UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística e Computação Científica

LUCIANO SANTANA BEGOT

Diferentes estratégias para determinar a região de Pareto

Campinas

2025

Luciano Santana Begot

Diferentes estratégias para determinar a região de Pareto

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática Aplicada.

Orientador: Roberto Andreani

Coorientador: Leonardo Delarmelina Secchin

Este trabalho corresponde à versão final da Dissertação defendida pelo aluno Luciano Santana Begot e orientada pelo Prof. Dr. Roberto Andreani.

Campinas

2025

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

B394d

Begot, Luciano Santana, 1997-
Diferentes estratégias para determinar a região de Pareto / Luciano Santana Begot. – Campinas, SP : [s.n.], 2025.

Orientador: Roberto Andreani.
Coorientador: Leonardo Delarmelina Secchin.
Dissertação (mestrado) – Universidade Estadual de Campinas (UNICAMP), Instituto de Matemática, Estatística e Computação Científica.

1. Otimização multiobjetivo. 2. Fronteira de Pareto. 3. Método dos pesos. 4. Método da descida máxima. 5. Método dos pesos com coleta. I. Andreani, Roberto, 1961-. II. Secchin, Leonardo Delarmelina, 1982-. III. Universidade Estadual de Campinas (UNICAMP). Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações complementares

Título em outro idioma: Different strategies for determining the Pareto region
Palavras-chave em inglês:
Multiobjective optimization
Pareto frontier
Weighted method
Steepest descent method
Weighted method with gathering
Área de concentração: Matemática Aplicada
Titulação: Mestre em Matemática Aplicada
Banca examinadora:
Roberto Andreani [Orientador]
María Laura Schuverdt
Leandro da Fonseca Prudente
Data de defesa: 18-02-2025
Programa de Pós-Graduação: Matemática Aplicada

Objetivos de Desenvolvimento Sustentável (ODS)
ODS: 9. Inovação e infraestrutura

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0009-0005-3270-8454
- Currículo Lattes do autor: http://lattes.cnpq.br/3801682744313985

Dissertação de Mestrado defendida em 18 de fevereiro de 2025 e aprovada pela banca examinadora composta pelos Profs. Drs.

Prof. Dr. ROBERTO ANDREANI

Profa. Dra. MARÍA LAURA SCHUVERDT

Prof. Dr. LEANDRO DA FONSECA PRUDENTE

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Agradecimentos

Gostaria de expressar minha profunda gratidão ao meu orientador, Prof. Dr. Roberto Andreani, por sua orientação exemplar, paciência e apoio ao longo desta jornada. Seus conhecimentos e conselhos foram fundamentais para a realização deste trabalho.

Ao meu coorientador, Prof. Dr. Leonardo Secchin, agradeço pela disponibilidade, pelas contribuições valiosas e pela visão crítica que enriqueceram significativamente esta pesquisa. Sua parceria foi essencial para a conclusão deste projeto.

Aos meus pais, agradeço por todo o apoio incondicional, amor e ensinamentos que sempre me guiaram. Vocês foram minha base e minha motivação para trilhar este caminho.

À minha namorada, Fabiane, agradeço pela paciência, compreensão e companheirismo constante. Sua presença ao meu lado foi um alicerce de força e inspiração ao longo deste percurso.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Resumo

A otimização multiobjetivo tem se destacado como uma área de crescente relevância para diversos problemas reais, pois muitos deles envolvem a minimização (ou maximização) simultânea de múltiplos objetivos, muitas vezes conflitantes entre si. Neste trabalho, parte-se da formalização do problema multiobjetivo em espaços de dimensão finita e apresentam-se definições fundamentais de otimalidade de Pareto, incluindo versões fraca e própria, além de discutir o conceito de ponto Pareto crítico. Discutem-se propriedades que impactam a forma como as fronteiras de Pareto podem ser caracterizadas. No que diz respeito aos métodos numéricos, estuda-se inicialmente o método dos pesos, o qual transforma o problema multiobjetivo em um problema de objetivo único, por meio de combinações convexas das funções objetivo. Embora seja eficaz quando o problema é convexo, esse método pode falhar em capturar regiões côncavas da fronteira de Pareto. Em seguida, analisa-se a aplicação do método de descida máxima, onde o conceito de ponto estacionário são adaptados para o contexto multiobjetivo a fim de garantir direções de busca. Como inovação, propõe-se o método dos pesos com coleta, que visa superar as limitações do método dos pesos tradicional. Essa estratégia envolve a coleta de pontos adicionais durante as iterações da busca linear, filtrando-os posteriormente para remover pontos dominados. Tal procedimento se mostra particularmente vantajoso em problemas não convexos, ampliando a exploração de regiões mais complexas da fronteira de Pareto. Em termos computacionais, implementam-se as abordagens citadas em linguagem Julia, adotando estratégias de *multi-start*. Para avaliar o desempenho dos algoritmos, realizam-se testes em um conjunto diversificado de problemas da literatura, variando dimensões, número de objetivos e convexidade do problema. A análise inclui métricas de espalhamento, de pureza (relação com a fronteira verdadeira) e de custo computacional. Os resultados mostram que o método dos pesos com coleta, aliado à verificação de Pareto criticidade, atinge melhor cobertura da fronteira em problemas não convexos, mantendo ainda competitividade em problemas convexos.

Palavras-chave: Otimização Multiobjetivo. Fronteira de Pareto. Método dos Pesos. Descida Máxima. Método dos Pesos com Coleta.

Abstract

Multi-objective optimization has emerged as a field of growing importance for a variety of real-world problems, which often involve the simultaneous minimization (or maximization) of multiple, frequently conflicting objectives. This work begins by formalizing the multi-objective problem in finite-dimensional spaces and presenting fundamental definitions of Pareto optimality, including weak and proper variants, as well as discussing the concept of Pareto critical points. We also examine properties that significantly influence how Pareto fronts can be characterized. Regarding numerical methods, we first study the weighted-sum method, which reduces the multi-objective problem to a single-objective one by taking linear combinations of the objective functions. Although effective for convex problems, it may fail to capture concave regions of the Pareto front. Next, we analyze the application of the steepest descent method, wherein the notion of a stationary point is adapted to the multi-objective setting to guide search directions. As an innovative strategy, we propose the weighted-sum method with gathering, aimed at overcoming the limitations of the traditional weighted-sum method. This approach entails collecting additional points during line search iterations and subsequently filtering out dominated ones. Such a procedure proves particularly beneficial in non-convex problems by enhancing the exploration of more complex regions of the Pareto front. From a computational perspective, the proposed methods are implemented in the Julia programming language, adopting a multi-start scheme. To evaluate the performance of these algorithms, experiments are conducted on a diversified set of benchmark problems, varying in dimensionality, number of objectives, and convexity. The analysis includes metrics of spread, purity (relative to the true front), and computational cost. The results show that the weighted-sum method with gathering, combined with a Pareto criticality check, achieves a broader coverage of the Pareto front in non-convex problems, while remaining competitive in convex scenarios.

Keywords: Multiobjective Optimization. Pareto Front. Weighted-Sum Method. Steepest Descent. Weighted-Sum Method with Gathering.

Lista de ilustrações

Figura 1 – Os conjuntos S e Z e os vetores objetivos Pareto ótimos. 18
Figura 2 – Ilustração do problema do Exemplo 1.2.1 19
Figura 3 – Ilustração do problema do Exemplo 1.4.1 23
Figura 4 – Ilustração do problema do Exemplo 1.4.2. 24
Figura 5 – Figura ilustrando o problema do Exemplo 1.5.1. 26
Figura 6 – Região factível limitada pela circunferência e a caixa, onde a região hachurada interna à circunferência é a região factível do problema do Exemplo 1.5.2. 28
Figura 7 – Método dos pesos com problemas convexos e não convexos. 34
Figura 8 – Distâncias calculadas por $\delta_{i,j}$ 46
Figura 9 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.1. . 48
Figura 10 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.2. . 49
Figura 11 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.3. . 49
Figura 12 – Aproximação da fronteira de Pareto com método dos pesos com coleta para o Exemplo 4.1.2. 50
Figura 13 – Aproximação da fronteira de Pareto com método dos pesos com coleta para o Exemplo 4.1.3. 50
Figura 14 – Aproximação da fronteira de Pareto para o problema MOP 2 para o método dos pesos (esquerda), método dos pesos com coleta utilizando $\lambda = 10^{-3}$ (centro) e o método da descida máxima (direita). . . 54
Figura 15 – Aproximação da fronteira de Pareto para o problema MOP 2 para o método dos pesos (esquerda), método dos pesos com coleta utilizando $\lambda = 10^{-1}$ (centro) e o método da descida máxima (direita). . . 54
Figura 16 – Perfil de desempenho comparando os métodos com base na métrica de tempo. 57
Figura 17 – Perfil de desempenho comparando os métodos com base na métrica de espalhamento Γ. 61
Figura 18 – Perfil de desempenho comparando os métodos com base na métrica de espalhamento Γ para os problemas não convexos. 61

Lista de tabelas

Tabela 1 – Razão de pontos Pareto Crítico. 51
Tabela 2 – Razão de pontos Pareto Crítico após a implementação da restrição. . 53
Tabela 3 – Resultados dos testes de tempo. 56
Tabela 4 – Resultados dos testes. 59
Tabela 5 – Problemas teste (Parte 1) 68
Tabela 6 – Problemas teste (Parte 2) 69

Lista de Algoritmos

1 Método dos Pesos com Gradiente e Backtracking 35
2 Método de Descida Máxima 39
3 Método dos Pesos com Coleta 42
4 Remover Pontos Dominados 42
5 Geração de pesos uniformemente distribuídos no simplex unitário. . . . 44
6 Checagem de Pareto Crítico 51

Sumário

Introdução **12**

1 Fundamentação Teórica **15**

1.1 Formulação do Problema e Notação Geral 15

1.2 Otimalidade de Pareto 17

1.3 Tomada de Decisões 21

1.4 Otimalidade de Pareto Fraca 22

1.5 Otimalidade de Pareto Própria 25

1.6 Dominância 30

2 Métodos **32**

2.1 Método dos Pesos 32

2.2 Método de Descida Máxima 36

2.2.1 Computando a direção de descida 36

2.2.2 Computando o comprimento de passo 38

2.3 Método dos Pesos Com Coleta 41

3 Implementação **43**

3.1 Método dos pesos 43

3.2 Método de descida máxima 44

3.3 Método dos pesos com coleta 45

3.4 Métricas 45

3.4.1 Métrica de Pureza 45

3.4.2 Métrica de Espalhamento 46

4 Testes Numéricos **47**

4.1 Problemas Ilustrativos 47

4.2 Checagem de Pareto Crítico 50

4.2.1 Escolha de λ 53

4.3 Tempo de Execução 55

4.4 Qualidade da Fronteira de Pareto 57

5 Considerações Finais **62**

REFERÊNCIAS **64**

Apêndices **67**

APÊNDICE A Problemas Teste **68**

Introdução

A otimização multiobjetivo é uma área que ganha cada vez mais relevância em diversos campos de aplicação, como engenharia, economia, logística e tomada de decisão em geral, como discutido em [1]. Diferentemente da otimização de objetivo único, em que se busca minimizar (ou maximizar) uma única função, problemas multiobjetivo lidam com múltiplos objetivos que, em muitos casos, são conflitantes entre si. Essa característica confere à otimização multiobjetivo um caráter intrinsecamente mais complexo, pois geralmente não existe uma única solução que seja ótima para todos os objetivos simultaneamente.

