

UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE CIÊNCIAS EXATAS  
PROGRAMA DE PÓS GRADUAÇÃO EM MATEMÁTICA

*Estudos sobre Modelos de Otimização Matemática  
utilizando Funções com Parâmetros Intervalares*

Thiago Parente da Silveira

Manaus- 2017

UNIVERSIDADE FEDERAL DO AMAZONAS  
INSTITUTO DE CIÊNCIAS EXATAS  
PROGRAMA DE PÓS GRADUAÇÃO EM MATEMÁTICA

*Estudos sobre Modelos de Otimização Matemática  
utilizando Funções com Parâmetros Intervalares*

Thiago Parente da Silveira

Dissertação de mestrado apresentada ao Programa de Pós Graduação em Matemática da UFAM como requisito para a obtenção parcial do grau de MESTRE em Matemática na área de concentração OTIMIZAÇÃO.

Orientador: Prof. Dr. Sandro Dimy  
Barbosa Bitar

Manaus- 2017

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S587e Silveira, Thiago Parente da  
Estudos sobre modelos de otimização matemática utilizando  
funções com parâmetros intervalares / Thiago Parente da Silveira.  
2017  
49 f.: il. color; 31 cm.

Orientador: Sandro Dimy Barbosa Bitar  
Dissertação (Mestrado em Matemática - Otimização) -  
Universidade Federal do Amazonas.

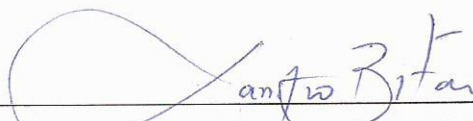
1. Otimização Intervalar. 2. Direção de Descida. 3. Função  
Intervalar Paramétrica. 4. Cálculo Intervalar. I. Bitar, Sandro Dimy  
Barbosa II. Universidade Federal do Amazonas III. Título

Thiago Parente da Silveira

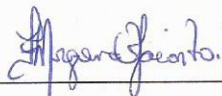
*Estudos sobre Modelos de Otimização Matemática  
utilizando Funções com Parâmetros Intervalares*

Dissertação de mestrado apresentada  
ao Programa de Pós Graduação em  
Matemática da UFAM como requisito para  
a obtenção parcial do grau de MESTRE  
em Matemática na área de concentração  
OTIMIZAÇÃO.

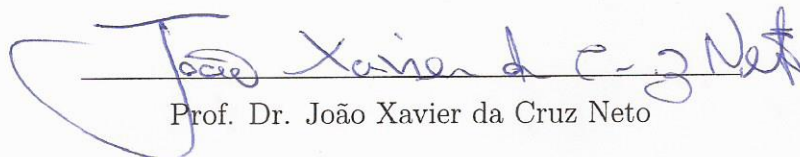
Manaus, 27 de abril de 2017



Prof. Dr. Sandro Dimy Barbosa Bitar



Prof<sup>ª</sup> Dra. Flávia Morgana de Oliveira Jacinto



Prof. Dr. João Xavier da Cruz Neto

*À Deus*

# Agradecimentos

A Deus, o maior matemático de todos. Ele nos mostrou isso ao usar a matemática como o idioma na qual escreveu o universo. Quanto mais estudamos a matemática, mais podemos conhecer à Deus.

Aos meus pais, que me apoiaram durante toda a minha vida, especialmente desde o início da minha carreira acadêmica. Que me ensinaram o quão importante são os estudos para a formação de uma pessoa e também que a disciplina é fundamental para que se alcancem os objetivos um dia planejados. Em vocês encontro forças para continuar mesmo quando já não acredito que vou conseguir avançar. Essa conquista é para vocês.

Ao Professor Doutor Sandro Bitar, que me acompanhou durante toda a minha faculdade e pós graduação. Pela sua orientação (não apenas acadêmica, mas pessoal também!), amizade e principalmente pela paciência, sem a qual este trabalho não se realizaria. Não encontro palavras no vocabulário para expressar meu sentimento de gratidão ao senhor.

Aos meus pastores Rafael Dias e Andressa Oliveira que ao longo desses anos me ajudam para o crescimento pessoal, obrigado por toda ajuda e incentivo que vocês me dão.

Aos meus amigos da pós graduação em matemática e de outros programas de pós graduação que sempre me incentivaram a seguir em frente com total determinação.

Aos professores do Programa de Pós Graduação em Matemática da Universidade Federal do Amazonas pelos seus ensinamentos e aos funcionários do programa, que durante esses anos contribuíram de algum modo para o meu enriquecimento pessoal e profissional.

Aos demais professores (que tenho a honra de tê-los como colegas de trabalho) do Departamento de Matemática da UFAM, pelo apoio dado nessa caminhada, pelos conselhos e experiências compartilhadas nas conversas na hora do almoço.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo suporte financeiro para realização deste trabalho.

A todos que contribuíram direta ou indiretamente para realização deste trabalho.

# Resumo

Nesse trabalho realizou-se um estudo sobre Otimização Intervalar, assumindo que o valor da Função Objetivo é um intervalo. Para tal, seguiu-se um desenvolvimento teórico para munir o espaço intervalar com uma estrutura de espaço vetorial parcialmente ordenado. Alguns Teoremas e Corolários, considerados essenciais à compreensão dos principais resultados, foram abordados. Condições para a existência de solução de um problema de otimização intervalar são apresentadas através de um problema auxiliar multiobjetivo, onde a função de valor intervalar é definida sob a forma paramétrica. Ao final do trabalho apresenta-se uma caracterização para direção de descida cuja função objetivo possui coeficientes intervalares, sendo o problema de otimização irrestrito e, além disso, apresenta-se também um algoritmo genérico baseado em busca direcional para esse problema.

**Palavras-chaves:** Otimização Intervalar, Direção de Descida, Função Intervalar Paramétrica, Cálculo Intervalar.

# Abstract

In this work a study on Interval Optimization was carried out, assuming that the value of the Objective Function is an interval. For this, a theoretical development is followed to provide spatial space with a partially ordered vector space structure. Some Theorems and Corollary, are essential for the understanding of the main results, were addressed. For a solution solution of an interval optimization problem are presented through a multiobjective auxiliary problem, where a variable value function is defined in a parametric form. At the end of the work we present a characterization for the direction of the descent of an object with interval coefficients, being the problem of optimization unrestricted and, in addition, a generic algorithm based on directional search for this problem is also presented.

**Keywords:** Interval Optimization, Descent Direction, Parametric Interval Function, Interval Calculus.



# Sumário

<b>Introdução</b> . . . . .	<b>1</b>
<b>1 O Espaço Vetorial <math>M^n</math></b> . . . . .	<b>3</b>
1.1 Fundamentos da aritmética intervalar . . . . .	3
1.1.1 Operações aritméticas . . . . .	3
1.1.2 Propriedades Algébricas . . . . .	8
1.2 Espaço Vetorial Intervalar Generalizado $n$ Dimensional . . . . .	10
1.2.1 Bijeção entre $M^n$ e $\mathbb{R}^{2n}$ . . . . .	10
1.3 Topologia em $M^n$ . . . . .	13
1.3.1 Métricas em $M^n$ . . . . .	14
1.4 Relação de Ordem . . . . .	15
<b>2 Elementos de Otimização Intervalar</b> . . . . .	<b>20</b>
2.1 Funções intervalares . . . . .	20
2.2 Cálculo em ambiente intervalar . . . . .	21
2.3 Funções intervalares convexas . . . . .	25
2.3.1 Relação entre função intervalar convexa e matriz definida positiva	27
2.4 Problema de Otimização Intervalar . . . . .	29
2.5 Problema de Programação Convexa Intervalar . . . . .	31
<b>3 Busca Direcional para Funções com Coeficientes Intervalares</b> . . . . .	<b>34</b>
3.1 Caracterização da Direção de Descida . . . . .	34
<b>Conclusão</b> . . . . .	<b>40</b>
<b>Referências</b> . . . . .	<b>41</b>

# Introdução

Os primeiros resultados importantes sobre a Análise Intervalar começaram a surgir em meados do século *XX*. Um dos precursores do desenvolvimento dessa teoria foi Moore com a publicação de um livro sobre análise intervalar em 1966 [5]. Uma das necessidades desse estudo se dava ao fato da limitação dos computadores não conseguirem representar números irracionais. No entanto, a aritmética desenvolvida para o ambiente intervalar limitava bastante as operações algébricas.

Em 2014, Costa [2] apresentou um método para dotar o espaço intervalar generalizado  $n$  dimensional com uma estrutura de espaço vetorial com o auxílio de uma bijeção com o espaço euclidiano  $\mathbb{R}^{2n}$ . Uma aplicação mais recente da análise intervalar está na resolução de problemas de otimização em que alguns coeficientes (sejam da função objetivo ou das restrições) resultam de valores arredondados e/ou de informações incompletas. É muito natural usar a análise intervalar para trabalhar com este tipo de problema.

Ao longo dos anos, a função de valor intervalar  $F$  com domínio em  $D \subseteq \mathbb{R}^n$  era pensada na forma  $F(x) = [\underline{f}(x), \overline{f}(x)]$  ([2], [5]), onde  $\underline{f}(x) \leq \overline{f}(x)$  para todo  $x \in D$ . Todas as propriedades da  $F$  eram obtidas a partir do estudo de  $\underline{f}$  e  $\overline{f}$ . No entanto, no desenvolvimento do cálculo intervalar alguns resultados relacionados às derivadas de  $F$  não eram totalmente explicados analisando somente as derivadas de  $\underline{f}$  e  $\overline{f}$ .

Para contornar tal situação, em 2012 Panda [1] propõe que a função de valor intervalar seja definida na forma paramétrica e assim suas propriedades sejam estudadas. A vantagem de se considerar na forma paramétrica acontece quando se analisa a positividade da matriz da Hessiana.

É possível definir relações de ordem parciais através de bijeções já conhecidas e com isso é definido um Problema de Otimização Intervalar (IOP). O problema de otimização intervalar pode ser convertido para um problema de otimização geral na forma paramétrica e com isso é possível determinar a existência de solução para o IOP.

A convexidade desempenha um papel importante para provar a existência de solução de um problema de otimização clássico. Daí a necessidade de estudar o caso convexo do problema de otimização intervalar. Como o conjunto de intervalos não é totalmente ordenado, a convexidade tem que ser estudada com respeito a uma ordem

parcial.

Em particular, é estudado também o problema de otimização quadrática intervalar. Ao final do trabalho, apresenta-se uma caracterização para direção de descida para o problema de otimização intervalar irrestrito e uma ideia para algoritmo baseado em busca direcional.

Apesar de Costa [3] e Panda [1] desenvolverem um estudo sobre otimização considerando em que a função objetivo assume valor intervalar, os resultados obtidos são diferentes devido ao fato dos autores abordarem as funções intervalares de maneiras distintas. Nesse trabalho, foi feito um detalhamento substancial sobre o trabalho de Panda [1], pois acredita-se que na prática, na maioria das vezes as incertezas recaem sobre os coeficientes da função objetivo.

# 1 O Espaço Vetorial $M^n$

Neste capítulo será feita a fundamentação da aritmética intervalar, que é a generalização da aritmética usual. Para isso, vamos estender o conceito de número real e analisar os objetos desse conjunto tratando-os como intervalos limitados da reta.

## 1.1 Fundamentos da aritmética intervalar

Inicialmente vamos definir operações aritméticas no espaço intervalar. Tais operações foram definidas por Moore [7].

### 1.1.1 Operações aritméticas

Considere os intervalos fechados e limitados da reta real que são denotados por  $[a, b]$ , que nada mais é do que o conjunto

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

Apesar de existirem outros tipos de intervalos (abertos, semi abertos), neste trabalho será considerado essencialmente intervalos fechados.

**Definição 1.1.1.** *Considere um intervalo real fechado  $X = [a, b]$ . Um número intervalar  $X$  é tal intervalo fechado, ou seja, consiste do conjunto  $\{x \in \mathbb{R} : a \leq x \leq b\}$ .*

Um *número intervalar* é também chamado simplesmente de *intervalo*. As extremidades de um número intervalar  $X$  serão denotadas por  $\underline{X}$  e  $\overline{X}$ . Assim,

$$X = [\underline{X}, \overline{X}]. \quad (1.1)$$

Dois números intervalares  $X$  e  $Y$  são ditos iguais se eles são os mesmos conjuntos. Isso acontece se suas extremidades, respectivas, são iguais:

$$X = Y \Leftrightarrow \underline{X} = \underline{Y}, \overline{X} = \overline{Y}. \quad (1.2)$$

**Definição 1.1.2.** *Dizemos que  $X$  é degenerado se  $\underline{X} = \overline{X}$ . Tal intervalo contém um único número real  $x$ . Por convenção, aceitamos a identificação de um intervalo*

degenerado  $[x, x]$  com o número real  $x$ . Nesse sentido, podemos escrever a seguinte identidade

$$0 = [0, 0]. \quad (1.3)$$

A interseção de dois intervalos  $X$  e  $Y$  é vazia se  $\overline{Y} < \underline{X}$  ou  $\overline{X} < \underline{Y}$ . Nesse caso escrevemos

$$X \cap Y = \emptyset,$$

indicando que  $X$  e  $Y$  não tem pontos em comum.

Quando  $X \cap Y \neq \emptyset$ , pode-se definir a interseção como sendo o intervalo:

$$X \cap Y = [\max\{\underline{X}, \underline{Y}\}, \min\{\overline{X}, \overline{Y}\}]. \quad (1.4)$$

Nesse caso, a união de  $X$  e  $Y$  também é um intervalo:

$$X \cup Y = [\min\{\underline{X}, \underline{Y}\}, \max\{\overline{X}, \overline{Y}\}]. \quad (1.5)$$

Em geral, a união de dois intervalos não é um intervalo. Entretanto, a *envoltória intervalar*  $X \underline{\cup} Y$  de dois intervalos, que é definida por

$$X \underline{\cup} Y = [\min\{\underline{X}, \underline{Y}\}, \max\{\overline{X}, \overline{Y}\}], \quad (1.6)$$

é sempre um intervalo. Além disso, tem-se

$$X \cup Y \subseteq X \underline{\cup} Y \quad (1.7)$$

para quaisquer intervalos  $X$  e  $Y$ .

**Exemplo 1.1.1.** Se  $X = [-1, 0]$  e  $Y = [1, 2]$ , então  $X \underline{\cup} Y = [-1, 2]$ . Apesar de  $X \cup Y$  ser um conjunto desconexo (ou seja, não é um intervalo), a relação (1.7) é mantida.

**Definição 1.1.3.** A medida  $\omega(X)$  de um intervalo  $X$  é definida por

$$\omega(X) = \overline{X} - \underline{X}. \quad (1.8)$$

**Definição 1.1.4.** O valor absoluto de  $X$ , denotado por  $|X|$ , é o máximo dos valores absolutos de suas extremidades:

$$|X| = \max\{|\underline{X}|, |\overline{X}|\}. \quad (1.9)$$

Note que  $|x| \leq |X|$  para todo  $x \in X$

**Definição 1.1.5.** O ponto médio  $m(X)$  de  $X$  é dado por

$$m(X) = \frac{1}{2}(\overline{X} + \underline{X}). \quad (1.10)$$

**Exemplo 1.1.2.** Sejam  $X = [0, 2]$  e  $Y = [-1, 1]$ . A interseção e a união de  $X$  e  $Y$  são os intervalos

$$\begin{aligned} X \cap Y &= [\max\{0, -1\}, \min\{2, 1\}] = [0, 1], \\ X \cup Y &= [\min\{0, -1\}, \max\{2, 1\}] = [-1, 2]. \end{aligned}$$

Temos  $\omega(X) = \omega(Y) = 2$  e, por exemplo,

$$|X| = \max\{0, 2\} = 2$$

O ponto médio de  $Y$  é  $m(Y) = 0$ .

