\\ [0.1896, 0.3536]\end{bmatrix}$
15	*1.7504*	*0.7488*	$\begin{bmatrix}[4.6246, 6.1854]\\ [0.2564, 0.4772]\end{bmatrix}$
16	*1.7638*	*0.7087*	$\begin{bmatrix}[4.6131, 6.1154]\\ [0.3489, 0.6482]\end{bmatrix}$
17	*1.7799*	*0.6604*	$\begin{bmatrix}[4.6040, 6.0401]\\ [0.4808, 0.8914]\end{bmatrix}$
18	*1.7998*	*0.6005*	$\begin{bmatrix}[4.6000, 5.9606]\\ [0.6767, 1.2513]\end{bmatrix}$
19	*1.8256*	*0.5231*	$\begin{bmatrix}[4.6066, 5.8802]\\ [0.9857, 1.8162]\end{bmatrix}$
20	*1.8612*	*0.4163*	$\begin{bmatrix}[4.6375, 5.8108]\\ [1.5212, 2.7897]\end{bmatrix}$

La gráfica de algunos puntos no dominados de Pareto y las gráficas de las componentes, al comparar f_1^L *y* f_1^U*, como también* f_2^L *y* f_2^U*, son como siguen:*



aúnque fue más significativo en el ejemplo mono-objetivo, aquí no tiene tanta relevancia. Por el contrario al observar su gráfica en 3-D, se obtiene la visión del sentido geométrico de la soluciones de Pareto.



4.3. Método de optimización por metas.

En los problemas de optimización multiobjetivo descrito en (2.9), Gembicki [16] propuso un método de optimización por métas, el cual involucra un conjunto de metas propuestas, $f^* = \{f_1^*, \cdots, f_m^*\}$, las cuales están asociadas a un conjunto de objetivos, $f(x) = \{f_1(x), \cdots, f_m(x)\}$. En la formulación del problema se permite que los objetivos subestimen o sobre estimen las metas, de acuerdo a una imprecisión relativa alrededor de las metas iniciales que admite el decisor. El grado relativo de subestimación o sobre estimación de las metas, es dado a través de un vector de coeficientes, $\omega = (\omega_1, \cdots, \omega_m)$. La optimización por metas se presenta como un problema de optimización estándar usando la siguiente formulación:

$$\begin{array}{ll} \text{mín} & \gamma \\ \text{s. a.} & \begin{cases} \gamma \in \mathbb{R} \\ f_j(x) - \omega_j \gamma \leq f_i^* \\ g_i(x) \leq 0,\ i = 1, \cdots, k \\ x \in S \end{cases} \end{array} \tag{4.3}$$

el término $\omega_j \gamma$ nos produce o genera una cierta holgura con respecto a las metas propuestas. El vector de pesos, w, permite al decisor expresar un medida de imprecisión relativa entre los objetivos.

Basados en lo anterior y en la posibilidad de convertir funciones restricciones intervalo valuadas, a su sistema equivalente de funciones restricciones reales, podemos plantear el problema de optimización por metas como sigue:
Dados ω_j y $f_j^* \in \mathcal{I}, \quad j = 1, \cdots, m$

$$\begin{array}{ll} \text{mín} & \gamma \\ \text{s. a.} & \begin{cases} \gamma \in \mathbb{R} \\ f_j(x) \leq f_j^* + \omega_j \gamma, \quad j = 1, \cdots, m \\ g_i(x) \leq 0, \quad i = 1, \cdots, k \\ x \geq 0 \end{cases} \end{array}$$

y su problema multiobjetivo real equivalente:

$$\begin{array}{ll} \text{mín} & \gamma \\ \text{s. a.} & \begin{cases} \gamma \in \mathbb{R} \\ f_j^{\eta}(x) - \omega_j^{\eta} \gamma \leq (f_j^*)^{\eta}, \\ j = 1, \cdots, m \ \text{ y } \eta = L, U \\ g_i(x) \leq 0, \quad i = 1, \cdots, k \\ x \geq 0 \end{cases} \end{array}$$

Ejemplo 4.3.1. *Dado el problema:*

$$\begin{array}{ll} \text{mín} & f(x) = \begin{bmatrix} \left[x_1^2 + x_2^2 + 1, x_1^2 + 2x_2^2 + 2\right] \\ \left[(x_1^2 + 1)(x_2 - 1)^2, (x_1 + 1)^2 (x_2 - 1)^2\right] \end{bmatrix} \\ s.\ a. & \begin{cases} \begin{bmatrix} -1 & -1 \\ -6 & -2 \end{bmatrix} x \leq \begin{bmatrix} -1 \\ -12 \end{bmatrix} \\ x \geq 0 \end{cases} \end{array}$$

al aplicar el método de optimización por metas, con metas $f^* = \begin{bmatrix} 2, & 8, & 3, & 6 \end{bmatrix}^\top$ *e imprecisión relativa* $w = \begin{bmatrix} 1, & 1, & 1, & 1 \end{bmatrix}^\top$.