Para lidar com esse cenário, conceitos fundamentais como *otimalidade de Pareto* e diferentes critérios de dominância tornam-se indispensáveis. Esses conceitos permitem comparar soluções em termos de *trade-offs* entre os objetivos, estabelecendo bases formais para identificar pontos que sejam, de alguma forma, "equilibrados". No contexto de problemas reais, em que decisões estratégicas demandam uma análise cuidadosa de múltiplos fatores, a existência de uma *fronteira de Pareto* rica e bem distribuída é particularmente útil: ela fornece ao tomador de decisão um conjunto amplo de alternativas potenciais, evidenciando claramente as compensações entre os objetivos.

A fim de facilitar a abordagem de problemas desse tipo, o presente trabalho dedica-se à fundamentação teórica e à análise de métodos numéricos para otimização multiobjetivo em espaços de dimensão finita. Inicia-se pela apresentação rigorosa do problema multiobjetivo e dos conceitos essenciais de Pareto, tais como otimalidade de Pareto, otimalidade de Pareto fraca e otimalidade de Pareto própria. Essas definições fornecem as bases para a comparação de soluções e para o estudo de propriedades como convexidade, que influenciam o comportamento das fronteiras de Pareto.

Em seguida, são introduzidas duas estratégias clássicas para obtenção de pontos Pareto ótimos: (i) o **método dos pesos**, que transforma o problema multiobjetivo em um problema de objetivo único por meio de combinações convexas das funções, e (ii) a **adaptação de métodos de otimização tradicionais**, exemplificada aqui pelo *método de descida máxima*, no qual os conceitos de ponto crítico são reinterpretados no contexto multiobjetivo. Essas duas abordagens são cuidadosamente descritas, com ênfase em suas limitações e potencialidades. Em especial, observa-se que o método dos pesos é bastante eficaz em problemas convexos, garantindo a recuperação de toda a fronteira de Pareto, mas pode falhar em capturar regiões côncavas quando o problema é não convexo. Já o método de descida máxima requer atenção para descartar pontos

de máximo e pontos dominados, que acabam surgindo naturalmente durante a busca por direções de descida.

Para aprimorar a capacidade de explorar regiões mais gerais da fronteira de Pareto (incluindo eventuais partes não convexas), propõe-se o **método dos pesos com coleta**. Nele, pontos adicionais são "coletados" ao longo da busca linear de cada iteração do método de pesos convencional e, ao final, aplica-se um filtro para descartar pontos dominados. Essa estratégia demonstrou, em experimentos numéricos, uma capacidade superior para explorar diferentes regiões da fronteira de Pareto, preenchendo lacunas que permanecem ocultas quando se utiliza apenas uma soma ponderada fixa.

Outro aspecto central do trabalho está na caracterização computacional das soluções. Após encontrar pontos que se apresentam como não dominados, busca-se confirmar se são efetivamente *Pareto críticos*, isto é, se podem ser considerados "ótimos" sob a ótica multiobjetivo. Nesse sentido, o conceito de *ponto Pareto crítico* desempenha papel análogo ao de um "ponto estacionário" em otimização de objetivo único, porém adequado ao contexto multiobjetivo.

Por fim, o texto discute a implementação em linguagem Julia dos métodos apresentados (método dos pesos, método de descida máxima e método dos pesos com coleta), bem como os testes numéricos realizados sobre um amplo conjunto de problemas da literatura. Nessas experimentações, avaliam-se métricas de espalhamento e de pureza, além de indicadores de desempenho computacional (tempo de execução e quantidade de pontos Pareto críticos). Tais resultados evidenciam que a abordagem proposta (método dos pesos com coleta), aliada a um filtro para garantir Pareto criticidade, tem um desempenho competitivo em problemas convexos e se mostra especialmente vantajosa em problemas não convexos, onde a fronteira de Pareto costuma apresentar regiões mais complexas de serem identificadas.

Estruturalmente, este trabalho organiza-se da seguinte forma:

- **Capítulo de Fundamentação Teórica**: apresenta o problema multiobjetivo, formalizando conceitos de Pareto e discutindo convexidade, quasiconvexidade e pseudoconvexidade.
- **Capítulo de Métodos**: introduz e detalha as estratégias do método dos pesos, do método de descida máxima e do método dos pesos com coleta, além das condições de *Pareto criticidade*.
- **Capítulo de Implementação**: descreve o ambiente de desenvolvimento e a forma de implementação em Julia, contemplando estratégias de *multi-start* e regras de busca linear.

- **Capítulo de Resultados e Testes Numéricos**: reúne experimentos computacionais para problemas ilustrativos e de maior porte, analisando métricas de qualidade da fronteira e custo computacional.
- **Considerações Finais**: discute as principais contribuições do trabalho, limitações e perspectivas de pesquisa futura.

Com isso, espera-se que o presente estudo não apenas reforce a importância teórica e prática da otimização multiobjetivo, mas também apresente uma alternativa viável para lidar com problemas complexos, fornecendo ao tomador de decisões um conjunto vasto e diversificado de soluções de alta qualidade.

1 Fundamentação Teórica

1.1 Formulação do Problema e Notação Geral

Iniciamos estabelecendo uma base conceitual e teórica, limitando nosso enfoque a espaços euclidianos de dimensões finitas. Inicialmente, apresentamos a formulação do problema a ser tratado, acompanhada de notações gerais. Na sequência, discutimos diversos conceitos e definições relacionados à otimização multiobjetivo, explorando também as conexões existentes entre eles. É importante destacar que os termos e conceitos utilizados no contexto da otimização multiobjetivo podem variar de autor para autor. Utilizamos [2, 3] como principais referências para este capítulo.

O problema multiobjetivo geral é

$$\begin{aligned} &\text{Minimizar } \{f_1(x), f_2(x), \ldots, f_m(x)\} \\ &\text{sujeito a } x \in S \end{aligned} \tag{1.1}$$

onde temos m ($\geqslant 2$) funções objetivo $f_i : \mathbb{R}^n \to \mathbb{R}$. Denotamos o vetor de funções objetivo por $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))^T$. Os vetores de decisão (variáveis) $x = (x_1, x_2, \ldots, x_n)^T$ pertencem à região factível (conjunto) S, que é um subconjunto do espaço de variáveis de decisão $\mathbb{R}^n$.

O objetivo do problema é minimizar todas as funções objetivo simultaneamente. Caso as funções objetivo não entrassem em conflito, seria possível encontrar uma solução onde todas atingem seus mínimos globais, tornando desnecessário o uso de métodos especializados. Para evitar cenários triviais, supomos que não existe uma única solução ótima para todas as funções objetivo, o que implica que as funções são, pelo menos em parte, conflitantes.

Denotamos a imagem do conjunto viável S pelo vetor

$$Z = f(S),$$

referindo-nos a ele como a região viável no espaço objetivo. Os elementos z de $Z \subseteq \mathbb{R}^m$ são chamados de vetores objetivo, de modo que

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix},$$

onde $z_i = f_i(x)$ para $i = 1, \ldots, m$.

Para simplificar a análise, consideramos que todas as funções objetivo são minimizadas. Caso uma função objetivo f_i precise ser maximizada, isso pode ser tratado minimizando a função $-f_i$.

No que segue, introduziremos algumas definições e notações importantes.

Definição 1.1.1. *Para dois vetores $x, x^* \in \mathbb{R}^n$, dizemos que $x \leqslant x^*$ se, e somente se, $x_i \leqslant x_i^*$ para todo $i = 1, \ldots, n$. Analogamente, $x < x^*$ significa que $x_i < x_i^*$ para todo $i = 1, \ldots, n$.*

Definição 1.1.2. *A soma $\sum_{i=1}^{m} \beta_i x^i$ é chamada de combinação convexa dos vetores $x_1, x_2, \ldots, x_m \in S$ se $\beta_i \geqslant 0$ para todo i e $\sum_{i=1}^{m} \beta_i = 1$.*

Definição 1.1.3. *O fecho convexo de um conjunto $X \subset \mathbb{R}^n$, denotado por* conv X*, é o conjunto de todas as combinações convexas de vetores em X.*

Diferentes categorias de problemas de otimização multiobjetivo podem ser estabelecidas.

Definição 1.1.4. *Um problema de otimização multiobjetivo é classificado como **linear** se todas as funções objetivo e as restrições que definem a região factível forem lineares.*

*Caso pelo menos uma função objetivo ou restrição seja não linear, o problema é denominado problema de otimização multiobjetivo **não linear**.*

Para definirmos problemas de otimização multiobjetivo convexos, apresentamos as definições de funções e conjuntos convexos.