A noção de intervalo degenerado permite ver o conjunto dos intervalos fechados como uma extensão do conjunto dos números reais. Dado um número real  $x$ , temos a associação

$$x \mapsto [x, x] \quad (1.11)$$

que leva os números reais a intervalos degenerados.

Vamos agora definir as operações aritméticas básicas entre intervalos. O ponto chave nessas definições é o fato de que operar intervalos é operar conjuntos.

**Definição 1.1.6.** Sejam  $X = [\underline{X}, \overline{X}]$ ,  $Y = [\underline{Y}, \overline{Y}]$  dois intervalos. A soma de  $X$  e  $Y$  é o conjunto

$$X + Y = \{x + y : x \in X, y \in Y\}. \quad (1.12)$$

que é um intervalo. De fato,

$$\underline{X} \leq x \leq \overline{X} \text{ e } \underline{Y} \leq y \leq \overline{Y}.$$

Podemos ver pela soma das inequações que  $x + y \in X + Y$  e

$$\underline{X} + \underline{Y} \leq x + y \leq \overline{X} + \overline{Y}.$$

Consequentemente, a fórmula

$$X + Y = [\underline{X} + \underline{Y}, \overline{X} + \overline{Y}] \quad (1.13)$$

pode ser usada para calcular (1.12)

**Exemplo 1.1.3.** Sejam  $X = [0, 2]$  e  $Y = [-1, 1]$  como no Exemplo 1.1.2. Então

$$X + Y = [0, 2] + [-1, 1] = [0 + (-1), 2 + 1] = [-1, 3].$$

A diferença de dois intervalos  $X$  e  $Y$  é o conjunto

$$X - Y = \{x - y : x \in X, y \in Y\}. \quad (1.14)$$

Temos

$$\underline{X} \leq x \leq \bar{X} \text{ e } -\bar{Y} \leq -y \leq -\underline{Y}$$

para obter

$$\underline{X} - \bar{Y} \leq x - y \leq \bar{X} - \underline{Y}.$$

Segue que

$$X - Y = [\underline{X} - \bar{Y}, \bar{X} - \underline{Y}] \quad (1.15)$$

Note que

$$X - Y = X + (-Y),$$

onde

$$-Y = [-\bar{Y}, -\underline{Y}] = \{y : -y \in Y\}.$$

**Exemplo 1.1.4.** Se  $X = [-1, 0]$  e  $Y = [1, 2]$ , então

$$-Y = [-2, -1]$$

$$\text{e } X - Y = X + (-Y) = [-3, -1].$$

O produto de  $X$  e  $Y$  é dado por

$$X \cdot Y = \{x \cdot y : x \in X, y \in Y\}. \quad (1.16)$$

Em termos das extremidades, o produto  $X \cdot Y$  de dois intervalos  $X$  e  $Y$  é dado por

$$X \cdot Y = [\min S, \max S], \quad (1.17)$$

onde  $S = \{\underline{X} \cdot \underline{Y}, \underline{X} \cdot \overline{Y}, \overline{X} \cdot \underline{Y}, \overline{X} \cdot \overline{Y}\}$ .

Se um dos intervalos for degenerado, tem-se

$$\alpha \cdot [\underline{X}, \overline{X}] = [\alpha \cdot \underline{X}, \alpha \cdot \overline{X}] \quad (1.18)$$

se  $\alpha \geq 0$ , ou  $\alpha \cdot [\underline{X}, \overline{X}] = [\alpha \cdot \overline{X}, \alpha \cdot \underline{X}]$  se  $\alpha < 0$ .

**Exemplo 1.1.5.** Sejam  $X = [-1, 0]$  e  $Y = [1, 2]$ . Então

$$S = \{-1 \cdot 1, -1 \cdot 2, 0 \cdot 1, 0 \cdot 2\} = \{-1, -2, 0\}$$

e  $X \cdot Y = [\min S, \max S] = [-2, 0]$ . Temos também, por exemplo,  $2Y = [2, 2] \cdot [1, 2] = [2, 4]$ .

O quociente  $X/Y$  é definido como

$$X/Y = \{x/y : x \in X, y \in Y\}, \quad (1.19)$$

desde que  $0 \notin Y$ .

Assim como acontece com os números reais, a divisão pode ser definida via multiplicação pelo inverso multiplicativo do segundo termo. Isto é, podemos calcular a equação (1.19) usando

$$X/Y = X \cdot (1/Y), \quad (1.20)$$

onde

$$1/Y = \{y : 1/y \in Y\} = [1/\overline{Y}, 1/\underline{Y}], \quad (1.21)$$

novamente, assumindo que  $0 \notin Y$ .

Qualquer intervalo  $X$  pode ser expresso como

$$\begin{aligned} X &= m(X) + [-\frac{1}{2}\omega(X), \frac{1}{2}\omega(X)] \\ &= m(X) + \frac{1}{2}\omega(X)[-1, 1]. \end{aligned} \quad (1.22)$$

De fato, se  $X = [\underline{X}, \overline{X}]$  então  $m(X) = \frac{1}{2}(\underline{X} + \overline{X})$  e  $\omega(X) = \overline{X} - \underline{X}$ . Logo



$$\begin{aligned}
m(X) + \frac{1}{2}\omega(X)[-1, 1] &= \frac{1}{2}(\underline{X} + \overline{X}) + \frac{1}{2}(\overline{X} - \underline{X})[-1, 1] \\
&= \frac{1}{2}(\underline{X} + \overline{X}) + \left[ -\frac{1}{2}(\overline{X} - \underline{X}), \frac{1}{2}(\overline{X} - \underline{X}) \right] \\
&= \left[ \frac{1}{2}(\underline{X} + \overline{X}), \frac{1}{2}(\underline{X} + \overline{X}) \right] + \left[ -\frac{1}{2}(\overline{X} - \underline{X}), \frac{1}{2}(\overline{X} - \underline{X}) \right] \\
&= \left[ \frac{1}{2}(\underline{X} + \overline{X}) - \frac{1}{2}(\overline{X} - \underline{X}), \frac{1}{2}(\underline{X} + \overline{X}) + \frac{1}{2}(\overline{X} - \underline{X}) \right] \\
&= [\underline{X}, \overline{X}].
\end{aligned}$$

Na definição a seguir, veremos que é possível trabalhar com n-uplas e matrizes onde cada entrada é um intervalo fechado e limitado da reta. No próximo capítulo será apresentado um Problema de Otimização Intervalar que estará definido no espaço vetorial intervalar n-dimensional.

**Definição 1.1.7.** *Um vetor intervalar n-dimensional é uma n-upla ordenada onde cada coordenada é um intervalo.*

$$X = (X_1, X_2, \dots, X_n) \quad (1.23)$$

onde cada  $X_i$  é da forma  $X_i = [\underline{X}_i, \overline{X}_i]$

Uma matriz intervalar  $A$  é uma matriz onde cada entrada  $a_{ij}$  é um intervalo, isto é,  $a_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$ .

### 1.1.2 Propriedades Algébricas

É fácil ver que a adição e multiplicação de intervalos são comutativas e associativas. Tem-se:

$$\begin{aligned}
X + Y &= Y + X, & X + (Y + Z) &= (X + Y) + Z \\
XY &= YX, & X(YZ) &= (XY)Z
\end{aligned}$$

para quaisquer intervalos  $X, Y$  e  $Z$ .

Os intervalos degenerados  $0 = [0, 0]$  e  $1 = [1, 1]$  os elementos neutros na adição e na multiplicação respectivamente.

$$\begin{aligned}
0 + X &= X + 0 = X \\
1 \cdot X &= X \cdot 1 = X
\end{aligned}$$

para todo intervalo  $X$ .

Perceba que  $-X$  não é o inverso aditivo para  $X$  no ambiente intervalar. De fato,

$$X + (-X) = [\underline{X}, \overline{X}] + [-\overline{X}, -\underline{X}] = [\underline{X} - \overline{X}, \overline{X} - \underline{X}],$$

e é igual a  $[0, 0]$  se, e somente se,  $\underline{X} = \overline{X}$ . Se  $X$  não tem medida nula, então

$$X - X = \omega(X)[-1, 1]. \quad (1.24)$$

Similarmente,  $X/X = 1$  somente se  $\omega(X) = 0$ . Em geral,

$$X/X = \begin{cases} [\underline{X}/\overline{X}, \overline{X}/\underline{X}], & \text{se } 0 < \underline{X} \\ [\overline{X}/\underline{X}, \underline{X}/\overline{X}], & \text{se } \overline{X} < 0 \end{cases}$$

Não temos inverso aditivo ou multiplicativo exceto para intervalos degenerados. Entretanto, sempre temos as inclusões  $0 \in X - X$  e  $1 \in X/X$ .

A lei da distributividade

$$x(y + z) = xy + xz$$

da aritmética real também não se mantém para o caso intervalar. Um contra exemplo pode ser obtido tomando  $X = [1, 2]$ ,  $Y = [1, 1]$  e  $Z = [-1, -1]$ :

$$\begin{aligned} X(Y + Z) &= [1, 2] \cdot ([1, 1] + [-1, -1]) \\ X(Y + Z) &= [1, 2] \cdot [0, 0] \\ X(Y + Z) &= [0, 0], \end{aligned}$$

porém,

$$\begin{aligned} XY + XZ &= [1, 2] \cdot [1, 1] + [1, 2] \cdot [-1, -1] \\ XY + XZ &= [1, 2] - [1, 2] \\ XY + XZ &= [-1, 1]. \end{aligned}$$

Contudo, existe a *lei da subdistributividade*:

$$X(Y + Z) \subseteq XY + XZ \quad (1.25)$$

A lei da distributividade é mantida em casos especiais. Em particular, para qualquer número real  $x$  temos

$$x(Y + Z) = xY + xZ. \quad (1.26)$$

A multiplicação intervalar pode ser distribuída sobre uma soma de intervalos, desde que esses intervalos tenham o mesmo sinal:

$$X(Y + Z) = XY + XZ \quad (1.27)$$

desde que  $\underline{Y}, \underline{Z} > 0$  ou  $\bar{Y}, \bar{Z} < 0$ .

A lei do corte para adição

$$X + Z = Y + Z \implies X = Y \quad (1.28)$$

é mantida na aritmética intervalar. Contudo,  $XZ = YZ$  não implica  $X = Y$ . Considere  $X = [-1, 1]$ ,  $Y = [-1, 0]$  e  $Z = [-1, 1]$ . Temos

$$X \cdot Z = [-1, 1] \cdot [-1, 1] = [-1, 1] \text{ e } Y \cdot Z = [-1, 0] \cdot [-1, 1] = [-1, 1],$$

ou seja,  $X \cdot Z = Y \cdot Z$ , porém  $X \neq Y$ .

## 1.2 Espaço Vetorial Intervalar Generalizado $n$ Dimensional

Nesta seção vamos apresentar uma extensão do conjunto dos intervalos estudados até agora e equipar esse novo conjunto com uma estrutura algébrica de espaço vetorial. Isso será feito através de uma bijeção entre esse novo espaço e o espaço vetorial  $\mathbb{R}^{2n}$ .

### 1.2.1 Bijeção entre $M^n$ e $\mathbb{R}^{2n}$

Denotaremos o conjunto dos intervalos fechados e limitados por

$$I(\mathbb{R}) = \{[a, b] : a \leq b, a, b \in \mathbb{R}\}$$

com a aritmética desenvolvida na seção anterior. Sabemos que esse espaço não é um espaço vetorial porque nem todo elemento possui inverso aditivo nesse espaço. Assim, alguns conceitos de cálculos intervalares não são extensões naturais dos conceitos clássicos.

Em 2014, Costa [3] usa um conjunto  $M \supset I(\mathbb{R})$  com novas operações definidas. Tal conjunto é dado por

$$M \equiv I(\mathbb{R}) \cup \overline{I(\mathbb{R})},$$

onde  $\overline{I(\mathbb{R})} = \{[a, b] : [b, a] \in I(\mathbb{R})\}$ . O conjunto  $M$  é chamado *espaço intervalar generalizado*.

Os elementos de  $I(\mathbb{R})$  são chamados *intervalos próprios* e os elementos de  $\overline{I(\mathbb{R})}$  são chamados de *intervalos impróprios*. Vamos considerar também o conjunto  $M^n = M \times M \times \dots \times M$  com  $n$  fatores, para  $n \in \mathbb{N}$ . O conjunto  $M^n$  dotado de uma estrutura algébrica de espaço vetorial é chamado *Espaço Vetorial Intervalar Generalizado  $n$  Dimensional*.

A necessidade da extensão do espaço intervalar se deve a alguns motivos. Dentre eles podemos destacar alguns, por exemplo, a não existência de inverso aditivo para um intervalo não degenerado. Outra caso que pode ocorrer é o seguinte, quando uma função intervalar  $F : \mathbb{R} \rightarrow I(\mathbb{R})$  é definida da forma  $F(x) = [\underline{F}(x), \overline{F}(x)]$  onde  $\underline{F}(x) \leq \overline{F}(x)$  para todo  $x \in \mathbb{R}^n$ , a função fica bem definida. Porém, quando se analisam as derivadas de  $F$ , a responsabilidade fica inteiramente sobre as funções  $\underline{F}$  e  $\overline{F}$ . O gradiente de  $F$  é definido como sendo  $\nabla F(x) = [\nabla \underline{F}(x), \nabla \overline{F}(x)]$ . Mas o fato de que  $\underline{F}(x) \leq \overline{F}(x)$  não implica em que  $\nabla \underline{F}(x) \leq \nabla \overline{F}(x)$ , isto é, o gradiente de  $F$  pode não estar bem definido.

Vamos construir uma bijeção entre  $M^n$  e  $\mathbb{R}^{2n}$  que será importante para o desenvolvimento desse trabalho.

**Exemplo 1.2.1.** *Seja a aplicação  $\psi : M \rightarrow \mathbb{R}^2$  definida por*

$$\psi([\underline{X}, \overline{X}]) = (\lambda_1 \underline{X} + \lambda_2 \overline{X}, \beta_1 \underline{X} + \beta_2 \overline{X}) \quad (1.29)$$

com  $\lambda_1, \lambda_2, \beta_1, \beta_2 \in \mathbb{R}$  são parâmetros fixos, porém arbitrários, tais que  $\lambda_1 \beta_2 - \lambda_2 \beta_1 \neq 0$ .