» [xMe2 fvalMe2 facme2]=fgoalattmulintval(Funmul,[4 5]',[2 8 3 6]',[1 1 1 1]',A,b,zeros(2,1),[],[],[])
Optimization terminated:
*magnitude of search direction less than 2*options.TolX and maximum constraint violation is less than options.TolCon. Active inequalities (to within options.TolCon = 1e-006):*

lower	upper	ineqlin	ineqnonlin
		2	1

$xMe2 = \begin{pmatrix} 1.8000 \\ 0.6000 \end{pmatrix}$

$fvalMe2 = \begin{bmatrix} [4.6000, 5.9600] \\ [0.6784, 1.2544] \end{bmatrix}$

$facme2 = 2.6000$

El punto $xMe2 = \begin{pmatrix} 1.8000 \\ 0.6000 \end{pmatrix}$ *es un punto de la frontera de Pareto, al verificar las condiciones sobre este punto. Además, la segunda desigualdad de las resticciones es satisfecha en igualdad, con lo cual hace parte de la frontera de la región factible siendo en ella un vértice.*

Conclusiones

En este trabajo se logra conjugar las teorías de optimización intervalo-valuada y la optimización multi-objetivo.

Se enuncia y se demuestra un teorema de diferenciación basado en el teorema de embebimiento o extensión de semigrupos topológicos en espacios vectoriales, el cual permite generar una metodología para resolver el problema propuesto en este trabajo.

Se enuncian y se demostran teoremas de condición necesaria para encontrar soluciones al problema de optimización multi-intervalo valuado.

Este trabajo se diferencia de la teoría de multifunciones, al proponer una métrica distinta a la métrica de Hausdorff, además de unos ordenes parciales diferentes de los ordenes usuales que se han usado en la optimización de multifunciones.

Se generan soluciones algoritmicas a este problema, implementados en dos métodos numéricos.

Problemas abiertos

Enumerar algunos problemas abiertos.

Generar teoremas de suficiencia para la existencia de soluciones del problema de optimización multi-intervalo valuado, con restricciones real-valuadas.

Determinar las condiciones bajo las cuales las sumas escalizantes generan todo el conjunto eficiente.

Construir algoritmos para el problema de optimización multi-intervalo valuado en el caso no convexo.

Proponer una metodología de solución para el problema de optimización multi-intervalo valuado con funciones restricciones intervalo valuadas.

Definir los conceptos de dualidad para el problema tratado en este trabajo.

Proponer algoritmos primal-dual para la hallar soluciones aproximadas.

Bibliografía

[1] Wu, Hsien-Chung, On interval-valued nonlinear programming problems, J. Math. Ana. Appl. (2007), doi:10.1016/j.jmaa.2007.05.023.

[2] Wu, Hsein-Chung, The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function, European Journal of Operational Research, 176 (2007), pag. 46-59.

[3] Alefeld, G., Hersberger, J., Introduction to interval Computations, Academic Press, 1983.

[4] Ishibuchi, H., Tanaka, H., Multiobjective programming in optimization of the interval objective function, Eurepean Journal of Operational Research 48 (1990), pp 219-225.

[5] Rådström, Hans, An embedding theorem for spaces of convex sets, American mathematical Society, Vol 3, N^o, Feb, 1953, pp165-169.

[6] Banks, H. T., Jacobs, M. Q., A differential calculus for multifunctions, Journal of Mathematical Analysis and Applications 29(1970), pp. 246-272.

[7] Luenberger, David G., Optimization by vector space methods, Jhon Wiley & Sons, Inc, 1976.

[8] Ehrgott, M., Multicriteria Optimization, University of Kaiserslautern, Deparment of Mathematics, 1999.

[9] In:Słowiński, R., Teghem, J., Stochastic versus Fuzzy Approaches to Multiobjective mathematical Prog, Editors, Kluwer Academic publishers, Boston, 1998.

[10] Programación Vectorial Lineal y Entera, Jorge S., Jesús M., Tesís Doctoral, Universidad de la Laguna, 2002.

[11] Moore, R. E., Method and Applications of Interval Analysis, SIAM, Philadelphia, 1979.

[12] Stancu-Minasian, I. M., Stochastic Programming with Multiple Objective Functions, D. Reidel Publishing Company, Dordrecht, 1984.

[13] Ekeland, I., Témam, R., Convex Analysis and Variational Problems, SIAM, 1999.

[14] Rockafellar, R. T., Convex Analysis, Princeton University Press, 1972.

[15] Apostol, T. M.; Análisis Matemático, segunda edición, Editorial Reverté, S.A, 1977.

[16] Gembicki, F.W., "Vector Optimization for Control with Performance and Parameter Sensitivity Indices,"Ph.D. Thesis, Case Western Reserve Univ., Cleveland, Ohio, 1974.

[17] Falk, J.E.;Exact solution of Inexact linear programs, Operations Research 24 (1976), pag. 783-787.