Definição 1.1.5. *Uma função $f_i : \mathbb{R}^n \to \mathbb{R}$ é convexa se, para todos $x^1, x^2 \in \mathbb{R}^n$, a seguinte desigualdade é satisfeita:*

$$f_i(\beta x^1 + (1-\beta)x^2) \leqslant \beta f_i(x^1) + (1-\beta) f_i(x^2),$$

para todo $0 \leqslant \beta \leqslant 1$.

Um conjunto $S \subset \mathbb{R}^n$ é dito convexo se $x^1, x^2 \in S$ implica que

$$\beta x^1 + (1-\beta)x^2 \in S,$$

para todo $0 \leqslant \beta \leqslant 1$.

Definição 1.1.6. *Um problema de otimização multiobjetivo é classificado como **convexo** se todas as funções objetivo e a região factível forem convexas.*

A seguir, introduzimos os conceitos de quasiconvexidade e pseudoconvexidade de funções. Para definir pseudoconvexidade, também é necessário abordar as definições de funções diferenciáveis e continuamente diferenciáveis.

Definição 1.1.7. *Uma função $f_i : \mathbb{R}^n \to \mathbb{R}$ é diferenciável em x^* se*

$$f_i(x^* + d) - f_i(x^*) = \nabla f_i(x^*)^T d + \|d\|\epsilon(x^*, d),$$

onde $\nabla f_i(x^*)$ é o gradiente de f_i em x^* e $\epsilon(x^*, d) \to 0$ quando $\|d\| \to 0$.

Adicionalmente, f_i é continuamente diferenciável em x^* se todas as suas derivadas parciais forem contínuas nesse ponto.

Definição 1.1.8. *Uma função $f_i : \mathbb{R}^n \to \mathbb{R}$ é quasiconvexa se, para todo $x^1, x^2 \in \mathbb{R}^n$ e $0 \leqslant \beta \leqslant 1$, valer:*

$$f_i(\beta x^1 + (1 - \beta)x^2) \leqslant \max[f_i(x^1), f_i(x^2)].$$

Além disso, se f_i for diferenciável em $\mathbb{R}^n$, ela será pseudoconvexa se, para todo $x^1, x^2 \in \mathbb{R}^n$ tal que $\nabla f_i(x^1)^T(x^2 - x^1) \geqslant 0$, a condição $f_i(x^2) \geqslant f_i(x^1)$ for satisfeita.

Note que toda função pseudoconvexa também é quasiconvexa.

Definição 1.1.9. *Um problema de otimização multiobjetivo é dito **não diferenciável** se pelo menos uma das funções objetivo ou uma das restrições que definem a região factível não forem diferenciáveis.*

1.2 Otimalidade de Pareto

O conceito de otimalidade desempenha um papel central na teoria da otimização. Nesta seção, abordaremos como este conceito se molda aos problemas multiobjetivo. Em problemas com um único objetivo, o foco está quase sempre no espaço das variáveis de decisão. No entanto, em otimização multiobjetivo, o espaço dos objetivos frequentemente assume maior relevância. Isso ocorre porque, além de geralmente ter uma dimensão inferior ao espaço das variáveis de decisão, ele também serve como base para as definições de otimalidade utilizadas nesse contexto.

A interação entre múltiplos objetivos, que podem ser conflitantes, impede a existência de uma única solução que satisfaça simultaneamente todos eles de forma ideal. Isso confere aos problemas de otimização multiobjetivo uma característica intrinsecamente ambígua, já que o espaço dos objetivos não possui uma hierarquia clara. De fato, ele é apenas parcialmente ordenado, o que significa que podemos comparar vetores como $(0, 0)^T$ e $(2, 2)^T$, mas não é evidente como ordenar $(0, 2)^T$ em relação a $(2, 1)^T$.

Ainda assim, é possível identificar vetores objetivos que se destacam por uma propriedade específica: nenhuma de suas componentes pode ser melhorada sem que pelo menos uma outra seja comprometida. Esses vetores representam soluções

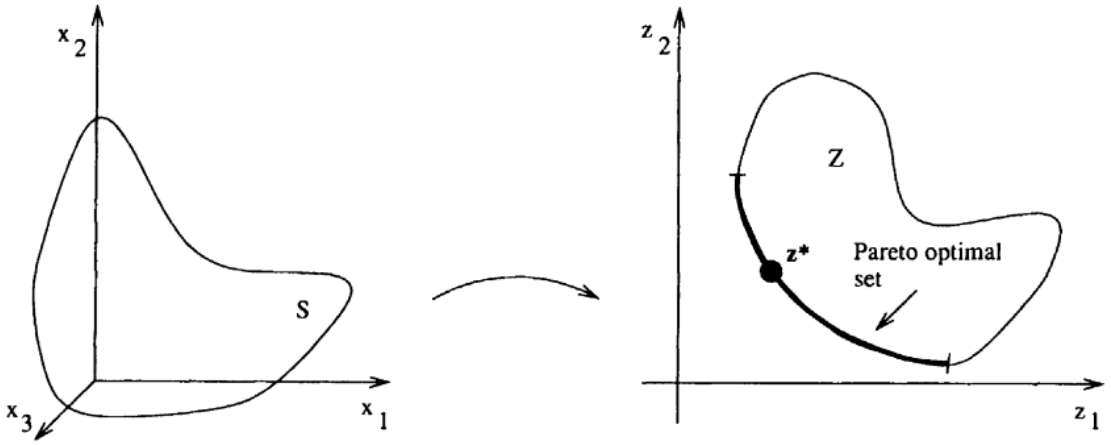

Figura 1 – Os conjuntos S e Z e os vetores objetivos Pareto ótimos.

Fonte: Kaisa [2].

conhecidas como pontos ótimos de Pareto, em homenagem a Vilfred Pareto, economista e sociólogo franco-italiano.

A seguir, apresentamos uma definição mais rigorosa de otimalidade de Pareto:

Definição 1.2.1. *Um vetor de decisão $x^* \in S$ é Pareto ótimo se não existe outro vetor de decisão $x \in S$ tal que $f_i(x) \leqslant f_i(x^*)$ para todo $i = 1, \cdots, m$ e $f_j(x) < f_j(x^*)$ para pelo menos um índice j.*

Um vetor objetivo $z^* \in Z$ é Pareto ótimo se não existe outro vetor objetivo $z \in Z$ tal que $z_i \leqslant z_i^*$ para todo $i = 1, \cdots, m$ e $z_j < z_j^*$ para pelo menos um índice j. Equivalentemente, z^* é Pareto ótimo se o seu vetor de decisão correspondente é Pareto ótimo.

A definição de otimalidade de Pareto implica que transitar de uma solução Pareto ótima para outra exige realizar compensações (trade-off) entre os objetivos conflitantes, caracterizando um princípio central da otimização multiobjetivo.

A Figura 1 ilustra uma representação gráfica da região factível $S \subset \mathbb{R}^3$ e de sua imagem, a região objetiva factível $Z \subset \mathbb{R}^2$. A linha destacada na figura representa os vetores objetivos que são Pareto ótimos, e z^* é mostrado como um exemplo de vetor objetivo com tal propriedade.

Exemplo 1.2.1. *Sejam $f_1 = (x+1)^2$ e $f_2 = (x-1)^2$. Consideremos os seguinte problema:*

$$\textit{Minimizar } \{f_1(x), f_2(x)\}$$
$$\textit{sujeito a } x \in \mathbb{R}$$

Graficamente, o problema corresponde a:

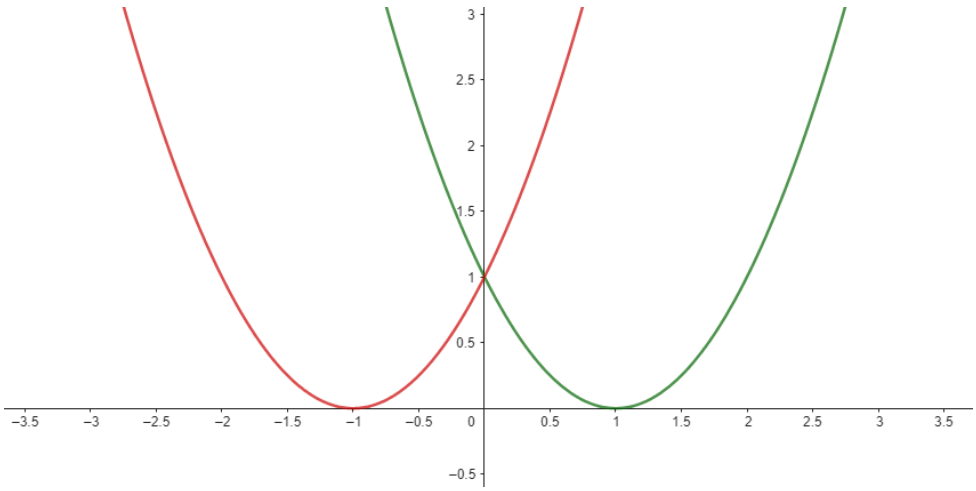

Figura 2 – Ilustração do problema do Exemplo 1.2.1

Análise dos Pontos:

- *Para* $x = -1$*:*
  - $f_1(-1) = 0$ *e* $f_2(-1) = 4$.
  - *Este ponto é Pareto ótimo, pois:*
    - *Não existe outro* x *tal que* $f_1(x) < f_1(-1)$ *(já que* $f_1(-1)$ *é o mínimo global de* $f_1(x)$*).*
    - *Qualquer tentativa de melhorar* $f_2(x)$ *(diminuir seu valor) levaria a um aumento de* $f_1(x)$*.*
- *Para* $x = 1$*:*
  - $f_1(1) = 4$ *e* $f_2(1) = 0$.
  - *Este ponto também é Pareto ótimo, pois:*
    - *Não existe outro* x *tal que* $f_2(x) < f_2(1)$ *(já que* $f_2(1)$ *é o mínimo global de* $f_2(x)$*).*
    - *Qualquer tentativa de melhorar* $f_1(x)$ *levaria a um aumento de* $f_2(x)$*.*
- *Para* $x \in [-1, 1]$*:*
  - *Para qualquer ponto neste intervalo, os valores de* $f_1(x)$ *e* $f_2(x)$ *são ambos menores ou iguais a 4.*
  - *Não existe outro* $x \in \mathbb{R}$ *que melhore* $f_1(x)$ *ou* $f_2(x)$ *sem piorar o outro, pois:*
    - *Quando* x *se aproxima de* -1*,* $f_1(x)$ *melhora, mas* $f_2(x)$ *piora.*

  - *Quando x se aproxima de 1, $f_2(x)$ melhora, mas $f_1(x)$ piora.*
- *Portanto, todos os pontos $x \in [-1, 1]$ são Pareto ótimos.*

Portanto, o conjunto de soluções Pareto ótimas corresponde ao intervalo $x \in [-1, 1]$.