A aplicação  $\psi$  é uma bijeção e sua inversa  $\psi^{-1}$  é dada por

$$\psi^{-1}(a, b) = \left[ \frac{\beta_2 a - \lambda_2 b}{\lambda_1 \beta_2 - \lambda_2 \beta_1}, \frac{\lambda_1 b - \beta_1 a}{\lambda_1 \beta_2 - \lambda_2 \beta_1} \right] \quad (1.30)$$

Um caso particular dessa bijeção pode ser obtida quando se considera  $\lambda_1 = \beta_2 = 1$  e  $\lambda_2 = \beta_1 = 0$ , e obtemos

$$\psi_1([\underline{X}, \overline{X}]) = (\underline{X}, \overline{X}) \quad (1.31)$$

**Exemplo 1.2.2.** *Considere a bijeção  $\psi : M \rightarrow \mathbb{R}^2$  dada no Exemplo 1.2.1 e faça  $\varphi_i = \psi$  para todo  $i \in \{1, 2, \dots, n\}$ . Podemos definir uma bijeção  $\overline{\varphi} : M^n \rightarrow \mathbb{R}^{2n}$  como sendo*

$$\begin{aligned} \bar{\varphi}([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n]) = \\ ((\lambda_1 \underline{X}_1 + \lambda_2 \overline{X}_1, \beta_1 \underline{X}_1 + \beta_2 \overline{X}_1), \dots, (\lambda_1 \underline{X}_n + \lambda_2 \overline{X}_n, \beta_1 \underline{X}_n + \beta_2 \overline{X}_n)), \end{aligned}$$

com inversa dada por

$$\bar{\varphi}^{-1}((a_1, b_1), \dots, (a_n, b_n)) = (\psi^{-1}(a_1, b_1), \dots, \psi^{-1}(a_n, b_n))$$

Vamos agora definir operações em  $M^n$  induzidas por uma bijeção.

**Definição 1.2.1.** Dados o espaço vetorial usual  $(\mathbb{R}^{2n}, +, \cdot)$  e uma bijeção  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$ , denotaremos por  $(M^n, +_\varphi, \cdot_\varphi)$  o espaço com as operações  $+_\varphi : M^n \times M^n \rightarrow M^n$  e  $\cdot_\varphi : \mathbb{R} \times M^n \rightarrow M^n$  dadas por

$$\begin{aligned} &([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n]) +_\varphi ([\underline{Y}_1, \overline{Y}_1], \dots, [\underline{Y}_n, \overline{Y}_n]) \\ &:= \varphi^{-1}(\varphi([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n]) + \varphi([\underline{Y}_1, \overline{Y}_1], \dots, [\underline{Y}_n, \overline{Y}_n])) \end{aligned}$$

e

$$\begin{aligned} &\alpha \cdot_\varphi ([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n]) \\ &:= \varphi^{-1}(\alpha \cdot \varphi([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n])) \end{aligned}$$

**Teorema 1.2.1.** Seja  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$  uma bijeção. Então o espaço  $(M^n, +_\varphi, \cdot_\varphi)$  é um espaço vetorial.

*Demonstração.* A prova segue do fato de que  $(\mathbb{R}^{2n}, +, \cdot)$  é um espaço vetorial e da Definição 1.2.1 para  $+_\varphi$  e  $\cdot_\varphi$ .  $\square$

**Corolário 1.2.1.** O espaço vetorial  $(M^n, +_\varphi, \cdot_\varphi)$  é isomorfo a  $(\mathbb{R}^{2n}, +, \cdot)$ .

O seguinte resultado estabelece uma equivalência entre a operação de soma e multiplicação por escalar definida por Moore [7] e por Costa [3].

**Teorema 1.2.2.** Considere a bijeção dada no Exemplo 1.2.1. As operações definidas de acordo com a Definição 1.2.1 são equivalentes as operações dadas nas equações (1.13) e (1.18).

*Demonstração.* Sejam  $X = [\underline{X}, \overline{X}]$  e  $Y = [\underline{Y}, \overline{Y}]$ . Assim sendo, tem-se:

$$\begin{aligned}
X +_{\psi} Y &= \psi^{-1}(\psi(X) + \psi(Y)) \\
&= \psi^{-1}\left((\lambda_1 \underline{X} + \lambda_2 \overline{X}, \beta_1 \underline{X} + \beta_2 \overline{X}) + (\lambda_1 \underline{Y} + \lambda_2 \overline{Y}, \beta_1 \underline{Y} + \beta_2 \overline{Y})\right) \\
&= \psi^{-1}\left(\lambda_1 \underline{X} + \lambda_2 \overline{X} + \lambda_1 \underline{Y} + \lambda_2 \overline{Y}, \beta_1 \underline{X} + \beta_2 \overline{X} + \beta_1 \underline{Y} + \beta_2 \overline{Y}\right) \\
&= \psi^{-1}\left(\lambda_1(\underline{X} + \underline{Y}) + \lambda_2(\overline{X} + \overline{Y}), \beta_1(\underline{X} + \underline{Y}) + \beta_2(\overline{X} + \overline{Y})\right) \\
&= \left[ \frac{\beta_2(\lambda_1(\underline{X} + \underline{Y}) + \lambda_2(\overline{X} + \overline{Y})) - \lambda_2(\beta_1(\underline{X} + \underline{Y}) + \beta_2(\overline{X} + \overline{Y}))}{\lambda_1\beta_2 - \lambda_2\beta_1}, \right. \\
&\quad \left. \frac{\lambda_1(\beta_1(\underline{X} + \underline{Y}) + \beta_2(\overline{X} + \overline{Y})) - \beta_1(\lambda_1(\underline{X} + \underline{Y}) + \lambda_2(\overline{X} + \overline{Y}))}{\lambda_1\beta_2 - \lambda_2\beta_1} \right] \\
&= \left[ \frac{(\underline{X} + \underline{Y})(\lambda_1\beta_2 - \lambda_2\beta_1)}{\lambda_1\beta_2 - \lambda_2\beta_1}, \frac{(\overline{X} + \overline{Y})(\lambda_1\beta_2 - \lambda_2\beta_1)}{\lambda_1\beta_2 - \lambda_2\beta_1} \right] \\
&= [\underline{X} + \underline{Y}, \overline{X} + \overline{Y}]
\end{aligned}$$

Sem perda de generalidade, considere  $\alpha \geq 0$ . Logo

$$\begin{aligned}
\alpha \cdot_{\psi} X &= \psi^{-1}(\alpha \cdot \psi(X)) \\
&= \psi^{-1}\left(\alpha \cdot (\lambda_1 \underline{X} + \lambda_2 \overline{X}, \beta_1 \underline{X} + \beta_2 \overline{X})\right) \\
&= \psi^{-1}\left(\alpha\lambda_1 \underline{X} + \alpha\lambda_2 \overline{X}, \alpha\beta_1 \underline{X} + \alpha\beta_2 \overline{X}\right) \\
&= \left[ \frac{\beta_2(\alpha\lambda_1 \underline{X} + \alpha\lambda_2 \overline{X}) - \lambda_2(\alpha\beta_1 \underline{X} + \alpha\beta_2 \overline{X})}{\lambda_1\beta_2 - \lambda_2\beta_1}, \right. \\
&= \left. \frac{\lambda_1(\alpha\beta_1 \underline{X} + \alpha\beta_2 \overline{X}) - \beta_1(\alpha\lambda_1 \underline{X} + \alpha\lambda_2 \overline{X})}{\lambda_1\beta_2 - \lambda_2\beta_1} \right] \\
&= \left[ \frac{\alpha(\lambda_1\beta_2 - \lambda_2\beta_1)\underline{X}}{\lambda_1\beta_2 - \lambda_2\beta_1}, \frac{\alpha(\lambda_1\beta_2 - \lambda_2\beta_1)\overline{X}}{\lambda_1\beta_2 - \lambda_2\beta_1} \right] \\
&= [\alpha \cdot \underline{X}, \alpha \cdot \overline{X}]
\end{aligned}$$

□

### 1.3 Topologia em $M^n$

Como consequência dos resultados obtidos até agora, vamos mostrar que  $M^n$  é um espaço com produto interno e induzir norma, métrica e consequentemente uma topologia em  $M^n$

### 1.3.1 Métricas em $M^n$

**Teorema 1.3.1.** *Seja  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$  uma bijeção linear e  $\langle \cdot, \cdot \rangle$  um produto interno em  $\mathbb{R}^{2n}$ . Considere a seguinte aplicação:*

$$\begin{aligned} (\cdot, \cdot)_\varphi &: M^n \times M^n \rightarrow \mathbb{R} \\ (X, Y)_\varphi &= \langle \varphi(X), \varphi(Y) \rangle. \end{aligned}$$

*Essa aplicação é um produto interno em  $(M^n, +_\varphi, \cdot_\varphi)$ .*

*Demonstração.* Sejam  $X, Y, Z \in M^n$  e  $\alpha \in \mathbb{R}$ . Então:

- i)  $(X, X)_\varphi = \langle \varphi(X), \varphi(X) \rangle = \|\varphi(X)\|^2 \geq 0$  e  $(X, X)_\varphi = 0 \Leftrightarrow X = 0$ .
- ii)  $(X, Y)_\varphi = \langle \varphi(X), \varphi(Y) \rangle = \langle \varphi(Y), \varphi(X) \rangle = (Y, X)_\varphi$
- iii)  $(\alpha \cdot_\varphi X +_\varphi Y, Z)_\varphi = \langle \varphi(\alpha \cdot_\varphi X +_\varphi Y), \varphi(Z) \rangle =$   
 $\alpha \langle \varphi(X), \varphi(Z) \rangle + \langle \varphi(Y), \varphi(Z) \rangle = \alpha(X, Z)_\varphi + (Y, Z)_\varphi$

□

**Corolário 1.3.1.** *O espaço  $(M^n, +_\varphi, \cdot_\varphi, \|\cdot\|_\varphi)$  onde  $\|X\|_\varphi = (X, X)_\varphi^{\frac{1}{2}}$  é um espaço normado.*

**Corolário 1.3.2.** *O espaço  $(M^n, d_\varphi)$  onde  $d_\varphi : M^n \times M^n \rightarrow \mathbb{R}$  definida por*

$$d_\varphi(X, Y) = \|\varphi(X) - \varphi(Y)\|_\varphi,$$

*é um espaço métrico. Mais geralmente, dada uma métrica  $\bar{d}$  em  $\mathbb{R}^{2n}$ , podemos definir uma métrica  $\bar{d}_M$  em  $M^n$  definida por  $d_M(X, Y) = \bar{d}(\varphi(X), \varphi(Y))$*

Sejam  $A, B \subseteq \mathbb{R}^n$ . A métrica de Hausdorff entre  $A$  e  $B$  é definida como sendo

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

onde  $\|\cdot\|$  é a norma Euclidiana em  $\mathbb{R}^n$ . Sejam  $X, Y$  intervalos, isto é,  $X = [\underline{X}, \bar{X}]$  e  $Y = [\underline{Y}, \bar{Y}]$ , segundo Wu [9] tem-se que

$$d_H(X, Y) = \max\{|\underline{X} - \underline{Y}|, |\bar{X} - \bar{Y}|\}$$

Agora que o espaço  $M$  está munido de métrica, pode-se trabalhar o conceito de sequência em  $M$ . Uma sequência em  $M$ , é uma aplicação  $x : \mathbb{N} \rightarrow M$ .

Sejam  $\{X^k\}$  uma sequência em  $M$  e  $d$  uma métrica em  $M$ . Dizemos que  $\{X^k\}$  é convergente, se existe  $X^* \in M$  tal que para todo  $\varepsilon > 0$ , existe  $n_0 \in \mathbb{N}$  tal que  $d(X^k, X^*) < \varepsilon$  sempre que  $k > n_0$ . Quando isso acontecer, denotaremos como

$$X^* = \lim_{k \rightarrow \infty} X^k$$

ou simplesmente  $X^k \rightarrow X^*$

**Proposição 1.3.1.** *Considere a métrica de Hausdorff em  $M$ . Seja  $\{X^k\}$  uma sequência em  $M$  convergente para  $X^* \in M$ . Então  $\{\underline{X^k}\} \rightarrow \underline{X^*}$  e  $\{\overline{X^k}\} \rightarrow \overline{X^*}$ .*

*Demonstração.* De fato, temos que  $d_H(X^k, X^*) < \varepsilon$  implica em  $|\underline{X^k} - \underline{X^*}| < \varepsilon$  e  $|\overline{X^k} - \overline{X^*}| < \varepsilon$ . □

## 1.4 Relação de Ordem

Vamos começar definindo uma relação de ordem parcial em  $M^n$ . Isso será feito inicialmente com auxílio de uma bijeção entre  $M^n$  e  $\mathbb{R}^{2n}$ .

**Definição 1.4.1.** *Uma relação binária  $R$  entre dois conjuntos  $U$  e  $V$  é um subconjunto do produto cartesiano  $U \times V$ . Se  $U = V$ , podemos dizer simplesmente que  $R$  é uma relação binária em  $U$ . Se  $u \in U$  está relacionado com  $v \in V$ , dizemos que  $(u, v) \in R$ . Uma relação binária  $A$  definida em  $\mathbb{R}^{2n}$  é chamada **relação de ordem parcial não estrita** em  $\mathbb{R}^{2n}$  se possui as seguintes propriedades:*

- *Reflexiva, se  $(x, x) \in A$  para todo  $x \in \mathbb{R}^{2n}$ .*
- *Anti-Simétrica, se  $(x, y) \in A$  e  $(y, x) \in A$  implica que  $x = y$  para  $x, y \in \mathbb{R}^{2n}$ .*
- *Transitiva, se  $(x, y) \in A$  e  $(y, z) \in A$  implica que  $(x, z) \in A$  para  $x, y, z \in \mathbb{R}^{2n}$ .*

*Uma relação binária  $R$  em  $U$  tem a propriedade da dicotomia, se dados  $u, v \in U$  então  $(u, v) \in R$  ou  $(v, u) \in R$ .*

*Se uma relação binária é reflexiva, anti-simétrica, transitiva e além disso, se essa relação binária também satisfaz a propriedade da dicotomia, então é uma relação de ordem total não estrita em  $\mathbb{R}^{2n}$ .*



Uma relação binária  $A$  definida em  $\mathbb{R}^{2n}$  é chamada **relação de ordem parcial estrita** em  $\mathbb{R}^{2n}$  se possui as seguintes propriedades:

- Anti reflexiva, se  $(x, x) \notin A$  para todo  $x \in \mathbb{R}^{2n}$ .
- Assimétrica, se  $(x, y) \in A$  não implica em  $(y, x) \in A$  para todo  $x, y \in \mathbb{R}^{2n}$ .
- Transitiva, se  $(x, y) \in A$  e  $(y, z) \in A$  implica que  $(x, z) \in A$  para  $x, y, z \in \mathbb{R}^{2n}$ .

Além disso, se essa relação binária também satisfaz a propriedade da tricotomia, então é uma relação de ordem total estrita em  $\mathbb{R}^{2n}$ .

**Exemplo 1.4.1.** Vamos denotar por  $\leq_{\mathbb{R}^{2n}}$  uma relação de ordem parcial em  $\mathbb{R}^{2n}$ . Por exemplo, a ordem parcial usual em  $\mathbb{R}^{2n}$  definida por

$$x \leq_{\mathbb{R}^{2n}} y \text{ se, e somente se, } x_i \leq y_i \text{ para todo } i = 1, 2, \dots, 2n$$

onde  $x = (x_1, x_2, \dots, x_{2n}), y = (y_1, y_2, \dots, y_{2n}) \in \mathbb{R}^{2n}$ .