Geralmente, encontramos uma multiplicidade de soluções Pareto ótimas, muitas vezes em quantidade infinita. Um conjunto dessas soluções é chamado de conjunto Pareto ótimo, o qual pode ser não convexo ou até mesmo não conexo.

Outro conceito fundamental em otimização multiobjetivo é a otimalidade de Pareto local, definida formalmente a seguir:

Definição 1.2.2. *Dizemos que um vetor de decisão $x^* \in S$ é localmente Pareto ótimo se existe um $\delta > 0$ tal que x^* é Pareto ótimo no conjunto $S \cap B(x^*, \delta)$, onde $B(x^*, \delta)$ representa uma bola aberta de raio δ centrada em x^*.*

De maneira análoga, um vetor objetivo $z^* \in Z$ é localmente Pareto ótimo se o vetor de decisão correspondente a ele for localmente Pareto ótimo.

Vale notar que toda solução Pareto ótima global também é localmente Pareto ótima. A recíproca é garantida no contexto de problemas de otimização multiobjetivo convexos, conforme o teorema a seguir:

Teorema 1.2.1. *Em um problema de otimização multiobjetivo convexo, todo ponto localmente Pareto ótimo é, necessariamente, globalmente Pareto ótimo.*

Demonstração. Seja $x \in x^*$ localmente Pareto ótimo. Então existe algum $\delta > 0$ e uma vizinhança $B(x^*, \delta)$ de x^* tal que não existe $x \in S \cap B(x^*, \delta)$ para o qual $f_i(x) \leqslant f_i(x^*)$ para todo $i = 1, \cdots, m$ e, para pelo menos um índice j, $f_j(x) < f_j(x^*)$.

Vamos supor que x^* não é globalmente Pareto ótimo. Neste caso, existe algum outro ponto $x^0 \in S$ tal que

$$f_i(x^0) \leqslant f_i(x^*) \text{ para todo } i = 1, \cdots, k \text{ e } f_j(x^0) < f_j(x^*) \text{ para algum } j. \tag{1.2}$$

Definimos $\hat{x} = \beta x^0 + (1 - \beta)x^*$, onde $0 < \beta < 1$ é selecionado de modo que $\hat{x} \in B(x^*, \delta)$. A convexidade de S implica que $\hat{x} \in S$. Pela convexidade das funções objetivo e aplicando(1.2), obtemos

$$f_i(\hat{x}) \leqslant \beta f_i(x^0) + (1 - \beta) f_i(x^*) \leqslant \beta f_i(x^*) + (1 - \beta) f_i(x^*) = f_i(x^*) \text{ para todo } i = 1, \cdots, k.$$

Como x^* é localmente Pareto ótimo e $\hat{x} \in B(x^*, \delta)$, devemos ter $f_i(\hat{x}) = f_i(x^*)$ para todo i.

Assim, $f_i(x^*) \leqslant \beta f_i(x^0) + (1-\beta) f_i(x^*)$ para todo $i = 1, \cdots, m$. Como $\beta > 0$, podemos dividir por β e obter $f_i(x^*) \leqslant f_i(x^0)$ para todo i. De (1.2) temos $f_j(x^*) > f_j(x^0)$ para algum j. Isto é uma contradição. Portanto, x^* é globalmente Pareto ótimo. □

Teorema 1.2.2. *Considere um problema de otimização multiobjetivo onde a região viável é convexa e as funções objetivo são quasiconvexas, com pelo menos uma delas sendo estritamente quasiconvexa. Nesse caso, qualquer solução que seja localmente Pareto ótima também será globalmente Pareto ótima.*

Demonstração. Seja $x^* \in S$ localmente ótimo de Pareto. Assim, existe algum $\delta > 0$ e uma vizinhança $B(x^*, \delta)$ de x^* tal que não existe $x \in S \cap B(x^*, \delta)$ para o qual $f_i(x) < f_i(x^*)$ para todo $i = 1, \ldots, k$ e, para pelo menos um índice j, $f_j(x) < f_j(x^*)$. Suponha que x^* não seja globalmente ótimo de Pareto. Nesse caso, existe algum outro ponto $x^o \in S$ tal que

$$f_i(x^o) \leqslant f_i(x^*), \quad \text{para todo } i = 1, \ldots, m, \quad \text{e } f_j(x^o) < f_j(x^*), \text{ para algum } j. \tag{1.3}$$

Definimos $\hat{x} = \beta x^o + (1-\beta) x^*$, onde $0 < \beta < 1$ é selecionado tal que $\hat{x} \in B(x^*, \delta)$. A convexidade de S implica que $\hat{x} \in S$. Utilizando (1.3) e a quasiconvexidade das funções objetivo, para cada índice i tal que $f_i(x^o) = f_i(x^*)$, obtemos

$$f_i(\hat{x}) \leqslant \max\{f_i(x^o), f_i(x^*)\} = f_i(x^*),$$

e para cada índice j tal que $f_j(x^o) < f_j(x^*)$, temos

$$f_j(\hat{x}) \leqslant \max\{f_j(x^o), f_j(x^*)\} = f_j(x^*).$$

Como pelo menos uma das funções objetivo é estritamente quasiconvexa, pelo menos uma das desigualdades acima é estrita. Aqui temos uma contradição com a otimalidade local de Pareto de x^*. Assim, x^* é globalmente ótimo de Pareto. □

Na prática, as soluções encontradas computacionalmente geralmente correspondem a pontos localmente Pareto ótimos, a menos que condições adicionais, como a convexidade, estejam presentes.

1.3 Tomada de Decisões

Do ponto de vista matemático, cada ponto Pareto ótimo é uma solução válida para um problema de otimização multiobjetivo. Contudo, na prática, muitas vezes é necessário selecionar uma única solução. Essa escolha entre as várias opções do conjunto Pareto ótimo depende de informações adicionais que não estão contidas

nas funções objetivo. Assim, em comparação à otimização de objetivo único, um novo componente é introduzido na otimização multiobjetivo.

Essa nova dimensão envolve a participação de um tomador de decisões. Esse agente pode ser uma pessoa ou um grupo, responsável por compreender o problema de forma mais ampla e expressar preferências entre diferentes soluções possíveis. É geralmente atribuído ao tomador de decisões o papel de determinar a solução final.

A resolução de problemas de otimização multiobjetivo exige colaboração entre o tomador de decisões e um analista. O analista, que pode ser uma pessoa ou um software, cuida da formulação e solução matemática do problema, fornecendo informações que orientam o tomador de decisões. A escolha final da solução é feita com base nas preferências expressas pelo tomador de decisões.

Neste trabalho, não consideramos a etapa de tomada de decisão, ou seja, não assumimos a presença de um tomador de decisões. Nosso foco está em calcular o maior número possível de pontos Pareto ótimos, levando em consideração sempre a qualidade da fronteira de Pareto gerada. A introdução de um tomador de decisões implicaria na existência de uma função que represente suas preferências e selecione um ponto específico dentro da fronteira de Pareto calculada. Problemas que envolvem essa consideração são conhecidos como problemas de dois níveis, como discutido, por exemplo, em [4].

Apesar de não incorporar preferências explícitas de um tomador de decisões, destacamos que a abordagem proposta neste trabalho é altamente relevante. Ao produzir uma fronteira de Pareto densa e de alta qualidade, fornecemos uma maior gama de opções para um eventual tomador de decisões no futuro, aumentando a utilidade prática dos resultados gerados.

1.4 Otimalidade de Pareto Fraca

Além da otimalidade de Pareto, outros conceitos associados são amplamente utilizados, incluindo a otimalidade de Pareto fraca e a otimalidade de Pareto própria. Esses conceitos estão inter-relacionados: o conjunto de soluções propriamente Pareto ótimas é um subconjunto do conjunto de soluções Pareto ótimas, que, por sua vez, é um subconjunto do conjunto de soluções Pareto ótimas fracas.

Definimos que um vetor é Pareto ótimo fraco se não houver outro vetor que melhore todas as suas componentes simultaneamente. Formalmente, isso é descrito da seguinte forma:

Definição 1.4.1. *Dizemos que um vetor de decisão $x^* \in S$ é Pareto ótimo fraco se não existir outro vetor de decisão $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todos os índices $i = 1, \cdots, m$.*

Analogamente, um vetor objetivo $z^* \in Z$ é considerado Pareto ótimo fraco se não houver outro vetor objetivo $z \in Z$ tal que $z_i < z_i^*$ para todos os índices $i = 1, \cdots, m$. Equivalentemente, essa condição é satisfeita se o vetor de decisão associado for Pareto ótimo fraco.

Assim como ocorre com a otimalidade de Pareto, também podemos distinguir entre otimalidade fraca de Pareto global e local. Na prática, as soluções computacionalmente acessíveis geralmente são localmente Pareto ótimas fracas. Por simplicidade, é comum referir-se a pontos localmente Pareto ótimos fracos apenas como pontos Pareto ótimos fracos.