Denotaremos por  $\prec_{\mathbb{R}^{2n}}$ , a relação binária em  $\mathbb{R}^{2n}$ , tal que, dados  $u = (u_1, u_2, \dots, u_{2n})$  e  $v = (v_1, v_2, \dots, v_{2n})$  em  $\mathbb{R}^{2n}$ , tem-se

$$u \prec_{\mathbb{R}^{2n}} v \Leftrightarrow u \leq_{\mathbb{R}^{2n}} v \text{ e } u \neq v.$$

Denotaremos por  $<_{\mathbb{R}^{2n}}$ , a relação binária em  $\mathbb{R}^{2n}$ , tal que, dados  $u = (u_1, u_2, \dots, u_{2n})$  e  $v = (v_1, v_2, \dots, v_{2n})$  em  $\mathbb{R}^{2n}$ , tem-se

$$u <_{\mathbb{R}^{2n}} v \Leftrightarrow u \leq_{\mathbb{R}^{2n}} v \text{ e } u_i < v_i \text{ para todo } i \in \{1, 2, \dots, 2n\}.$$

**Proposição 1.4.1.** Seja  $\leq_{\mathbb{R}^{2n}}$  uma relação de ordem parcial (total) em  $\mathbb{R}^{2n}$ . Então

- $\prec_{\mathbb{R}^{2n}}$  também é uma relação de ordem parcial (total) estrita em  $\mathbb{R}^{2n}$ .
- $<_{\mathbb{R}^{2n}}$  também é uma relação de ordem parcial estrita em  $\mathbb{R}^{2n}$ .

*Demonstração.* Ver Costa [2]. □

**Definição 1.4.2.** Dada uma relação de ordem parcial (total)  $\leq_{\mathbb{R}^{2n}}$  em  $\mathbb{R}^{2n}$  e uma bijeção  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$ , denotaremos por  $\leq_{\varphi}$ ,  $\prec_{\varphi}$  e  $<_{\varphi}$  as relações binárias definida em  $M^n$  como sendo

$$\begin{aligned}
A \leq_{\varphi} B &\Leftrightarrow \varphi(A) \leq_{\mathbb{R}^{2n}} \varphi(B); \\
A \prec_{\varphi} B &\Leftrightarrow \varphi(A) \prec_{\mathbb{R}^{2n}} \varphi(B); \\
A <_{\varphi} B &\Leftrightarrow \varphi(A) <_{\mathbb{R}^{2n}} \varphi(B);
\end{aligned}$$

onde  $A, B \in M^n$ .

**Proposição 1.4.2.** *Seja  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$  uma bijeção. Dada uma relação de ordem parcial (total)  $\leq_{\mathbb{R}^{2n}}$  em  $\mathbb{R}^{2n}$ , as relações binárias  $\leq_{\varphi}$  e  $\prec_{\varphi}$  dadas na Definição 1.4.2, são relações de ordem parcial (total) em  $M^n$ . Além disso,  $<_{\varphi}$  é uma relação de ordem parcial em  $M^n$ .*

*Demonstração.* A prova segue diretamente do fato de que  $\leq_{\mathbb{R}^{2n}}$  é uma relação de ordem parcial (total) em  $\mathbb{R}^{2n}$ , da Definição 1.4.1 e da Definição 1.4.2.  $\square$

Note que para cada bijeção  $\varphi : M^n \rightarrow \mathbb{R}^{2n}$  temos diferentes relações  $\leq_{\varphi}$  em  $M^n$ .

**Exemplo 1.4.2.** *Vamos considerar a bijeção  $\psi_1 : M^n \rightarrow \mathbb{R}^2$  de acordo com a Definição 1.2.1 e a relação de ordem usual  $\leq_{\mathbb{R}^{2n}}$  em  $\mathbb{R}^{2n}$ . Então temos a relação de ordem  $\leq_{\psi_1}$  definida por*

$$\begin{aligned}
[\underline{X}, \overline{X}] \leq_{\psi_1} [\underline{Y}, \overline{Y}] &\Leftrightarrow \psi_1([\underline{X}, \overline{X}]) \leq_{\mathbb{R}^{2n}} \psi_1([\underline{Y}, \overline{Y}]) \\
&\Leftrightarrow (\underline{X}, \overline{X}) \leq_{\mathbb{R}^{2n}} (\underline{Y}, \overline{Y}) \\
&\Leftrightarrow \underline{X} \leq \underline{Y} \text{ e } \overline{X} \leq \overline{Y}.
\end{aligned}$$

A relação de ordem  $\leq_{\psi_1}$  compara dois intervalos próprios, isto é, é uma relação de ordem em  $I(\mathbb{R})$ . Essa relação compara as extremidades respectivas de dois intervalos.

**Exemplo 1.4.3.** *Se considerarmos a bijeção  $\overline{\varphi}$  definida no Exemplo 1.2.2 e a relação de ordem usual  $\leq_{\mathbb{R}^{2n}}$  em  $\mathbb{R}^{2n}$ , temos a relação de ordem parcial  $\leq_{\overline{\varphi}}$  em  $M^n$  definida por*

$$\begin{aligned}
([\underline{X}_1, \overline{X}_1], \dots, [\underline{X}_n, \overline{X}_n]) &\leq_{\overline{\varphi}} ([\underline{Y}_1, \overline{Y}_1], \dots, [\underline{Y}_n, \overline{Y}_n]) \\
\Leftrightarrow (\psi([\underline{X}_1, \overline{X}_1]), \dots, \psi([\underline{X}_n, \overline{X}_n])) &\leq_{\mathbb{R}^{2n}} (\psi([\underline{Y}_1, \overline{Y}_1]), \dots, \psi([\underline{Y}_n, \overline{Y}_n]))
\end{aligned}$$

**Exemplo 1.4.4.** *Existe uma relação de ordem total, chamada ordem lexicográfica em  $\mathbb{R}^2$  é dada por*

$$(x_1, x_2) \prec_{\mathbb{R}^2} (y_1, y_2) \Leftrightarrow x_1 < y_1 \text{ ou } x_1 = y_1 \text{ e } x_2 \leq y_2$$

Essa pode ser generalizada para  $\mathbb{R}^{2n}$  da seguinte forma

$$x \prec_{\mathbb{R}^{2n}} y \Leftrightarrow x_1 < y_1 \text{ ou } \exists i \in \{1, 2, \dots, 2n-1\} \text{ tal que } x_1 = y_1, \dots, x_i = y_i \text{ e } x_{i+1} \leq y_{i+1}$$

para todo  $x = (x_1, x_2, \dots, x_{2n}), y = (y_1, y_2, \dots, y_{2n}) \in \mathbb{R}^{2n}$ .

Alguns autores propuseram relações de ordem parcial em  $I(\mathbb{R})$  (Moore [5], Ishibuchi [4]), dadas das seguintes maneiras:

$$A \preceq_{LR} B \text{ se } \underline{A} \leq \underline{B} \text{ e } \overline{A} \leq \overline{B}; A \prec_{LR} B \text{ se } A \preceq_{LR} B \text{ e } A \neq B.$$

$$A \preceq_{LC} B \text{ se } \underline{A} \leq \underline{B} \text{ e } m(A) \leq m(B); A \prec_{LC} B \text{ se } A \preceq_{LC} B \text{ e } A \neq B.$$

$$A \preceq_{RC} B \text{ se } m(A) \leq m(B) \text{ e } \overline{A} \leq \overline{B}; A \prec_{RC} B \text{ se } A \preceq_{RC} B \text{ e } A \neq B.$$

Em 2012, Panda [1] propõe uma outra forma de se estabelecer relação de ordem em ambiente intervalar. Para tal, se considera um intervalo  $A = [\underline{A}, \overline{A}]$  na forma paramétrica, tal que  $A$  pode ser expresso como  $A(t)$ , onde  $A(t) = \underline{A} + t(\overline{A} - \underline{A}) = (1-t)\underline{A} + t\overline{A}$ . Sob essa abordagem, as operações algébricas também são definidas em termos das extremidades do intervalo. Sejam  $A$  e  $B$  dois intervalos, as operações ficam determinadas como sendo

$$A \circledast B = \{a(t_1) \circledast b(t_2) | t_1, t_2 \in [0, 1]\}$$

onde  $\circledast = +, -, \cdot$  ou  $/$  (no caso em que  $\circledast = /$ , necessariamente  $b(t_2) \neq 0$ ).

**Definição 1.4.3.** Para  $A, B \in M$ ,

$$i) A \preceq B \text{ se } a(t_1) \leq b(t_2), t_1, t_2 \in [0, 1];$$

$$ii) A \preceq_{\omega} B \text{ se } a(t) \leq b(t), t \in [0, 1].$$

Além disso,  $A \prec B$  se  $a(t_1) < b(t_2)$ ,  $t_1, t_2 \in [0, 1]$  e  $A \prec_{\omega} B$  se  $a(t) < b(t)$ ,  $t \in [0, 1]$ .

Note o seguinte:

$$i) \text{ Para } (t_1, t_2) = (0, 0) \text{ e } (1, 1), \preceq \text{ torna-se } \preceq_{LR};$$

$$ii) \text{ Para } (t_1, t_2) = (0, 0) \text{ e } (\frac{1}{2}, \frac{1}{2}), \preceq \text{ torna-se } \preceq_{LC};$$

- iii) Para  $(t_1, t_2) = (1, 1)$  e  $(\frac{1}{2}, \frac{1}{2})$ ,  $\preceq$  torna-se  $\preceq_{RC}$ ;
- iv) Para  $t = 0$  e  $t = 1$ ,  $\preceq_\omega$  torna-se  $\preceq_{LR}$ ;
- v) Para  $t = 0$  e  $t = \frac{1}{2}$ ,  $\preceq_\omega$  torna-se  $\preceq_{LC}$ ;
- vi) Para  $t = 1$  e  $t = \frac{1}{2}$ ,  $\preceq_\omega$  torna-se  $\preceq_{RC}$ ;

Segundo Panda [1], a relação  $\preceq_\omega$  é um caso particular de  $\preceq$ . De fato,  $A \preceq B$  sempre implica  $A \preceq_\omega B$ , mas a recíproca não é verdadeira. Considere os intervalos  $A = [1, 3]$  e  $B = [2, 4]$ . Na forma paramétrica temos  $A(t) = 1 + 2t$  e  $B(t) = 2 + 2t$  onde  $t \in [0, 1]$ . Logo,  $1 + 2t \leq 2 + 2t$  o que implica em  $A(t) \leq B(t)$  para todo  $t \in [0, 1]$ , isto é,  $A \preceq_\omega B$ . No entanto, não é verdade que  $1 + 2t_1 \leq 2 + t_2$  para todo  $t_1, t_2 \in [0, 1]$ , para verificar isso basta considerar  $t_1 = 1$  e  $t_2 = 0$ . Teremos que  $A(1) = 3$  e  $B(0) = 2$ .

**Proposição 1.4.3.** *As relações de ordem  $\preceq_\omega$  e  $\leq_{\psi_1}$  são equivalentes.*

*Demonstração.* Sejam  $A, B \in M$ . Se  $\preceq_\omega B$ , então

$$A(t) \leq B(t), \text{ para todo } t \in [0, 1].$$

ou seja,

$$A(0) \leq B(0) \text{ que implica } \underline{A} \leq \underline{B} \text{ e,}$$

$$A(1) \leq B(1) \text{ que implica } \overline{A} \leq \overline{B},$$

que implica em  $A \leq_{\psi_1} B$  de acordo com o Exemplo 1.4.2.

Suponha agora que  $A \leq_{\psi_1} B$ . Assim sendo,  $\underline{A} \leq \underline{B}$  e  $\overline{A} \leq \overline{B}$ . Para  $t \in [0, 1]$  temos

$$(1-t)\underline{A} \leq (1-t)\underline{B} \text{ e}$$

$$t\overline{A} \leq t\overline{B}$$

que implica em  $(1-t)\underline{A} + t\overline{A} \leq (1-t)\underline{B} + t\overline{B}$ , ou seja,  $A(t) \leq B(t)$  para  $t \in [0, 1]$ . □

## 2 Elementos de Otimização Intervalar

Geralmente é difícil, do ponto de vista prático, determinar os coeficientes de uma função objetivo como um número real, pois, na maioria das vezes, estes possuem incertezas. Para tal, consideramos a otimização intervalar como uma abordagem para lidar com as incertezas dos coeficientes da função objetivo para modelos de programação matemática.

### 2.1 Funções intervalares

Nessa seção apresentamos alguns resultados relacionados a funções que assumem valores intervalares.

Costa [3] define funções intervalares  $F : \mathbb{R}^n \rightarrow M$  com valor  $F(x) = [\underline{F}(x), \overline{F}(x)]$ , onde  $\underline{F}, \overline{F} : \mathbb{R}^n \rightarrow \mathbb{R}$  e  $\underline{F}(x) \leq \overline{F}(x)$  para todo  $x$ , assim como foi proposto por Moore [5]. No entanto, Panda [1] representa funções intervalares de outra maneira.

Considere uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , onde os coeficientes de  $f$  são números reais. Seja  $c$  o conjunto de todos os parâmetros presentes em  $f$ . Sem perda de generalidade podemos considerar que  $c$  é um conjunto ordenado com respeito a ordem em que os parâmetros estão dispostos em  $f(x)$ . Por exemplo, se  $f(x_1, x_2) = 3x_1^2 + 2x_1e^{4x_2}$ , então  $c = (3, 2, 4)$ . Em geral para  $c = (c_1, c_2, c_3)$ , denotaremos  $f_c(x_1, x_2) = c_1x_1^2 + c_2x_1e^{c_3x_2}$ . Suponha agora que  $c_1, c_2$  e  $c_3$  variam nos intervalos  $C_1, C_2$  e  $C_3$ , respectivamente,  $C_j = [\underline{C}_j, \overline{C}_j]$ , que sob a forma paramétrica fica  $C_j(t_j) = (1 - t_j)\underline{C}_j + t_j\overline{C}_j$ ,  $0 \leq t_j \leq 1$ ,  $j = 1, 2, 3$ . Logo  $C(t) = (c_1(t_1), c_2(t_2), c_3(t_3))$ . Denotaremos por  $C_v^k \in M^k$  como sendo

$$C_v^k = \{c(t) | c(t) = (c_1(t_1), c_2(t_2), \dots, c_k(t_k)), c_j(t_j) = (1 - t_j)\underline{c}_j + t_j\overline{c}_j, \\ t = (t_1, t_2, \dots, t_k), 0 \leq t_j \leq 1, j = 1, 2, \dots, k\}.$$

Tem-se que  $f_{c(t)}(x_1, x_2) = c_1(t_1)x_1^2 + c_2(t_2)x_1e^{c_3(t_3)x_2^2}$ . Cada  $c_j(t_j)$  é uma função linear. Como  $f_{c(t)}$  é contínua em  $t$  para todo  $x$  e  $t$  varia num conjunto compacto, dado um intervalo vetorial  $C_v^3$  o conjunto  $\{f_{c(t)}(x_1, x_2) | f_{c(t)} : \mathbb{R}^2 \rightarrow \mathbb{R}, c(t) \in C_v^3\}$  é um intervalo,

$$\left[ \min_t f_{c(t)}(x_1, x_2), \max_t f_{c(t)}(x_1, x_2) \right].$$

Vamos agora generalizar esse conceito. Dado  $C_v^k \in M^k$ , seja  $f_{c(t)} : \mathbb{R}^n \rightarrow \mathbb{R}$ . Para todo  $c(t)$ ,  $f_{c(t)}$  é uma função de  $x$ . Suponha que para todo  $x$ ,  $f_{c(t)}$  é contínua em  $t$ . Então para um dado intervalo vetorial  $C_v^k$ , definimos uma função intervalar  $F_{C_v^k}(x) : \mathbb{R}^n \rightarrow M$  por

$$F_{C_v^k}(x) = \{f_{c(t)}(x) | f_{c(t)} : \mathbb{R}^n \rightarrow \mathbb{R}, c(t) \in C_v^k\} \quad (2.1)$$

Como para cada  $x$  fixado,  $f_{c(t)}(x)$  é contínua em  $t$  então,  $\min_{c(t) \in C_v^k} f_{c(t)}(x)$ , que é  $\min_{t \in [0,1]^k} f_{c(t)}(x)$  e  $\max_{c(t) \in C_v^k} f_{c(t)}(x)$  que é  $\max_{t \in [0,1]^k} f_{c(t)}(x)$  existem pois  $[0, 1]^k$  é um conjunto compacto. Nesse caso

$$F_{C_v^k}(x) = \left[ \min_{t \in [0,1]^k} f_{c(t)}(x), \max_{t \in [0,1]^k} f_{c(t)}(x) \right]. \quad (2.2)$$

Para  $x, y \in \mathbb{R}^n$ , definimos operações algébricas de funções intervalares como segue.