Exemplo 1.4.1. *Sejam $f_1 = (x+1)^2$, $f_2 = (x-1)^2$ e $f_3 = 1$. Consideremos os seguinte problema:*

$$\begin{aligned} &\textit{Minimizar } \{f_1(x), f_2(x), f_3(x)\} \\ &\textit{sujeito a } x \in \mathbb{R} \end{aligned}$$

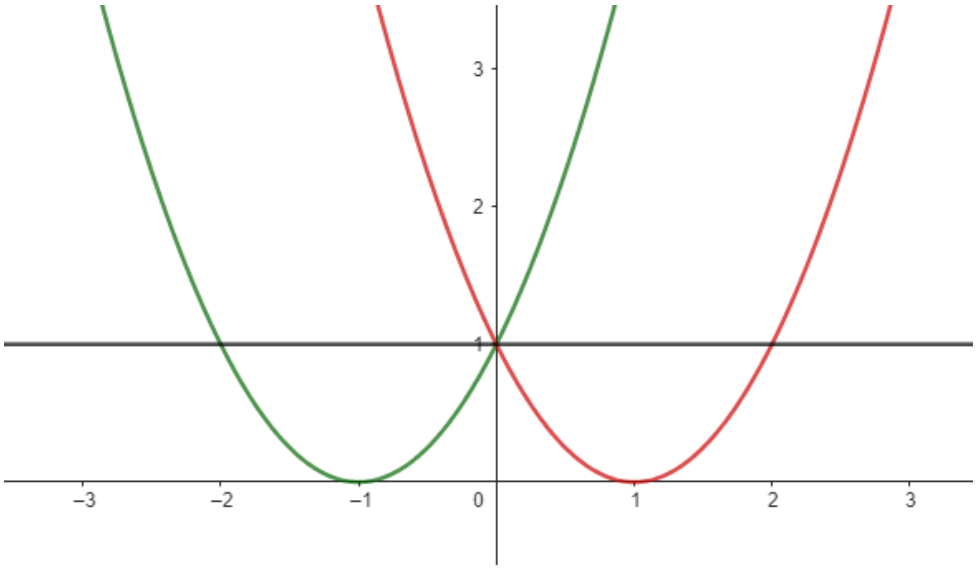

Figura 3 – Ilustração do problema do Exemplo 1.4.1

Análise de Otimalidade de Pareto:

- *Assim como no exemplo anterior:*
  - *Para $x = -1$, temos $f_1(-1) = 0$, $f_2(-1) = 4$ e $f_3(-1) = 1$. Este ponto é Pareto ótimo, pois não é possível melhorar $f_1(x)$ sem piorar $f_2(x)$ ou $f_3(x)$.*
  - *Para $x = 1$, temos $f_1(1) = 4$, $f_2(1) = 0$ e $f_3(1) = 1$. Este ponto também é Pareto ótimo pelo mesmo raciocínio.*

  - *Para $x \in [-1, 1]$, qualquer ponto nesse intervalo caracteriza um equilíbrio entre $f_1(x)$ e $f_2(x)$, mantendo $f_3(x)$ constante. Portanto, todos os pontos $x \in [-1, 1]$ são Pareto ótimos.*

- *No entanto, ao introduzirmos $f_3(x)$ como uma constante, temos uma nova observação:*
  - *Para qualquer $x \in \mathbb{R}$, não é possível reduzir simultaneamente $f_1(x)$, $f_2(x)$ $f_3(x)$;*
  - *Como $f_3(x)$ não pode ser reduzido (é constante em 1 para todo x), qualquer ponto $x \in \mathbb{R}$ é **Pareto ótimo fraco**.*

Logo, apesar do conjunto Pareto ótimo se manter inalterado quando comparado ao Exemplo 1.2.1, o conjunto dos pontos Pareto ótimo fracos é diferente e corresponde a todos os números reais.

Exemplo 1.4.2. *Considere $f_1 = x$ e $f_2 = y$ e o seguinte problema:*

$$\begin{aligned} &\text{Minimizar } \{f_1(x), f_2(x)\} \\ &\text{sujeito a } 0 < x < 1 \\ &\qquad 0 \leqslant y \leqslant 1 \end{aligned}$$

Este problema equivale a minimizar os valores de x e y dentro da caixa aberta, como mostra a Figura 4:

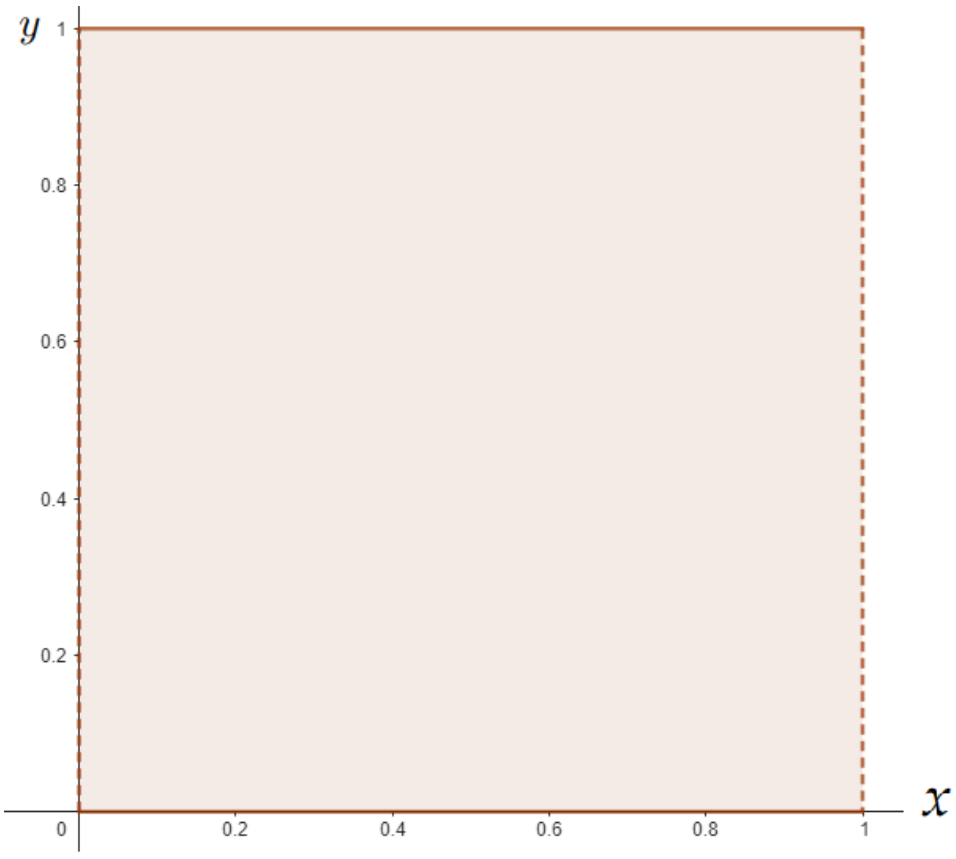

Figura 4 – Ilustração do problema do Exemplo 1.4.2.

Análise de Otimalidade de Pareto:

- *Não existem pontos Pareto ótimos no sentido estrito para este problema:*
  - *Para qualquer ponto (x, y) dentro da região viável, é possível encontrar outro ponto com $x' < x$ e $y' \leqslant y$, diminuindo x enquanto y é mantido.*
  - *Dessa forma, nenhum ponto pode ser considerado Pareto ótimo, já que todos são dominados por outro ponto na direção de x.*
- *No entanto, há pontos Pareto ótimos fracos:*
  - *Os pontos pertencentes à "borda inferior"da caixa, $(0, 1) \times \{0\}$, são Pareto ótimos fracos.*
  - *Para qualquer ponto (x, y) com $y = 0$ e $0 < x < 1$, não é possível melhorar y (reduzir abaixo de 0). Portanto, mesmo que seja possível diminuir x, esses pontos são Pareto ótimos fracos.*

Logo, neste exemplo não temos uma região Pareto ótima, mas temos uma região Pareto ótima fraca.

Embora tenham valor teórico, as soluções fracamente Pareto ótimas podem abranger um conjunto extenso demais, onde algumas soluções podem carecer de critérios mais rigorosos. Apesar disso, sua importância técnica não pode ser ignorada, já que frequentemente são mais simples de calcular ou identificar do que os pontos estritamente Pareto ótimos.

Por outro lado, a otimalidade de Pareto própria oferece uma abordagem mais restritiva e criteriosa, que pode ser mais adequada para contextos em que um conjunto mais refinado de soluções é desejado.

1.5 Otimalidade de Pareto Própria

Kuhn e Tucker, em [5], foram pioneiros ao notar que algumas soluções Pareto ótimas apresentavam características indesejáveis: era necessário um aumento muito grande em um objetivo para se conseguir um decréscimo pequeno em outro. Para contornar isto, eles propuseram a introdução de soluções propriamente Pareto ótimas e sugeriram a distinção entre soluções Pareto ótimas próprias e impróprias. A essência da otimalidade de Pareto própria é a restrição contra compensações ilimitadas entre os objetivos.

Existem várias interpretações formais para a otimalidade de Pareto própria. A definição mais intuitiva é a seguinte, apresentada em [6]:

Definição 1.5.1. *Um vetor de decisão $x^* \in S$ é propriamente Pareto ótimo se este é Pareto ótimo e se existe algum número real $M > 0$ tal que para cada f_i e cada $x \in S$ satisfazendo $f_i(x) < f_i(x^*)$, existe pelo menos um f_j tal que $f_j(x^*) < f_j(x)$ e*

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} \leqslant M.$$

Um vetor objetivo $z^* \in Z$ é propriamente Pareto ótimo se o seu vetor de decisão correspondente é propriamente Pareto ótimo.

De forma simplificada, uma solução é considerada propriamente Pareto ótima quando, para pelo menos um par de objetivos, uma redução pequena em um objetivo só é alcançada à custa de um aumento relativamente grande no outro objetivo.

Exemplo 1.5.1. *Considere $f_1 = -x^2$ e $f_2 = x^3$ e o seguinte problema:*

$$\begin{gathered} \textit{Minimizar } \{f_1(x), f_2(x)\} \\ \textit{sujeito a } x \geqslant 0 \end{gathered}$$

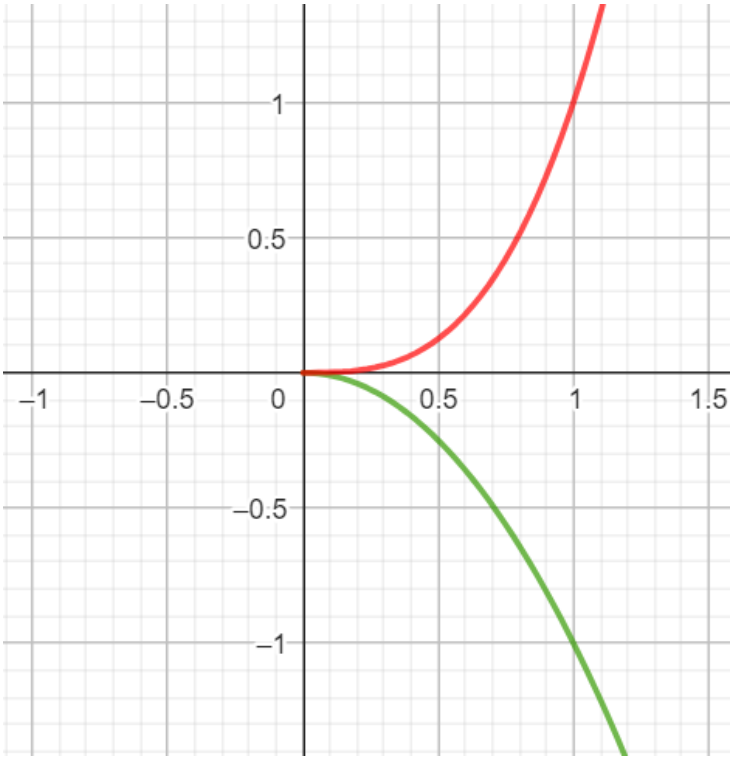

Figura 5 – Figura ilustrando o problema do Exemplo 1.5.1.

Análise de Otimalidade de Pareto:

- *Para $x \geqslant 0$, todos os pontos são Pareto ótimos fracos. Isso ocorre porque, para valores positivos de x, existe um equilíbrio entre $f_1(x)$ e $f_2(x)$, mas qualquer pequeno aumento de x pode resultar em grandes mudanças em $f_1(x)$ e em $f_2(x)$.*

- *No entanto, ao nos aproximarmos de $x = 0$ por valores de x positivos, o ganho em $f_1(x)$ pode se tornar arbitrariamente grande em relação à perda em $f_2(x)$. Portanto, o ponto $x_0 = 0$ não é propriamente Pareto ótimo.*

Cálculo da Razão entre as Funções:

- *Tomando $x^\epsilon = \epsilon$, onde $0 < \epsilon < 1$, temos:*

$$\frac{f_1(x^*) - f_1(x^\epsilon)}{f_2(x^\epsilon) - f_2(x^*)} = \frac{0 - (-\epsilon^2)}{\epsilon^3} = \frac{1}{\epsilon}.$$

- *Esta fração tende ao infinito conforme ϵ tende a 0, indicando que $x_0 = 0$ não é propriamente Pareto ótimo.*

- *De fato, para qualquer $M > 0$, existe um índice $i \in \{1, 2\}$ e um ponto $x \in S$ tal que:*

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} > M,$$

para todo $j \in \{1, 2\}$, onde $f_j(x) > f_j(x^*)$.

Logo, apesar de todos os pontos do domínio serem Pareto ótimos, o ponto $x_0 = 0$ não é propriamente Pareto ótimo, pois a relação entre $f_1(x)$ e $f_2(x)$ se torna arbitrariamente grande conforme x se aproxima de 0.

Exemplo 1.5.2. *Considere o seguinte problema de otimização multiobjetivo:*

$$\begin{aligned} &\text{Minimizar } \{f_1(x), f_2(x)\}, \\ &\text{sujeito a } (x_1 - 1)^2 + (x_2 - 1)^2 \leqslant 1, \\ &0 \leqslant x_1 \leqslant 1, \\ &0 \leqslant x_2 \leqslant 1, \end{aligned}$$

onde $f_1(x) = x_1$ e $f_2(x) = x_2$.

O problema consiste em minimizar os valores de x_1 e x_2 dentro de uma circunferência de raio 1 centrada em $(1, 1)$, limitada à região de uma caixa definida pelas restrições $0 \leqslant x_1 \leqslant 1$ e $0 \leqslant x_2 \leqslant 1$, conforme ilustra a Figura 6.

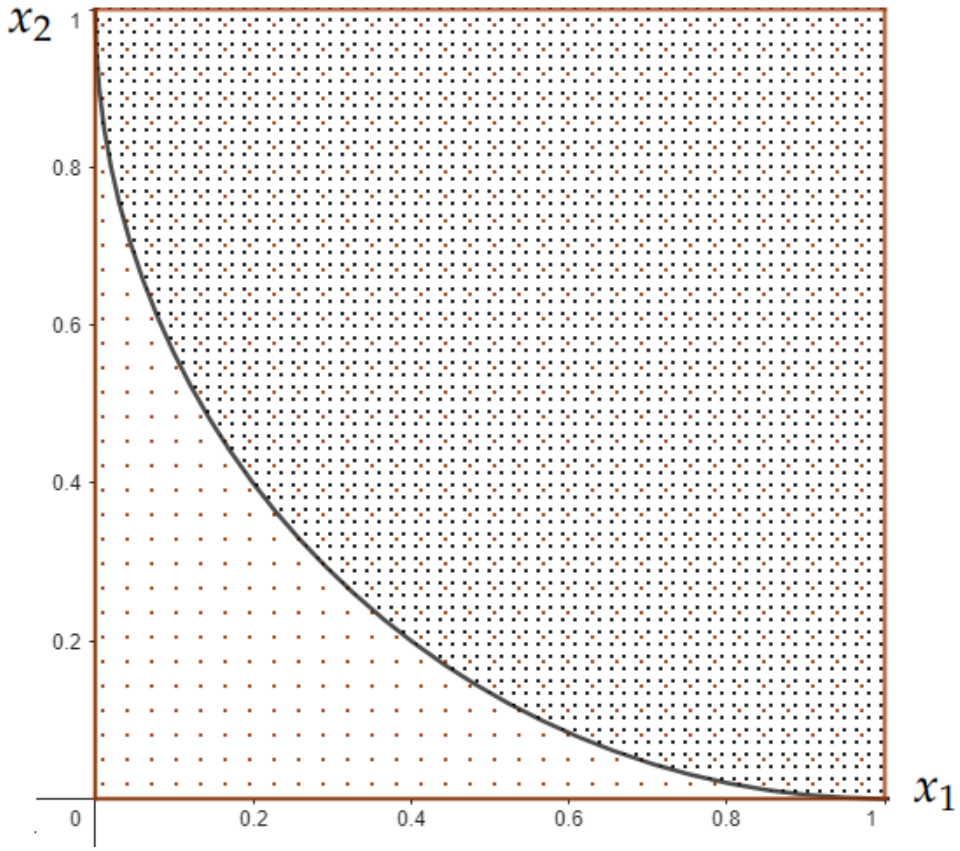

Figura 6 – Região factível limitada pela circunferência e a caixa, onde a região hachurada interna à circunferência é a região factível do problema do Exemplo 1.5.2.

Análise de Otimalidade de Pareto:

- *Os pontos na borda da circunferência que pertencem à região factível, ou seja,* $(x_1, x_2) \in \mathbb{R}^2$ *tal que* $(x_1 - 1)^2 + (x_2 - 1)^2 = 1$*, são todos pontos ótimos de Pareto.*
- *Contudo, ao nos aproximarmos do ponto* $x^* = (1, 0)$*, o equilíbrio entre* x_1 *e* x_2 *se altera. Para se obter um decréscimo suficientemente grande em* x_2*, é necessário um incremento cada vez maior em* x_1*. No limite, é necessário um incremento infinito em* x_1 *para reduzir* x_2 *de maneira significativa.*
- *Logo, o ponto* $x^* = (1, 0)$ *não é propriamente Pareto ótimo.*

Cálculo da Razão entre as Funções:

- *Precisamos demonstrar que, para qualquer* $M > 0$*, existe um índice* $i \in \{1, 2\}$ *e algum* $x \in S$ *tal que* $f_i(x) < f_i(x^*)$ *e:*

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} > M,$$

para todo $j \in \{1, 2\}$*, onde* $f_j(x) > f_j(x^*)$*.*

- *Seja $i = 1$. Consideramos x^ϵ definido por:*

$$x_1^\epsilon = 1 - \epsilon,$$
$$x_2^\epsilon = 1 - \sqrt{1 - \epsilon^2},$$

  onde $0 < \epsilon < 1$.

- *Sabemos que x^ϵ é ponto Pareto ótimo, pois:*

$$(x_1^\epsilon - 1)^2 + (x_2^\epsilon - 1)^2 = \epsilon^2 + (1 - \epsilon^2) = 1.$$

- *Note que $x^\epsilon \in S$, $x_1^\epsilon < x_1^*$ e $x_2^\epsilon > x_2^*$.*

- *Calculando a razão, temos:*

$$\frac{f_1(x^*) - f_1(x^\epsilon)}{f_2(x^\epsilon) - f_2(x^*)} = \frac{1 - (1 - \epsilon)}{1 - \sqrt{1 - \epsilon^2}} = \frac{\epsilon}{1 - \sqrt{1 - \epsilon^2}}.$$

- *Simplificando:*

$$\frac{\epsilon}{1 - \sqrt{1 - \epsilon^2}} = \frac{1 + \sqrt{1 - \epsilon^2}}{\epsilon}.$$

- *Essa fração tende ao infinito conforme $\epsilon \to 0$.*

Assim, vemos que ponto $x^* = (1, 0)$ não é propriamente Pareto ótimo, pois, ao nos aproximarmos de x^*, a relação entre o ganho em f_1 e a perda em f_2 torna-se arbitrariamente grande. Todos os outros pontos na borda da circunferência que pertencem à região factível são propriamente Pareto ótimos, exceto $(0, 1)$.

De maneira análoga, é possível provar que $(0, 1)$ também não é propriamente Pareto ótimo.

Introduziremos também a definição de propriamente Pareto ótimo no sentido de Kuhn e Tucker. Para isto, vamos supor que a região factível seja descrita apenas por restrições de desigualdade. Isto é, definimos

$$S = \Big\{ \mathbf{x} \in \mathbb{R}^n \; g(\mathbf{x}) = \big(g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x})\big)^T \leqslant 0 \Big\}.$$

Além disso, postulamos que tanto as funções objetivo quanto as funções de restrição são continuamente diferenciáveis em cada ponto $\mathbf{x} \in S$. Esta definição é aplicável apenas a problemas diferenciáveis.