**Definição 2.1.1.** Para  $C_v^k, D_v^k \in M^k$  e  $\otimes \in \{+, -, \cdot, /\}$ ,

- i)  $(F_{C_v^k} \otimes F_{C_v^k})(y) = \{f_{c(t)}(x) \otimes f_{c(t)}(y) | f_{c(t)} : \mathbb{R}^n \rightarrow \mathbb{R}, c(t) \in C_v^k\}$
- ii)  $(F_{C_v^k} \otimes F_{D_v^k})(x) = \{f_{c(t')}(x) \otimes f_{d(t'')}(y) | f_{c(t')}, f_{d(t'')} : \mathbb{R}^n \rightarrow \mathbb{R}, c(t') \in C_v^k, d(t'') \in D_v^k\}$

## 2.2 Cálculo em ambiente intervalar

Funções intervalares podem ser definidas de diferentes formas, e portanto o cálculo nesse ambiente pode ser desenvolvido. Na forma clássica uma função intervalar  $F$  pode ser expressada como sendo  $F(x) = [\underline{F}(x), \overline{F}(x)]$ . Então a existência de derivada de uma função intervalar  $F$  depende da existência da derivada das funções  $\underline{F}$  e  $\overline{F}$ . No entanto, Panda [1] faz uma outra abordagem para o cálculo em ambiente intervalar. Seja  $F_{C_v^k} : \mathbb{R}^n \rightarrow M$ . A existência da derivada de  $F_{C_v^k}$  depende da existência da derivada de  $f_{c(t)}$  para todo  $t$ .

**Teorema 2.2.1.** Suponha que  $\lim_{x \rightarrow x^*} f_{c(t)}(x)$  existe para todo  $t$  e é igual a  $a(t)$ . Sejam

$$a(t_*) = \min_{t \in [0,1]^k} \lim_{x \rightarrow x^*} f_{c(t)}(x) \text{ e } a(t^*) = \max_{t \in [0,1]^k} \lim_{x \rightarrow x^*} f_{c(t)}(x)$$

para alguns  $t_*, t^* \in [0, 1]^k$ . Então

$$\lim_{x \rightarrow x^*} F_{C_v^k}(x) = A$$

onde  $A = [a(t_*), a(t^*)]$ .

*Demonstração.* Desde que  $\lim_{x \rightarrow x^*} f_{c(t)}(x)$  exista para todo  $t$ , então dado  $\varepsilon > 0$ , existe  $\delta > 0$  tal que  $d(F_{C_v^k}(x), A) < \varepsilon$  sempre que  $\|x - x^*\| < \delta$ , onde  $d$  é uma métrica em  $M$  (por exemplo a métrica de Hausdorff  $d_H$ ). Então o limite de uma função intervalar  $F_{C_v^k}$  existe em um ponto  $x^*$  se o limite de  $f_{c(t)}$  existe para todo  $t$ .  $\square$

**Definição 2.2.1.** Uma função  $F_{C_v^k}$  é contínua em  $x^*$  se  $f_{c(t)}$  é contínua em  $x^*$  para todo  $t$ . Escrevemos

$$\lim_{x \rightarrow x^*} F_{C_v^k}(x) = F_{C_v^k}(x^*) = \left[ \min_t \lim_{x \rightarrow x^*} f_{c(t)}(x), \max_t \lim_{x \rightarrow x^*} f_{c(t)}(x) \right]$$

Para cada  $t \in [0, 1]^k$ , suponha que  $f_{c(t)}$  é diferenciável com respeito a  $x$ . Consequentemente

$$\lim_{h \rightarrow 0} \frac{f_{c(t)}(x+h) - f_{c(t)}(x)}{\|h\|}$$

existe. De acordo com a Definição 2.1.1 e o Teorema 2.2.1, temos

$$F_{C_v^k}(x+h) \ominus F_{C_v^k}(x) = \left\{ f_{c(t)}(x+h) - f_{c(t)}(x) \mid t \in [0, 1]^k \right\}$$

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{F_{C_v^k}(x+h) \ominus F_{C_v^k}(x)}{\|h\|} &= \left\{ \lim_{h \rightarrow 0} \frac{f_{c(t)}(x+h) - f_{c(t)}(x)}{\|h\|} \mid t \in [0, 1]^k \right\} \\ &= \left[ \min_{t \in [0, 1]^k} \lim_{h \rightarrow 0} \frac{f_{c(t)}(x+h) - f_{c(t)}(x)}{\|h\|}, \right. \\ &\quad \left. \max_{t \in [0, 1]^k} \lim_{h \rightarrow 0} \frac{f_{c(t)}(x+h) - f_{c(t)}(x)}{\|h\|} \right] \end{aligned}$$

A função intervalar  $F_{C_v^k}$  é diferenciável em  $x = x^*$  se  $f_{c(t)}$  é diferenciável em  $x = x^*$  para todo  $t \in [0, 1]^k$ . As derivadas parciais de  $F_{C_v^k}$  em  $x^*$  são definidas como segue:

$$\frac{\partial F_{C_v^k}(x^*)}{\partial x_i} = \left\{ \frac{\partial f_{c(t)}(x^*)}{\partial x_i} \mid t \in [0, 1]^k \right\}$$

Se  $\varphi_i : [0, 1]^k \rightarrow \mathbb{R}$  dada por  $\varphi_i(t) = \frac{\partial f_{c(t)}(x^*)}{\partial x_i}$  é contínua em  $t$  para  $i = 1, 2, \dots, n$ , então

$$\frac{\partial F_{C_v^k}(x^*)}{\partial x_i} = \left[ \min_{t \in [0, 1]^k} \frac{\partial f_{c(t)}(x^*)}{\partial x_i}, \max_{t \in [0, 1]^k} \frac{\partial f_{c(t)}(x^*)}{\partial x_i} \right]$$

O gradiente de uma função intervalar  $F_{C_v^k} : \mathbb{R}^n \rightarrow M$  em um ponto  $x^*$  é o vetor intervalar

$$\nabla F_{C_v^k}(x^*) = \left( \frac{\partial F_{C_v^k}(x^*)}{\partial x_1}, \frac{\partial F_{C_v^k}(x^*)}{\partial x_2}, \dots, \frac{\partial F_{C_v^k}(x^*)}{\partial x_n} \right)$$

De forma similiar a derivada parcial de segunda ordem de  $F_{C_v^k}$  em  $x = x^*$ , pode ser calculada como

$$\frac{\partial^2 F_{C_v^k}(x^*)}{\partial x_i \partial x_j} = \left\{ \frac{\partial^2 f_{c(t)}(x^*)}{\partial x_i \partial x_j} \middle| t \in [0, 1]^k \right\}$$

Se  $\varphi_{i,j} : [0, 1]^k \rightarrow \mathbb{R}$  dada por  $\varphi_{i,j}(t) = \frac{\partial^2 f_{c(t)}(x^*)}{\partial x_i \partial x_j}$  é contínua em  $t$  para  $i, j = 1, 2, \dots, n$ , então

$$\frac{\partial^2 F_{C_v^k}(x^*)}{\partial x_i \partial x_j} = \left[ \min_{t \in [0, 1]^k} \frac{\partial^2 f_{c(t)}(x^*)}{\partial x_i \partial x_j}, \max_{t \in [0, 1]^k} \frac{\partial^2 f_{c(t)}(x^*)}{\partial x_i \partial x_j} \right]$$

Seja  $\nabla^2 F_{C_v^k}$  a matriz da Hessiana de  $F_{C_v^k}$  em  $x = x^*$ , que é uma matriz quadrada cujas entradas são  $\frac{\partial^2 F_{C_v^k}(x^*)}{\partial x_i \partial x_j}$ . Consequentemente

$$\nabla^2 F_{C_v^k} = \left\{ \nabla^2 f_{c(t)}(x^*) \middle| t \in [0, 1]^k \right\} \quad (2.3)$$

**Exemplo 2.2.1.** Considere a função  $F : \mathbb{R}^{2+} \rightarrow I(\mathbb{R})$  dada por  $F(x_1, x_2) = [3, 4]x_1^2 + [4, 8]x_1x_2 + [4, 5]x_2^2$ . Sob a forma paramétrica temos  $f_{c(t)} : \mathbb{R}^{2+} \rightarrow \mathbb{R}$  dada por  $f_{c(t)}(x_1, x_2) = (3 + t_1)x_1^2 + (4 + 4t_2)x_1x_2 + (4 + t_3)x_2^2$ , onde  $t = (t_1, t_2, t_3) \in [0, 1]^3$ .

Vamos calcular o gradiente e a hessiana de  $F$ . Para cada  $t \in [0, 1]^3$ , temos

$$\frac{\partial f_{c(t)}}{\partial x_1}(x_1, x_2) = (6 + 2t_1)x_1 + (4 + 4t_2)x_2$$

e



$$\frac{\partial f_{c(t)}}{\partial x_2}(x_1, x_2) = (4 + 4t_2)x_1 + (8 + 2t_3)x_2$$

onde  $t = (t_1, t_2, t_3) \in [0, 1]^3$ .

Assim sendo,

$$\begin{aligned} \frac{\partial F}{\partial x_1}(x_1, x_2) &= \left[ \min_{t \in [0,1]^3} \frac{\partial f_{c(t)}}{\partial x_1}(x_1, x_2), \max_{t \in [0,1]^3} \frac{\partial f_{c(t)}}{\partial x_1}(x_1, x_2) \right] \\ &= \left[ \min_{t \in [0,1]^3} (6 + 2t_1)x_1 + (4 + 4t_2)x_2, \max_{t \in [0,1]^3} (6 + 2t_1)x_1 + (4 + 4t_2)x_2 \right] \\ &= [6x_1 + 4x_2, 8x_1 + 8x_2] \end{aligned}$$

e

$$\begin{aligned} \frac{\partial F}{\partial x_2}(x_1, x_2) &= \left[ \min_{t \in [0,1]^3} \frac{\partial f_{c(t)}}{\partial x_2}(x_1, x_2), \max_{t \in [0,1]^3} \frac{\partial f_{c(t)}}{\partial x_2}(x_1, x_2) \right] \\ &= \left[ \min_{t \in [0,1]^3} (8 + 2t_3)x_2 + (4 + 4t_2)x_1, \max_{t \in [0,1]^3} (8 + 2t_3)x_2 + (4 + 4t_2)x_1 \right] \\ &= [4x_1 + 8x_2, 8x_1 + 10x_2] \end{aligned}$$

O gradiente é dado por

$$\nabla F(x_1, x_2) = ([6x_1 + 4x_2, 8x_1 + 8x_2], [4x_1 + 8x_2, 8x_1 + 10x_2])$$

Se considerarmos o ponto  $(x_1, x_2) = (1, 1)$ , obtemos

$$\nabla F(1, 1) = ([10, 16], [12, 18])$$

As derivadas de segunda ordem são dadas por

$$\begin{aligned} \frac{\partial^2 f_{c(t)}}{\partial x_1 \partial x_2}(x_1, x_2) &= \frac{\partial^2 f_{c(t)}}{\partial x_2 \partial x_1}(x_1, x_2) = 4 + 4t_2, \\ \frac{\partial^2 f_{c(t)}}{\partial x_1^2}(x_1, x_2) &= 6 + 2t_1 \quad e \quad \frac{\partial^2 f_{c(t)}}{\partial x_2^2}(x_1, x_2) = 8 + 2t_3 \end{aligned}$$

A hessiana de cada  $f_{c(t)}$  é

$$\nabla^2 f_{c(t)}(x_1, x_2) = \begin{bmatrix} 6 + 2t_1 & 4 + 4t_2 \\ 4 + 4t_2 & 8 + 2t_3 \end{bmatrix}$$

Para a função  $F$  temos

$$\begin{aligned} \frac{\partial^2 F}{\partial x_1^2}(x_1, x_2) &= \left[ \min_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_1^2}(x_1, x_2), \max_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_1^2}(x_1, x_2) \right] \\ &= \left[ \min_{t \in [0,1]^3} (6 + 2t_1), \max_{t \in [0,1]^3} (6 + 2t_1) \right] \\ &= [6, 8] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 F}{\partial x_2^2}(x_1, x_2) &= \left[ \min_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_2^2}(x_1, x_2), \max_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_2^2}(x_1, x_2) \right] \\ &= \left[ \min_{t \in [0,1]^3} (8 + 2t_3), \max_{t \in [0,1]^3} (8 + 2t_3) \right] \\ &= [8, 10] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 F}{\partial x_1 \partial x_2}(x_1, x_2) &= \left[ \min_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_1 \partial x_2}(x_1, x_2), \max_{t \in [0,1]^3} \frac{\partial^2 f_{c(t)}}{\partial x_1 \partial x_2}(x_1, x_2) \right] \\ &= \left[ \min_{t \in [0,1]^3} (4 + 4t_2), \max_{t \in [0,1]^3} (4 + 4t_2) \right] \\ &= [4, 8] \end{aligned}$$

A hessiana intervalar de  $F$  é a matriz

$$\nabla^2 F(x_1, x_2) = \begin{bmatrix} [6, 8] & [4, 8] \\ [4, 8] & [8, 10] \end{bmatrix}$$

## 2.3 Funções intervalares convexas

Da construção de função intervalar na forma paramétrica, fica claro que a convexidade de  $F_{C_v^k}$  depende da convexidade de  $f_{c(t)}$  e de uma ordem parcial.