Definição 1.5.2. *(De [5]) Um vetor de decisão $x^* \in S$ é propriamente Pareto ótimo se este é Pareto ótimo e não existe outro vetor $d \in \mathbb{R}^n$ tal que*

$$\nabla f_i(x^*)^T d \leqslant 0$$

para todo $i = 1, \cdots, m$ *e para algum* j *temos*

$$\nabla f_j(x^*)^T d < 0,$$

e

$$\nabla g_l(x^*)^T d \leqslant 0$$

para todo l *satisfazendo* $g_l(x^*) = 0$*, ou seja, para toda restrição ativa em* x^**.*

Um vetor objetivo $z^* \in Z$ *é propriamente Pareto ótimo se o vetor de decisão correspondente a ele é propriamente Pareto ótimo.*

Kuhn e Tucker também derivaram condições necessárias e suficientes para otimalidade de Pareto própria.

1.6 Dominância

Na otimização multiobjetivo frequentemente utilizamos o conceito de **dominância** para descrever relações entre soluções potenciais. Esse conceito oferece uma base formal para comparar vetores objetivos e decidir quais representam melhores soluções.

Definição 1.6.1. *Dizemos que um vetor* x_1 *é* ***dominado*** *por outro vetor* x_2 *se, para todas as funções objetivo* f_i*, temos* $f_i(x_2) \leqslant f_i(x_1)$ *e, além disso, existe pelo menos uma função* f_j *tal que* $f_j(x_2) < f_j(x_1)$*.*

Em outras palavras, um vetor x_2 domina um vetor x_1 se x_2 é, no mínimo, tão bom quanto x_1 em todos os objetivos, e estritamente melhor em pelo menos um deles. A ideia central é que, ao comparar duas soluções possíveis, a dominância permite identificar qual delas é mais vantajosa considerando todos os objetivos simultaneamente.

Considere, por exemplo, um problema de minimização com dois objetivos, no qual temos os vetores objetivos $z^1 = (3, 4)$ e $z^2 = (2, 5)$. Para determinar a relação de dominância entre esses vetores, verificamos as condições necessárias. O vetor z^1 dominaria z^2 caso todas as suas componentes fossem menores ou iguais às de z^2, com pelo menos uma delas sendo estritamente menor. No entanto, observamos que a primeira componente de z^1 é maior que a correspondente de z^2 (3 não é menor ou igual a 2). Assim, z^1 não domina z^2.

Por outro lado, consideremos um caso diferente, onde os vetores objetivos sejam $z^1 = (2, 4)$ e $z^2 = (3, 5)$. Neste cenário, ao comparar os dois vetores, percebemos que z^1 é inferior em todas as componentes em relação a z^2. Portanto, neste caso, z^1 domina z^2.

Assim, podemos definir pontos Pareto a partir do conceito de dominância:

Definição 1.6.2. *Um vetor de decisão $x^* \in S$ é chamado de **Pareto ótimo** se não existir nenhum outro vetor de decisão $x \in S$ tal que $f(x)$ domine $f(x^*)$. Em outras palavras, x^* é Pareto ótimo se não houver $x \in S$ para o qual $f_i(x) \leqslant f_i(x^*)$ para todo $i = 1, \ldots, m$, com pelo menos uma dessas desigualdades sendo estrita.*

Portanto, um ponto Pareto ótimo pode ser entendido como um ponto que não é dominado por nenhum outro do domínio.

2 Métodos

Neste capítulo, iremos incialmente apresentar duas estratégias para encontrar candidatos a pontos Pareto ótimos.

A primeira estratégia, apresentada inicialmente em trabalhos como [7, 8], consiste em reescrever (1.1) como um problema de único objetivo, em que minimiza-se uma soma ponderada dos objetivos. Decorre disto o que denominamos aqui de "método dos pesos".

A segunda consiste na adaptação de métodos usuais de otimização padrão [9, 10, 11, 12, 13, 14, 15, 16, 17, 18], dos quais destacamos o método de máxima descida descrito em [19]; nessas estratégias, o cálculo de direções e buscas lineares são adaptados tendo em mente o conceito de Pareto crítico.

Em seguida, apresentamos um método de elaboração própria, baseado no método dos pesos, ao qual denominamos método dos pesos com coleta.

2.1 Método dos Pesos

No método dos pesos abordamos o problema (1.1) associando cada função objetivo a um coeficiente de peso e minimizamos a soma ponderada dos objetivos. Desta forma, as múltiplas funções objetivo são transformadas em uma única função. Supomos que os coeficientes ω_i são números reais tais que $\omega_i \geqslant 0$ para todo $i = 1, \cdots, m$. Também é comum supormos que os pesos são normalizados, ou seja, $\sum_{i=1}^{m} \omega_i = 1$. Assim, o problema de otimização multiobjetivo é reescrito como o seguinte problema:

$$\begin{aligned} &\text{Minimizar } F_\omega(x) = \sum_{i=1}^{m} \omega_i f_i(x) \\ &\text{sujeito a } x \in S, \end{aligned} \tag{2.1}$$

onde $\omega_i \geqslant 0$ para todo i e $\sum_{i=1}^{m} \omega_i = 1$.

A seguir, vamos apresentar diversos resultados a respeito do método dos pesos.

Teorema 2.1.1. *Um minimizador global do problema (2.1) é Pareto ótimo fraco.*

Demonstração. Seja $x^* \in S$ um minimizador global do problema de pesos. Suponhamos que não seja Pareto ótimo fraco. Neste caso, existe uma solução $x \in S$ tal que

$f_i(x) < f_i(x^*)$ para todo $i = 1, \ldots, m$. De acordo com as suposições definidas para os coeficientes de peso, $w_j > 0$ para pelo menos um j. Assim, temos

$$\sum_{i=1}^{m} \omega_i f_i(x) < \sum_{i=1}^{m} \omega_i f_i(x^*).$$

Isso é uma contradição com a suposição de que x^* é uma solução do problema de ponderação. Portanto, x^* é Pareto ótimo fraco. □

Teorema 2.1.2. *Um minimizador global do problema (2.1) é Pareto ótimo se os coeficientes de peso são todos positivos.*

Demonstração. Seja $x^* \in S$ um minimizador global do problema de pesos com coeficientes de peso positivos. Suponhamos que não seja Pareto ótimo. Isso significa que existe uma solução $x \in S$ tal que $f_i(x) \leqslant f_i(x^*)$ para todo $i = 1, \ldots, m$ e $f_j(x) < f_j(x^*)$ para pelo menos um j. Como $\omega_i > 0$ para todo $i = 1, \ldots, m$, temos

$$\sum_{i=1}^{m} \omega_i f_i(x) < \sum_{i=1}^{m} \omega_i f_i(x^*).$$

Isso contradiz a suposição de que x^* é uma solução do problema de ponderação e, portanto, x^* deve ser Pareto ótimo. □

Teorema 2.1.3. *Se o minimizador global do problema (2.1) é único, este é Pareto ótimo.*

Demonstração. Seja $x^* \in S$ um minimizador global *único* do problema de pesos. Suponhamos que não seja Pareto ótimo. Neste caso, existe uma solução $x \in S$ do problema multiobjetivo tal que $f_i(x) \leqslant f_i(x^*)$ para todo $i = 1, \ldots, m$ e $f_j(x) < f_j(x^*)$ para pelo menos um j. Como todos os coeficientes de peso ω_i são não negativos, temos

$$\sum_{i=1}^{m} \omega_i f_i(x) \leqslant \sum_{i=1}^{m} \omega_i f_i(x^*).$$

Por outro lado, a unicidade de x^* significa que

$$\sum_{i=1}^{m} \omega_i f_i(x^*) < \sum_{i=1}^{m} \omega_i f_i(\hat{x})$$

para todo $\hat{x} \in S$. As duas desigualdades acima são contraditórias e, portanto, x^* deve ser Pareto ótimo. □

Portanto, de acordo com os teoremas anteriores, qualquer minimizador global do problema de pesos é Pareto ótimo se todos os coeficientes de peso forem estritamente positivos ou se esse minimizador global for único. A fraqueza do método dos pesos é que nem todas as soluções Pareto ótimas podem ser encontrados, a não ser que o problema seja convexo.

Teorema 2.1.4 ([2]). *Seja um problema de otimização multiobjetivo convexo. Se $x^* \in S$ é Pareto ótimo então existe um vetor peso ω ($\omega_i \geqslant 0$, $i = 1, \cdots, m$, $\sum_{i=1}^{m} \omega_i = 1$) tal que x^* é a solução do problema (2.1).*

Pelo teorema acima, qualquer solução Pareto ótima de um problema de otimização multiobjetivo convexo pode ser encontrada pelo método dos pesos. Note que o vetor peso não é necessariamente único.

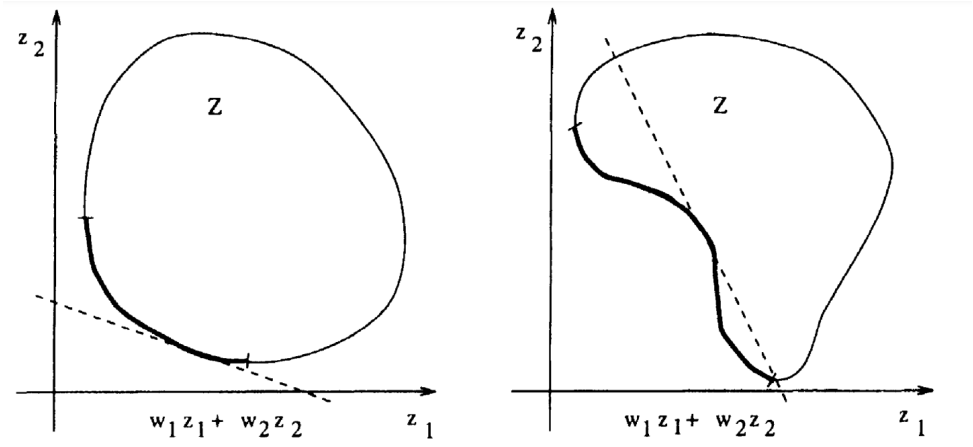

Figura 7 – Método dos pesos com problemas convexos e não convexos.