**Definição 2.3.1.** *Sejam  $D \subseteq \mathbb{R}^n$  um conjunto convexo e  $C_v^k \in M^k$ ,  $F_{C_v^k} : D \rightarrow M$ . Para  $x_1, x_2 \in D$ ,  $0 \leq \lambda \leq 1$ ,  $F_{C_v^k}$  é dita convexa com respeito a  $\preceq$  (ou  $\preceq_\omega$ ) se*

$$F_{C_v^k}(\lambda x_1 + (1 - \lambda)x_2) \preceq \text{(ou } \preceq_\omega) \lambda F_{C_v^k}(x_1) + (1 - \lambda)F_{C_v^k}(x_2).$$

Da maneira como foram definidas  $\preceq$  e  $\preceq_\omega$  e convexidade com relação a essas relações de ordem, uma maneira de observar que  $F_{C_v^k}$  é convexa com respeito a  $\preceq$  é

$$f_{c(t')}(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f_{c(t'')}(x_1) + (1 - \lambda)f_{c(t'')}(x_2)$$

para todo  $t', t'' \in [0, 1]^k$ , não necessariamente temos que  $t' = t''$ . Para que  $F_{C_v^k}$  seja convexa com relação a  $\preceq_\omega$  significa

$$f_{c(t)}(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f_{c(t)}(x_1) + (1 - \lambda)f_{c(t)}(x_2)$$

para todo  $t \in [0, 1]^k$ , note que  $t$  é o mesmo em ambos os lados.

De acordo com Panda [1], podemos concluir que  $F_{C_v^k}$  é convexa com relação a  $\preceq_\omega$  se, e somente se,  $f_{c(t)}$  é uma função convexa em  $D$  para todo  $t$ . O mesmo não se pode afirmar para a relação  $\preceq$ .

**Exemplo 2.3.1.** Considere a função  $F : \mathbb{R}^+ \rightarrow I(\mathbb{R})$  definida por  $F(x) = [1, 2]x$ . Na forma paramétrica temos  $f_{c(t)} : \mathbb{R} \rightarrow \mathbb{R}$  onde  $f_{c(t)}(x) = (1 + t)x$ ,  $t \in [0, 1]$ . Seja  $\lambda \in [0, 1]$ .

$$\begin{aligned} f_{c(t)}(\lambda x + (1 - \lambda)y) &= (1 + t)(\lambda x + (1 - \lambda)y) \\ &= (1 + t)\lambda x + (1 + t)(1 - \lambda)y \\ &= \lambda(1 + t)x + (1 - \lambda)(1 + t)y \\ &= \lambda f_{c(t)}x + (1 - \lambda)f_{c(t)}(y) \end{aligned}$$

assim sendo,  $f_{c(t)}$  é convexa para todo  $t \in [0, 1]$  e  $F$  é convexa segundo  $\preceq_\omega$ .

Para que  $F$  seja convexa segundo a relação de ordem  $\preceq$ , é necessário e suficiente que

$$f_{c(t')}(\lambda x + (1 - \lambda)y) \leq \lambda f_{c(t'')}(x) + (1 - \lambda)f_{c(t'')}(y)$$

onde  $t', t'' \in [0, 1]$ . Porém, temos o seguinte

$$f_{c(t')}(\lambda x + (1 - \lambda)y) = (1 + t')(\lambda x + (1 - \lambda)y)$$

e

$$\lambda f_{c(t'')}(x) + (1 - \lambda)f_{c(t'')}(y) = \lambda(1 + t'')x + (1 - \lambda)(1 + t'')y$$

Se considerarmos  $t' = 1$  e  $t'' = 0$  temos

$$\begin{aligned}
f_{c(1)}(\lambda x + (1 - \lambda)y) &= (1 + 1)(\lambda x + (1 - \lambda)y) \\
&= 2(\lambda x + (1 - \lambda)y) \\
&> (\lambda x + (1 - \lambda)y) \\
&= \lambda(1 + 0)x + (1 - \lambda)(1 + 0)y \\
&= \lambda f_{c(0)}(x) + (1 - \lambda)f_{c(0)}(y)
\end{aligned}$$

Portanto  $F$  não é convexa com relação a  $\preceq$ .

### 2.3.1 Relação entre função intervalar convexa e matriz definida positiva

Seja  $A_m \in M^{n \times n}$ ,  $A_m^m = \left( \frac{a_{ij} + \bar{a}_{ij}}{2} \right)_{n \times n}$ ,  $A_m^s = \left( \frac{a_{ij} - \bar{a}_{ij}}{2} \right)_{n \times n}$ , onde o  $(ij)$ -ésimo elemento de  $A_m$  é um intervalo  $[\underline{a}_{ij}, \bar{a}_{ij}]$ .

A matriz  $A_m$  pode ser expressada como  $A_m = [A_m^m - A_m^s, A_m^m + A_m^s]$ . Dada uma matriz real  $\bar{A}_{n \times n}$ , dizemos que  $\bar{A} \in A_m$  se  $a_{ij} \in [\underline{a}_{ij}, \bar{a}_{ij}]$ , onde  $\bar{A} = (a_{ij})$ . A matriz  $A_m$  é dita simétrica se  $A_m^m$  e  $A_m^s$  são matrizes simétricas.

Para cada matriz intervalar  $A_m$  podemos associar a uma matriz intervalar simétrica  $A'_m = [A'^m - A'^s, A'^m + A'^s]$ , onde

$$A'^m = \frac{(A_m^m)^T + A_m^m}{2} \text{ e } A'^s = \frac{(A_m^s)^T + A_m^s}{2}$$

Seja  $Y = \{z \in \mathbb{R}^n; |z_j| = 1, j = 1, 2, \dots, n\}$  e  $T_z$  uma matriz diagonal de ordem  $n$ . Para cada  $z \in Y$ , defina a matriz  $A_z = A_m^m - T_z A_m^s T_z$ . Então para cada  $i, j$ , temos  $(A_z)_{ij} = (A_m^m)_{ij} - z_i (A_m^s)_{ij} z_j$ . Então

$$(A_z)_{ij} = \begin{cases} (A_m^m - A_m^s)_{ij} & , z_i z_j = 1 \\ (A_m^m + A_m^s)_{ij} & , z_i z_j = -1 \end{cases}$$

**Definição 2.3.2.** Dizemos que  $A_m$  é definida positiva (semi definida positiva) se toda  $A(t) \in A_m$  é definida positiva (semi definida positiva), onde

$$A(t) = (a_{ij}(t_{ij}))_{n \times n}$$

é uma matriz real de ordem  $n$ , e

$$a_{ij}(t_{ij}) = (1 - t_{ij})\underline{a}_{ij} + t_{ij}\bar{a}_{ij}, \quad 0 \leq t_{ij} \leq 1, \quad i, j = 1, 2, \dots, n.$$

**Proposição 2.3.1.** *As seguintes afirmações são equivalentes:*

- i)  $A_m$  é definida positiva (semi definida positiva);*
- ii)  $A'_m$  é definida positiva (semi definida positiva);*
- iii)  $A_z$  é definida positiva (semi definida positiva) para cada  $z \in Y$ .*

*Demonstração.* Ver Rohn [8]. □

**Teorema 2.3.1.** *Dado  $C_v^k \in M^k$ , seja  $F_{C_v^k}$  uma função intervalar duas vezes diferenciável em um conjunto aberto e convexo  $D \subseteq \mathbb{R}^n$ . Então  $F_{C_v^k}$  é convexa com respeito a  $\preceq_\omega$  se, e somente se, a matriz Hessiana  $\nabla^2 F_{C_v^k}$  é semi definida positiva.*

*Demonstração.* Sabemos que  $F_{C_v^k}$  é convexa em  $D$  com respeito a  $\preceq_\omega$  se, e somente se, para todo  $x_1, x_2 \in D$ ,  $\lambda \in [0, 1]$ ,

$$\begin{aligned}
& F_{C_v^k}(\lambda x_1 + (1 - \lambda)x_2) \preceq_\omega \lambda F_{C_v^k}(x_1) + (1 - \lambda)F_{C_v^k}(x_2) \\
& \Leftrightarrow f_{c(t)}(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f_{c(t)}(x_1) + (1 - \lambda)f_{c(t)}(x_2) \\
& \Leftrightarrow f_{c(t)} \text{ é convexa em } D \text{ para todo } t \in [0, 1]^k \\
& \Leftrightarrow \nabla^2 f_{c(t)}(x) \text{ é semi definida positiva para todo } t \in [0, 1]^k \\
& \Leftrightarrow \nabla^2 F_{C_v^k}(x) \text{ é uma matriz intervalar semi definida positiva.}
\end{aligned}$$

□

Na forma clássica, a derivada de uma função intervalar  $F(x) = [\underline{F}(x), \overline{F}(x)]$  depende das derivadas de  $\underline{F}$  e  $\overline{F}$ . Assim  $\nabla F(x)$  é o vetor intervalar  $[\nabla \underline{F}(x), \nabla \overline{F}(x)]$  e  $\nabla^2 F(x)$  é a matriz intervalar  $([\nabla^2 \underline{F}(x), \nabla^2 \overline{F}(x)])_{ij}$ . Segundo WU [10], na forma clássica,  $F$  é uma função convexa com respeito a  $\preceq_{LC}$  se, e somente se,  $\underline{F}$  e  $m(F)$  são funções convexas;  $F$  é uma função convexa com respeito a  $\preceq_{RC}$  se, e somente se,  $m(F)$  e  $\overline{F}$  são funções convexas;  $F$  é uma função convexa com respeito a  $\preceq_{LR}$  se, e somente se,  $\underline{F}$  e  $\overline{F}$  são funções convexas. Isso significa que a convexidade de  $F$  depende da semi positividade das matrizes  $\nabla^2 \underline{F}$ ,  $\nabla^2 m(F)$  e  $\nabla^2 \overline{F}$  nos três diferentes casos. Em particular, o fato de que  $\nabla^2 \underline{F}$  e  $\nabla^2 \overline{F}$  são semi definidas positivas não nos dá a garantia de que a matriz intervalar  $([\nabla^2 \underline{F}(x), \nabla^2 \overline{F}(x)])_{ij}$  também seja semi definida positiva, pois uma matriz intervalar  $A_m$  é semi definida positiva se toda matriz real  $A \in A_m$  é semi definida positiva.

Assim, não é verdade que  $F$  é convexa com relação a  $\preceq_{LR}$  se, e somente se,  $([\nabla^2 \underline{F}(x), \nabla^2 \overline{F}(x)])_{ij}$  é semi definida positiva. Argumentos similares valem para  $\preceq_{RC}$  e  $\preceq_{LC}$ .

Essa dificuldade pode ser resolvida quando considera-se uma função intervalar calculada na forma paramétrica. Na forma paramétrica, a existência da derivada de  $F_{C_v^k}$  depende da existência da derivada de  $f_{c(t)}$  para todo valor de  $t$ . Assim,  $\nabla^2 F_{C_v^k}(x) = \{\nabla^2 f_{c(t)}(x) | f_{c(t)}(x) \in F_{C_v^k}(x)\}$ . A função intervalar  $F_{C_v^k}$  é uma função convexa se, e somente se,  $f_{c(t)}$  é uma função convexa para todo  $t$ . Isso significa que para todo  $t$ ,  $\nabla^2 f_{c(t)}(x)$  é uma matriz semi definida positiva. De acordo com Panda [1], isso implica  $\nabla^2 F_{C_v^k}(x)$  seja uma matriz intervalar semi definida positiva.

## 2.4 Problema de Otimização Intervalar

Nesta seção vamos considerar o Problema de Otimização Intervalar (IOP) como sendo

$$(IOP) \min F_{C_v^k}(x) \\ \text{sujeito a } G_{jD_v^{m_j}}(x) \preceq (\text{ou } \preceq_\omega) B_j, j = 1, 2, \dots, p$$

onde  $B_j \in M$ , as funções intervalares  $F_{C_v^k}, G_{jD_v^{m_j}} : \mathbb{R}^n \rightarrow M$  são definidas por

$$F_{C_v^k}(x) = \{f_{c(t)}(x) | f_{c(t)} : \mathbb{R}^n \rightarrow \mathbb{R}, t \in [0, 1]^k\} \text{ e,} \\ G_{jD_v^{m_j}}(x) = \{g_{jd(t'_j)}(x) | g_{jd(t'_j)} : \mathbb{R}^n \rightarrow \mathbb{R}; d(t'_j) \in D_v^{m_j}\}$$

Dependendo da relação de ordem escolhida, a região factível de (IOP) pode ser expressada como sendo o conjunto

$$S = \{x \in \mathbb{R}^n : G_{jD_v^{m_j}}(x) \preceq B_j, j = 1, 2, \dots, p\} \\ = \{x \in \mathbb{R}^n : g_{jd(t'_j)}(x) \leq b_j(t'_j), \forall j\} \\ = \{x \in \mathbb{R}^n : \max_{t_j \in [0, 1]^{m_j}} g_{jd(t_j)}(x) \leq b_j(0), \forall j\}$$

ou,

$$S = \{x \in \mathbb{R}^n : G_{jD_v^{m_j}}(x) \preceq_\omega B_j, j = 1, 2, \dots, p\} \\ = \{x \in \mathbb{R}^n : g_{jd(t_j)}(x) \leq b_j(t_j), \forall j\} \\ = \{x \in \mathbb{R}^n : \max_{t_j \in [0, 1]^{m_j}} g_{jd(t_j)}(x) \leq b_j(1), \min_{t_j \in [0, 1]^{m_j}} g_{jd(t_j)}(x) \leq b_j(0), \forall j\}$$

ou ainda,

$$\begin{aligned}
S &= \{x \in \mathbb{R}^n : G_{jD_v^{m_j}}(x) \preceq \widehat{B}_j, j = 1, 2, \dots, p\} \\
&= \{x \in \mathbb{R}^n : g_{jd(t_j)}(x) \leq b_j, \forall j\} \\
&= \{x \in \mathbb{R}^n : \max_{t_j \in [0,1]^{m_j}} g_{jd(t_j)}(x) \leq b_j, \forall j\}
\end{aligned}$$

onde  $d(t_j), d(t'_j) \in D_v^{m_j}$ ,  $b_j(t'_j), b_j(t_j) \in B_j$ ,  $b_j \in \widehat{B}_j = [b_j, b_j]$ ,  $t'_j \in [0, 1]^{m_j}$ ,  $t_j, t''_j \in [0, 1]$ .

Como  $\min_{x \in S} F_{C_v^k}(x) = \min_{x \in S} \{f_{c(t)}(x) | c(t) \in C_v^k, t \in [0, 1]^k\}$ , então (IOP) pode ser tratado como um problema de otimização multiobjetivo em  $t$  para todo  $x \in S$  sobre um domínio contínuo, que é o paralelepípedo retangular  $C_v^k$ . De acordo com Panda [1], isso significa que para todo  $c(t)$ , o problema de otimização

$$(IOP)_t \min_{x \in S} f_{c(t)}(x)$$

tem solução.

**Definição 2.4.1.** Dizemos que  $x^* \in S$  é uma solução eficiente de (IOP) se não existe  $x \in S$  com

$$f_{c(t)}(x) \leq f_{c(t)}(x^*), \forall t \in [0, 1]^k \text{ e } F_{C_v^k}(x) \neq F_{C_v^k}(x^*)$$

**Definição 2.4.2.** Dizemos que  $x^* \in S$  é uma solução eficiente própria de (IOP), se  $x^* \in S$  é uma solução eficiente e existe um número real  $\mu > 0$  tal que para algum  $t \in [0, 1]^k$  e todo  $x \in S$  com  $f_{c(t)}(x) < f_{c(t)}(x^*)$ , pelo menos um  $t' \in [0, 1]^k$ ,  $t \neq t'$  existe com  $f_{c(t')}(x) > f_{c(t')}(x^*)$  e

$$\frac{f_{c(t)}(x^*) - f_{c(t)}(x)}{f_{c(t')}(x) - f_{c(t')}(x^*)} \leq \mu$$

Considere o seguinte problema de otimização com respeito a uma função de peso  $\omega : [0, 1]^k \rightarrow \mathbb{R}^+$  dado por

$$(IOP)_I \min_{x \in S} \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x) dt_1 dt_2 \dots dt_k,$$

onde  $\omega(t) = \omega(t_1, t_2, \dots, t_k)$ .

Aqui  $t_1, t_2, \dots, t_k$  são mutuamente independentes e cada  $t_i$  varia de 0 a 1. Assim sendo,  $\int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x) dt_1 dt_2 \dots dt_k$  é uma função que depende  $x$  somente, digamos  $h(x)$ . Logo,  $(IOP)_I$  se torna  $\min_{x \in S} h(x)$ , que em geral é uma problema

de programação não linear e sem incertezas intervalares, que pode ser resolvido com técnicas de programação não linear.

**Teorema 2.4.1.** *Se  $x^* \in S$  é uma solução ótima de  $(IOP)_I$ , então  $x^*$  é uma solução eficiente própria de  $(IOP)$ .*

*Demonstração.* Seja  $x^* \in S$  uma solução ótima de  $(IOP)_I$ . Suponha que  $x^*$  não é uma solução eficiente própria de  $(IOP)$ . Assim, para algum  $t \in [0, 1]^k$  e algum  $x \in S$  têm-se  $f_{c(t)}(x) < f_{c(t)}(x^*)$ . Considere a função ponderadora  $\omega : [0, 1]^k \rightarrow \mathbb{R}^+$  contínua.

Tomando  $\mu = \max \left\{ \frac{\omega(t')}{\omega(t)} \right\}$ ,  $t \neq t'$ ,  $t, t' \in [0, 1]^k$ ,  $\omega(t) > 0$ , satisfazendo

$$\frac{f_{c(t)}(x^*) - f_{c(t)}(x)}{f_{c(t')}(x) - f_{c(t')}(x^*)} > \mu$$

para todo  $t' \in [0, 1]^k$  com  $f_{c(t')}(x) > f_{c(t')}(x^*)$ .

Assim,

$$f_{c(t)}(x^*) - f_{c(t)}(x) > \mu (f_{c(t')}(x) - f_{c(t')}(x^*)) > \frac{\omega(t')}{\omega(t)} (f_{c(t')}(x) - f_{c(t')}(x^*))$$

Logo  $\omega(t)f_{c(t)}(x^*) - \omega(t)f_{c(t)}(x) > \omega(t')f_{c(t')}(x) - \omega(t')f_{c(t')}(x^*)$ .

Ora,

$$\begin{aligned} & \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x^*) dt_1 dt_2 \dots dt_k - \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x) dt_1 dt_2 \dots dt_k \\ & > \\ & \int_0^1 \int_0^1 \dots \int_0^1 \omega(t') f_{c(t')}(x) dt'_1 dt'_2 \dots dt'_k - \int_0^1 \int_0^1 \dots \int_0^1 \omega(t') f_{c(t')}(x^*) dt'_1 dt'_2 \dots dt'_k. \end{aligned}$$

Portanto

$$\int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x^*) dt_1 dt_2 \dots dt_k > \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x) dt_1 dt_2 \dots dt_k$$

Isso contradiz a suposição de que  $x^* \in S$  é uma solução ótima de  $(IOP)_I$ .  $\square$

## 2.5 Problema de Programação Convexa Intervalar

Um problema de otimização intervalar (IOP) é dito problema de programação convexa intervalar se  $F_{C_v^k}$  e  $G_{D_v^{m_j}}$  são funções convexas com respeito a  $\preceq$  ou  $\preceq_\omega$ .



**Teorema 2.5.1.** *Se (IOP) é um problema de programação convexa intervalar, então  $(IOP)_I$  é um problema de programação convexa.*

*Demonstração.* Suponha que (IOP) é um problema de programação convexa intervalar. Então o conjunto factível  $S$  é um conjunto convexo e a função objetivo  $F_{C_v^k}$  é uma função convexa com respeito a  $\preceq$  ou  $\preceq_\omega$ . Para  $x_1, x_2 \in S$  e  $\lambda \in [0, 1]$ ,

$$F_{C_v^k}(\lambda x_1 + (1 - \lambda)x_2) \preceq \text{ (ou } \preceq_\omega) \lambda F_{C_v^k}(x_1) \oplus (1 - \lambda)F_{C_v^k}(x_2)$$

No caso de ambas as ordens parciais  $\preceq$  ou  $\preceq_\omega$ , a relação acima implica que para todo  $t \in [0, 1]^k$ ,  $f_{c(t)}(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f_{c(t)}(x_1) + (1 - \lambda)f_{c(t)}(x_2)$ . Multiplicando por  $\omega(t)$  e integrando com respeito a  $t_1, t_2, \dots, t_k$ , temos

$$\lambda \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(\lambda x_1 + (1 - \lambda)x_2) dt_1 dt_2 \dots dt_k \leq \\ \lambda \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x_1) dt_1 dt_2 \dots dt_k + (1 - \lambda) \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) f_{c(t)}(x_2) dt_1 dt_2 \dots dt_k$$

Isso implica que a função objetivo de  $(IOP)_I$  é uma função convexa e o seu conjunto factível é o conjunto  $S$ , que é um conjunto convexo. Assim,  $(IOP)_I$  é um problema de programação intervalar convexo.  $\square$

Um problema de programação quadrática intervalar é um problema da seguinte forma:

$$\begin{aligned} \min C_v^n \bullet x + \frac{1}{2} x^T Q_m x \\ \text{sujeito a } A_m x \preceq \text{ (ou } \preceq_\omega) B_v^p \\ x \geq 0, x \in \mathbb{R}^n \end{aligned}$$

onde  $C_v^n \bullet x = \sum_{j=1}^n C_j x_j$ ,  $C_v^n \in M^n$ ,  $B_v^p \in M^p$  e  $A_m = (A_{ij})_{p \times n}$  é uma matriz intervalar com  $A_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$ ,  $Q_m = (Q_{ij})_{n \times n}$  é uma matriz intervalar simétrica definida positiva

Qualquer elemento nos vetores intervalares  $C_v^n, B_v^p$  são vetores reais de ordem  $n$  e  $p$  respectivamente, que podem ser expressos sob a forma paramétrica, digamos  $c(t)$  e  $b(t)$ . Similarmente qualquer elemento nas matrizes intervalares  $A_m$  e  $Q_m$  são matrizes reais de ordem  $p \times n$  e  $n \times n$  respectivamente, que também podem ser expressadas sob a forma paramétrica como  $A(t) = (a_{ij}(t_{ij}))_{p \times n}$  e  $Q(t) = (q_{ij}(t_{ij}))_{n \times n}$ . O conjunto factível para esse problema pode ser calculado como

$$\begin{aligned}
S &= \{x \in \mathbb{R}^n; A(1)(x) \leq b(0), (A_m(x) \preceq B)\} \\
S &= \{x \in \mathbb{R}^n; A(1)(x) \leq b(1), A(0) \leq b(0), (A_m(x) \preceq_\omega B)\} \\
S &= \{x \in \mathbb{R}^n; A(1)(x) \leq b, (A_m(x) \preceq \hat{B}), \hat{B} = [b_j, b_j], b_j \in \mathbb{R}\}
\end{aligned}$$

Nos três casos,  $S$  é um conjunto convexo. A matriz intervalar Hessiana da função intervalar  $C_v^n \bullet x + \frac{1}{2}x^T Q_m x$  quando calculada de acordo com o capítulo anterior, resultará em  $Q_m$ , que é definida positiva. Assim sendo, a função  $C_v^n \bullet x + \frac{1}{2}x^T Q_m x$  é uma função convexa com relação a  $\preceq_\omega$ . Portanto o problema de programação quadrática intervalar é um problema de programação convexa intervalar. O problema de otimização correspondente com a função ponderadora é

$$(IOP)_I \min_{x \in S} \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) \left[ c(t')x + \frac{1}{2}x^T Q(t'')x \right] dt' dt''$$

onde  $c(t') \in C_v^n$ ,  $Q(t'') \in Q_m$ ,  $dt' = dt'_1 dt'_2 \dots dt'_n$ ,  $dt'' = dt''_{ij}$ ,  $i, j = 1, 2, \dots, n$ ,  $\omega : [0, 1]^{n^2+n} \rightarrow \mathbb{R}^+$ ,  $t = (t', t'')$ . Pelo Teorema 2.5.1,  $(IOP)_I$  é um problema de programação quadrática convexa. A solução ótima desse problema pode ser obtida resolvendo as condições de otimalidade de KKT, e essa solução ótima é uma solução eficiente de (IOP) pelo Teorema 2.4.1. Denote

$$h(x) = \int_0^1 \int_0^1 \dots \int_0^1 \omega(t) \left[ c(t')x + \frac{1}{2}x^T Q(t'')x \right] dt' dt''$$

Se  $S = \{x \in \mathbb{R}^n; A(1)x \leq b(0)\}$  quando consideramos  $A_m(x) \preceq B$ , então a função Lagrangiana é  $L(x, \lambda, \mu) = h(x) + \lambda^T (A(1)x - b(0)) - \mu^T x$ ,  $\lambda \in \mathbb{R}^m$ ,  $\mu \in \mathbb{R}^n$ ,  $\lambda \geq 0$ . As condições de otimalidade de KKT são  $\nabla_x L(x, \lambda, \mu) = 0$ ,  $\lambda^T (A(1)x - b(0)) = 0$ ,  $x \in S$ , que são equivalentes a

$$\int_0^1 \int_0^1 \dots \int_0^1 \omega(t) \left[ c(t') + Q(t'')x \right] dt' dt'' + \lambda^T A(1) = \mu \quad (2.4)$$

$$\lambda^T (A(1)x - b(0)) = 0, \mu^T x = 0, \lambda \geq 0, x \in S \quad (2.5)$$

Se  $S$  assume diferentes formas de acordo com a relação de ordem escolhida, então as condições de KKT podem ser determinadas de formas similares.

# 3 Busca Direcional para Funções com Coeficientes Intervalares

## 3.1 Caracterização da Direção de Descida

Nesta seção vamos considerar o problema de otimização intervalar irrestrita, dado por

$$\min_{x \in \mathbb{R}^n} F_{C^k}(x) \quad (3.1)$$

Mais adiante vamos caracterizar direção de descida utilizando o conceito de cone polar. Para isto temos a seguinte definição.

**Definição 3.1.1.** Um conjunto  $K \subset \mathbb{R}^n$  chama-se cone quando

$$d \in K \Rightarrow td \in K, \forall t \in \mathbb{R}^+$$

Pela definição, se  $K$  é um cone não vazio, necessariamente  $0 \in K$ . Informalmente, um cone é um conjunto de direções. Um cone não vazio é sempre ilimitado, com exceção do cone trivial  $\{0\}$ .

**Definição 3.1.2.** O cone dual (ou cone polar) de um cone  $K \subset \mathbb{R}^n$  é definido por

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, d \rangle \leq 0, \forall d \in K\}$$

No resultado a seguir tem-se a caracterização de um cone polar.

**Proposição 3.1.1.** Para todo cone  $\emptyset \neq K \subset \mathbb{R}^n$ , o cone dual  $K^*$  sempre é convexo e fechado.

*Demonstração.* Sejam  $x, y \in K^*$ , isto é,  $\langle x, d \rangle \leq 0$  e  $\langle y, d \rangle \leq 0$  para todo  $d \in K$ . Seja  $\alpha \in [0, 1]$ . Para qualquer  $d \in K$ , temos

$$\langle \alpha x + (1 - \alpha)y, d \rangle = \alpha \langle x, d \rangle + (1 - \alpha) \langle y, d \rangle \leq 0$$

ou seja,  $\alpha x + (1 - \alpha)y \in K^*$ .

Seja  $(x_n)_{n \in \mathbb{N}}$  uma sequência em  $K^*$  tal que  $x_n \rightarrow x$  quando  $n \rightarrow \infty$ . Tome  $d \in K$ . Como  $x_n \in K$ , então  $\langle x_n, d \rangle \leq 0$  para todo  $n$ . Passando ao limite, temos  $\langle x, d \rangle \leq 0$ , ou seja,  $x \in K$ . Portanto  $K$  é fechado e convexo.

□

**Definição 3.1.3.** *Seja  $F_{C_v^k} : \mathbb{R}^n \rightarrow M$ . Dizemos que  $d \in \mathbb{R}^n$  é uma direção de descida para  $F_{C_v^k}$  no ponto  $\bar{x} \in \mathbb{R}^n$ , se existe  $\bar{\lambda} > 0$  tal que*

$$F_{C_v^k}(\bar{x} + \lambda d) \prec_{\omega} \text{ (ou } \prec) F_{C_v^k}(\bar{x}), \lambda \in (0, \bar{\lambda})$$

*Se considerarmos  $\prec_{\omega}$ , a desigualdade acima é equivalente a*

$$f_{c(t)}(\bar{x} + \lambda d) < f_{c(t)}(\bar{x}), t \in [0, 1]^k, \lambda \in (0, \bar{\lambda}).$$

O seguinte resultado nos dá uma caracterização de uma direção de descida para um problema de otimização intervalar. Veremos que sob algumas hipóteses para  $f_{c(t)}$ , toda direção no interior do cone polar do cone gerado pelo gradiente da função intervalar é uma direção de descida.

**Teorema 3.1.1.** *Sejam  $F_{C_v^k} : \mathbb{R}^n \rightarrow M$  diferenciável,  $\bar{x} \in \mathbb{R}^n$  e  $d \in \mathbb{R}^n \setminus \{0\}$  tais que  $\nabla F_{C_v^k}(\bar{x}) \bullet d \prec_{\omega} 0$ . Considere a aplicação  $g : [0, 1]^k \rightarrow \mathbb{R}^n$  dada por  $g(t) = \nabla f_{c(t)}(\bar{x})$ . Suponha que  $g$  é contínua. Então existe  $\bar{\alpha} > 0$  tal que*

$$F_{C_v^k}(\bar{x} + \alpha d) \prec_{\omega} F_{C_v^k}(\bar{x}), \alpha \in (0, \bar{\alpha}).$$

*Em outras palavras,  $d$  é uma direção de descida para  $F_{C_v^k}$  em  $\bar{x}$ .*

*Demonstração.* Como a aplicação  $g$  é contínua, o conjunto  $g([0, 1]^k) \subset \mathbb{R}^n$  é compacto. Seja o problema

$$\max \langle r, d \rangle \tag{3.2}$$

$$r \in g([0, 1]^k)$$

Como  $\langle r, d \rangle$  é contínua e  $\langle \nabla f_{c(t)}(\bar{x}), d \rangle < 0$  para todo  $t \in [0, 1]^k$ , então existe  $r^* \in g([0, 1]^k)$  tal que

$$0 > \langle r^*, d \rangle \geq \langle r, d \rangle, \text{ para todo } r \in g([0, 1]^k)$$

Para cada  $t \in [0, 1]^k$ , considere a aproximação de primeira ordem

$$f_{c(t)}(\bar{x} + \alpha d) = f_{c(t)}(\bar{x}) + \alpha \langle \nabla f_{c(t)}(\bar{x}), d \rangle + o_t(\alpha)$$

onde

$$\lim_{\alpha \rightarrow 0} \frac{o_t(\alpha)}{\alpha} = 0$$

Tem-se então que

$$\begin{aligned} f_{c(t)}(\bar{x} + \alpha d) &= f_{c(t)}(\bar{x}) + \alpha \langle \nabla f_{c(t)}(\bar{x}), d \rangle + o_t(\alpha) \\ f_{c(t)}(\bar{x} + \alpha d) - f_{c(t)}(\bar{x}) &= \alpha \langle \nabla f_{c(t)}(\bar{x}), d \rangle + o_t(\alpha) \\ &\leq \alpha \langle r^*, d \rangle + o_t(\alpha) \\ &= \alpha \left( \langle r^*, d \rangle + \frac{o_t(\alpha)}{\alpha} \right) \end{aligned}$$

Tomando  $\alpha > 0$  suficientemente pequeno e  $\delta > 0$  conveniente de modo que

$$\frac{o_t(\alpha)}{\alpha} < \delta, \text{ para todo } t \in [0, 1]^k$$

e  $\delta < |\langle r^*, d \rangle|$ , então

$$f_{c(t)}(\bar{x} + \alpha d) - f_{c(t)}(\bar{x}) < 0, \text{ para todo } t \in [0, 1]^k$$

ou seja, existe  $\bar{\alpha} > 0$  tal que se  $\alpha \in (0, \bar{\alpha})$ ,

$$f_{c(t)}(\bar{x} + \alpha d) < f_{c(t)}(\bar{x}), \text{ para todo } t \in [0, 1]^k.$$

Portanto

$$F_{C_v^k}(\bar{x} + \alpha d) \prec_{\omega} F_{C_v^k}(\bar{x}), \alpha \in (0, \bar{\alpha}).$$

□

Seja  $F_{C_v^k}$  de acordo com as hipóteses do Teorema 3.1.1 e  $\bar{x} \in \mathbb{R}^n$ . Considere o problema (3.1). Seja  $K$  o cone gerado por  $\nabla F_{C_v^k}(\bar{x})$ , ou seja,

$$K \equiv \{y \in \mathbb{R}^n / y = \lambda \nabla f_{c(t)}(\bar{x}), \lambda \geq 0, t \in [0, 1]^k\}.$$

Note que se  $\text{int}(K^*) \neq \emptyset$ , então  $\bar{x}$  não é solução eficiente do problema (3.1), pois se  $\text{int}(K^*) \neq \emptyset$  então existe  $d \in \mathbb{R}^n \setminus \{0\}$  tal que

$$\langle d, \lambda \nabla f_{c(t)}(\bar{x}) \rangle < 0, \lambda > 0, t \in [0, 1]^k$$

isto é,

$$\nabla F_{C_v^k} \bullet d \prec_{\omega} 0.$$

Portanto, pelo Teorema 3.1.1,  $d$  é uma direção de descida para  $F_{C_v^k}$  em  $\bar{x}$  e de acordo com a Definição 2.4.1,  $\bar{x}$  não é solução eficiente para o problema (3.1).

**Exemplo 3.1.1.** *Seja a função  $F : \mathbb{R}^2 \rightarrow M$  dada por  $F(x_1, x_2) = [1, 3]x_1^2 + [5, 6]x_2^2$ . Na forma paramétrica temos*

$$f_{c(t)}(x_1, x_2) = c_1(t)x_1^2 + c_2(t)x_2^2$$

onde  $c_1(t_1) = (1 - t_1) \cdot 1 + 3t_1 = 1 + 2t_1$ ,  $c_2(t_2) = (1 - t_2) \cdot 5 + 6t_2 = 5 + t_2$  e  $t_1, t_2 \in [0, 1]$ .

Considere o problema

$$\min [1, 3]x_1^2 + [5, 6]x_2^2$$

$$\text{sujeito a. } x_1, x_2 \geq 0$$

Agora considere o ponto  $\bar{x} = (1, 1)$ . Temos que

$$\begin{aligned} f_{c(t)}(1, 1) &= (1 + 2t_1)(1)^2 + (5 + t_2)(1)^2 \\ &= 6 + 2t_1 + t_2 \\ &\geq 6 \end{aligned}$$

pois  $t_1, t_2 \in [0, 1]$ . Além disso,  $F(\bar{x}) = [6, 9]$ . Note que  $\bar{x} = (1, 1)$  não é solução eficiente do problema, basta considerar  $x^* = (0, 0)$  temos que  $f_{c(t)}(x^*) = 0$  e  $F(x^*) = [0, 0]$ , logo

$$f_{c(t)}(x^*) \leq f_{c(t)}(\bar{x})$$

e

$$F(x^*) \neq F(\bar{x}).$$

Vamos calcular o gradiente de  $F$  no ponto  $\bar{x}$ . Temos que

$$\nabla f_{c(t)}(x_1, x_2) = (2c_1(t)x_1, 2c_2(t)x_2)$$

Portanto

$$\begin{aligned}\nabla f_{c(t)}(1, 1) &= (2c_1(t), 2c_2(t)) \\ &= (2(1 + 2t_1), 2(5 + t_2)) \\ &= (2 + 4t_1, 10 + 2t_2)\end{aligned}$$

onde  $t_1, t_2 \in [0, 1]$ . Assim sendo, o gradiente da  $F$  em  $(1, 1)$  é  $\nabla F(1, 1) = ([2, 6], [10, 12])$ . Vamos considerar o cone  $K_{\bar{x}}$  gerado por  $\nabla F(1, 1)$ .

$$K_{\bar{x}} \equiv \{y \in \mathbb{R}^2 \mid y = \lambda \nabla f_{c(t)}(1, 1), \lambda \geq 0\} \quad (3.3)$$

Assim sendo, passamos a analisar agora o interior do cone polar  $K_{\bar{x}}^*$ . Ora,

$$K_{\bar{x}}^* \equiv \{v \in \mathbb{R}^2 \mid \langle v, \lambda \nabla f_{c(t)}(1, 1) \rangle \leq 0\}$$

A direção  $\bar{d} = (-1, -1)$  está no interior do cone polar  $K_{\bar{x}}^*$ ,

$$\begin{aligned}\langle \bar{d}, \lambda \nabla f_{c(t)}(1, 1) \rangle &= (-1)(2 + 4t_1) + (-1)(10 + 2t_2) \\ &= -12 - 4t_1 - 2t_2 \\ &< 0\end{aligned}$$

Note que  $\bar{d}$  é uma direção de descida para toda  $f_{c(t)}$  em  $(1, 1)$ , pois  $\langle \bar{d}, \nabla f_{c(t)}(1, 1) \rangle < 0$ , ou seja, para cada  $t \in [0, 1]^2$  existe  $\lambda_t > 0$  tal que

$$f_{c(t)}(\bar{x} + \lambda \bar{d}) < f_{c(t)}(\bar{x}), \lambda \in (0, \lambda_t)$$

Afirmo que  $\bar{d}$  é uma direção de descida para  $F_{C_{\bar{x}}^*}$  em  $\bar{x}$ . De fato, considere  $\bar{\lambda} = 2$ . Temos  $f_{c(t)}(\bar{x}) = 6 + 2t_1 + t_2$  e além disso:

$$\begin{aligned}\bar{x} + \bar{\lambda} \bar{d} &= (1, 1) + \bar{\lambda}(-1, -1) \\ &= (1 - \bar{\lambda}, 1 - \bar{\lambda})\end{aligned}$$

$$\begin{aligned}f_{c(t)}(\bar{x} + \bar{\lambda} \bar{d}) &= f_{c(t)}(1 - \bar{\lambda}, 1 - \bar{\lambda}) \\ &= (1 + 2t_1)(1 - \bar{\lambda})^2 + (5 + t_2)(1 - \bar{\lambda})^2 \\ &= (6 + 2t_1 + t_2)(1 - \bar{\lambda})^2\end{aligned}$$

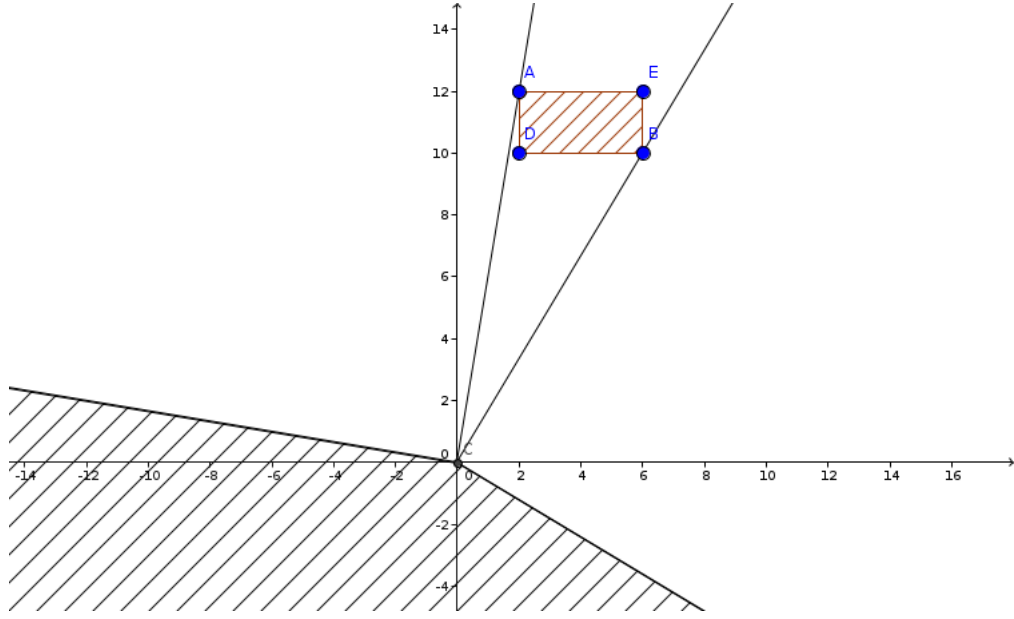


Figura 1 – Gradiente e cone polar ao cone gerado pelo gradiente de  $F$  em  $(1, 1)$

Para todo  $\lambda$  em  $(0, \bar{\lambda})$ , tem-se

$$\begin{aligned}
 (1 - \lambda)^2 &< 1 \\
 (6 + 2t_1 + t_2)(1 - \lambda)^2 &< (6 + 2t_1 + t_2) \\
 f_{c(t)}(\bar{x} + \lambda\bar{d}) &< f_{c(t)}(\bar{x}) \\
 &\Leftrightarrow \\
 F_{C_v^k}(\bar{x} + \lambda\bar{d}) &\prec_{\omega} F_{C_v^k}(\bar{x})
 \end{aligned}$$

□

Com base nos resultados obtidos até agora, podemos fazer um esboço de algoritmo genérico para otimização intervalar que usa direção de descida.

**Algoritmo 3.1:** Seja  $F_{C_v^p}$  de acordo com as hipóteses do Teorema 3.1.1 e  $x^k \in \mathbb{R}^n$ ,  $k \in \mathbb{N}$ . Seja  $K$  o cone gerado por  $\nabla F_{C_v^p}(x^k)$ , ou seja,  $K \equiv \{y \in \mathbb{R}^n / y = \lambda \nabla f_{c(t)}(x^k), \lambda \geq 0\}$ .

- 1) Se  $\text{int}(K^*) = \emptyset$ , parar.
- 2) Senão, obter  $d^k \in \text{int}(K^*)$  e  $t_k > 0$  tais que  $F_{C_v^p}(x^k + t_k d^k) \prec_{\omega} F_{C_v^p}(x^k)$ .
- 3) Faça  $x^{k+1} = x^k + t_k d^k$  e  $k = k + 1$ .
- 4) Volte ao passo 1.



# Conclusão

Os modelos de otimização intervalar tem sido pesquisados nos últimos anos, uma vez que esses modelos representam melhor alguns problemas de otimização que modelam casos reais com a presença de incertezas sobre os coeficientes tecnológicos.

Inicialmente foi feito um desenvolvimento teórico em assuntos relacionados a Análise Intervalar. O conjunto  $I(\mathbb{R})^n$  formado pelas n-uplas onde cada coordenada é um intervalo fechado e limitado não é um espaço vetorial. No entanto, considera-se uma extensão desse espaço para que, através de uma bijeção entre o espaço intervalar  $M^n = I(\mathbb{R})^n \cup \overline{I(\mathbb{R})^n}$  com  $\mathbb{R}^{2n}$  o espaço dos intervalos generalizado foi munido com uma estrutura vetorial parcialmente ordenado. Uma vez definida o que é função com valor intervalar e uma relação de ordem parcial, foi possível definir um problema de otimização com variáveis reais assumindo valores nesse ambiente e o que é uma solução para esse problema. As funções intervalares foram estudadas sob a forma paramétrica para obter resultados que relacionam a matriz da Hessiana com a convexidade da função.

Uma suficiência para a existência de solução eficiente própria para um problema de otimização intervalar, é verificar a existência de solução em um problema de otimização clássico na forma paramétrica. O problema de otimização convexo intervalar possui resultados semelhantes ao problema de otimização convexo clássico.

Ao final do trabalho apresentou-se uma caracterização para direção de descida para o problema de otimização intervalar irrestrito baseado na ideia de cone e, além disso, um algoritmo genérico baseado em busca direcional.

Com o intuito de progredir na pesquisa, abordando aspectos não estudados no presente trabalho ou de melhorar as formulações apresentadas, faz-se, em seguida, algumas sugestões e considerações para trabalhos futuros:

- i) Qualificar a escolha do tamanho do passo e da direção no algoritmo proposto;
- ii) Analisar a taxa de convergência do algoritmo de busca direcional;
- iii) Avaliar o teorema de convergência global.

# Referências

- [1] A. K. Bhurjee and G. Panda. Efficient solution of interval optimization problem. *Mathematical Methods of Operations Research*, 76(3):273–288, 2012.
- [2] T. M. Costa. *Espaços vetoriais e topológicos de intervalos generalizados com alguns conceitos de cálculo e otimização intervalar*. São José do Rio Preto: Tese de doutorado - Universidade Estadual Paulista, 2014.
- [3] T. M. Costa, Y. Chalco-Cano, W. A. Lodwick, and G. N. Silva. Generalized interval vector spaces and interval optimization. *Information Sciences*, 311:74–85, 2015.
- [4] H. Ishibuchi and H. Tanaka. Multiobjective programming in optimization of the interval objective function. *European journal of operational research*, 48(2):219–225, 1990.
- [5] R. E. Moore. *Interval analysis*. Prentice-Hall, Englewood Cliffs, 1966.
- [6] R. E. Moore. *Methods and applications of interval analysis*. Philadelphia: SIAM, 1979.
- [7] R. E. Moore, R. B. Kearfott, and M. J. Cloud. *Introduction to interval analysis*. Philadelphia: SIAM, 2009.
- [8] J. Rohn. Positive definiteness and stability of interval matrices. *SIAM Journal on Matrix Analysis and Applications*, 15(1):175–184, 1994.
- [9] H.-C. Wu. The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function. *European Journal of Operational Research*, 176(1):46–59, 2007.
- [10] H.-C. Wu. On interval-valued nonlinear programming problems. *Journal of Mathematical Analysis and Applications*, 338(1):299–316, 2008.