Fonte: Kaisa [2].

A Figura 7 ilustra o teorema anterior. No gráfico à esquerda, toda solução Pareto ótima ao longo da linha em negrito pode ser obtida ajustando os coeficientes de peso na função ponderada. No entanto, no gráfico à direita, não é possível encontrar as soluções localizadas na região côncava. De fato, ao deslocar a linha pontilhada para a esquerda, a função objetivo ponderada diminui. Como a função ponderada é uma combinação linear das funções objetivo, sua minimização sempre privilegia soluções na parte convexa do conjunto viável, deixando de explorar regiões côncavas da fronteira de Pareto.

Teorema 2.1.5 (Condição suficiente). *O minimizador global do problema (2.1) é propriamente Pareto ótimo se todos os coeficientes de peso forem positivos.*

Demonstração. Seja $x^* \in S$ um minimizador global do problema de pesos com coeficientes de peso positivos. No Teorema 2.1.2, mostramos que a solução é Pareto ótima. Vamos agora mostrar que x^* é propriamente Pareto ótimo com $M = (m-1) \max_{i,j}(\frac{\omega_j}{\omega_i})$.

Suponhamos que x^* não seja propriamente Pareto ótimo. Então, para algum i (que fixamos) e para $x \in S$ tal que $f_i(x^*) > f_i(x)$, temos

$$f_i(x^*) - f_i(x) > M(f_j(x) - f_j(x^*))$$

para todo j tal que $f_j(x^*) < f_j(x)$. Podemos agora escrever

$$f_i(x^*) - f_i(x) > (m-1)\frac{\omega_j}{\omega_i}(f_j(x) - f_j(x^*)).$$

Após multiplicar ambos os lados por $\dfrac{\omega_i}{m-1} > 0$, obtemos

$$\frac{\omega_i}{m-1}(f_i(x^*) - f_i(x)) > \omega_j(f_j(x) - f_j(x^*))(> 0 \geqslant \omega_l(f_l(x) - f_l(x^*))),$$

onde l difere do índice fixo i e dos índices j, que foram especificados anteriormente. Após este raciocínio, podemos somar sobre todos os $j \neq i$ e obter

$$\omega_i(f_i(x^*) - f_i(x)) > \sum_{\substack{j=1 \\ j\neq i}}^{m}(\omega_j(f_j(x) - f_j(x^*))),$$

o que significa

$$\sum_{j=1}^{m}\omega_j f_j(x^*) > \sum_{j=1}^{m}\omega_j f_j(x).$$

Aqui temos uma contradição com a suposição de que x^* é uma solução do problema de pesos. Assim, x^* tem que ser propriamente Pareto ótimo. □

Teorema 2.1.6 ([2]). *Se o problema de otimização multiobjetivo é convexo, então a condição do Teorema 2.1.5 é também necessária.*

Algoritmo 1 Método dos Pesos com Gradiente e Backtracking

---

1. Escolha $\alpha \in (0, 0.5)$, $\beta \in (0,1)$, uma tolerância $\epsilon > 0$, um vetor peso $\omega \in \mathbb{R}^m$ e um ponto inicial $x_0 \in \mathbb{R}^n$. Defina $k := 0$.
2. Defina a função objetivo escalar como $F_\omega(x) = \sum_{i=1}^{m}\omega_i f_i(x)$.
3. Calcule o gradiente $\nabla F_\omega(x_k) = \sum_{i=1}^{m} w_i \nabla f_i(x_k)$.
4. Se $\|\nabla F_\omega(x_k)\| < \epsilon$, pare.
5. Escolha o comprimento de passo inicial $t := 1$.
6. Calcule o comprimento de passo $t_k \in (0,1]$ como o máximo de

$$T_k := \{t = \frac{t}{2^j} \mid j \in \mathbb{N}, F_\omega(x_k - t\nabla F_\omega(x_k)) \leqslant F_\omega(x_k) - \alpha t\|\nabla F_\omega(x_k)\|^2\},$$

7. Defina $x_{k+1} := x_k - t_k \nabla F_\omega(x_k)$, $k := k+1$, e volte ao passo 3.

2.2 Método de Descida Máxima

Nesta seção iremos discorrer sobre o método de descida máxima. Proposto em [19], é uma adaptação direta do método do gradiente com busca linear inexata (Armijo com *backtracking*) para o contexto multiobjetivo. Os lemas e teoremas desta seção têm como referência [19]. Consideraremos apenas problemas irrestritos ($S = \mathbb{R}^n$).

Desta forma, a adaptação deve ser feita na condição de Armijo e no cálculo da direção de descida, que, em algum sentido, deve remeter ao gradiente. O método da descida máxima é de simples implementação e serve como arquétipo para a maioria dos métodos provenientes de adaptações de métodos clássicos.

O conceito usual de ponto estacionário, onde busca-se zerar a norma do gradiente da função objetivo, não se traduz bem para o contexto multiobjetivo. Portanto, afim de se obter um conceito de ponto estacionário computável, propõe-se o conceito de Ponto Pareto crítico, definido a seguir:

Definição 2.2.1. *Seja $F : \mathbb{R}^n \to \mathbb{R}^m$, $F(x) = (f_1(x), \cdots, f_m(x))$. Um ponto $x \in \mathbb{R}^n$ é chamado de **Pareto crítico** se não existe uma direção $v \in \mathbb{R}^n$ tal que $\nabla f_i(x)^t v < 0$ para todo $i = 1, \cdots, m$..*

Desta forma, a ideia geral deste método é escolher um ponto x e verificar se este é Pareto crítico. Se não for, buscamos uma direção de descida v e um tamanho de passo t. A partir do novo ponto encontrado, repetimos o mesmo processo até conseguirmos um ponto Pareto crítico.

2.2.1 Computando a direção de descida

Para computarmos esta direção de descida, em um dado ponto $x \in \mathbb{R}^n$ definimos a função $f_x : \mathbb{R}^n \to \mathbb{R}^m$ dada por

$$f_x(v) = \max\{(\nabla f_i(x)^t v) | i = 1, \cdots, m\}. \tag{2.2}$$

Desta forma, consideramos o problema de minimização irrestrito

$$\begin{gathered} \text{Minimizar } f_x(v) + \frac{1}{2}||v||^2 \\ \text{sujeito a } v \in \mathbb{R}^n. \end{gathered} \tag{2.3}$$

Temos que f_x é convexa por ser o máximo de funções lineares. Assim, no problema (2.3) a função objetivo é fechada, própria e fortemente convexa e portanto tem solução e esta é única.

Teorema 2.2.1. *Sejam $v(x)$ e $\alpha(x)$ a solução e o valor ótimo do problema (2.3), respectivamente.*

- *Se x é Pareto crítco, então $v(x) = 0$ e $\alpha(x) = 0$.*
- *Se x não é Pareto crítico, então $\alpha(x) < 0$ e*

$$f_x(v(x)) \leqslant -(1/2)||v(x)||^2,$$

$$\nabla f_i(x)^t v(x)) \leqslant f_x(v(x)), i = 1, \cdots, m.$$

- *Os mapas $x \to v(x)$ e $x \to \alpha(x)$ são contínuos.*

Em vez de resolver o Problema (2.3) de forma exata, pode ser interessante, do ponto de vista algorítmico, lidar com soluções inexatas. Assim, se x não é Pareto-crítico, diremos que v é uma solução aproximada de (2) com tolerância $\sigma \in]0, 1]$ se

$$f_x(v) + \tfrac{1}{2}\|v\|^2 \leqslant \sigma\,\alpha(x),$$

onde, conforme anteriormente, $f_x(v) := \max_i \left(\nabla f_i(x)^T v\right)$ e $\alpha(x)$ é o valor ótimo do Problema (2.3). Observe que, para $\sigma = 1$, somente a solução exata satisfaz a desigualdade acima.

Uma alternativa para o problema (2.3) para calcular a direção de descida é o problema a seguir:

$$\begin{aligned} &\text{Minimizar } f_x(v) \\ &\quad \text{sujeito a } ||v||_\infty \leqslant 1, \end{aligned} \tag{2.4}$$

o que pode ser reformulado como

$$\begin{aligned} &\text{Minimizar } \beta \\ &\quad \text{sujeito a } (\nabla f_i(x)^t v) \leqslant \beta, \qquad i = 1, \cdots, m, \\ &\qquad\qquad ||v||_\infty \leqslant 1, \end{aligned} \tag{2.5}$$

que é um problema linear.

Teorema 2.2.2. *Sejam $V(x), \beta(x)$ o conjunto solução e o valor ótimo do problema (2.4), respectivamente.*

- *Se x é Pareto crítico, então $0 \in V(x)$ e $\beta(x) = 0$.*
- *Se x não é Pareto crítico, então $\beta(x) < 0$ e para qualquer $v \in V(x)$ nós temos*

$$(\nabla f_i(x)^t v) \leqslant f_x(v) < 0, \qquad i = 1, \cdots, n.$$

- *O mapa $x \to \beta(x)$ é contínuo.*
- *Se x_k converge para $\overline{x}$, $v_k \in V(x_k)$ e v_k converge para $\overline{v}$ então $\overline{v} \in V(\overline{x})$.*

Lema 2.2.1. *Suponha que x não seja Pareto crítico e que v seja uma solução aproximada de (2.3) com tolerância $\sigma \in [0,1]$. Seja $A = JF(x)$, isto é, o jacobiano de F em x. Então,*

$$\|v\| \leqslant 2\|A\|_{\infty,2},$$

onde $\|A\|_{\infty,2}$ pode ser calculado equivalentemente como:

$$\|A\|_{\infty,2} := \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_2}$$

ou

$$\|A\|_{\infty,2} = \max_{i=1,\dots,m} \|A_i\|_2 = \max_{i=1,\dots,m} \left(\sum_{j=1}^{n} A_{i,j}^2 \right)^{1/2}.$$

onde A_i denota a i-ésima linha de A.

2.2.2 Computando o comprimento de passo

Para calcularmos o comprimento do passo usamos uma regra similar à regra de Armijo. Suponha que temos uma direção de descida $v \in \mathbb{R}^n$ com $JF(x)v < 0$. Seja $\gamma \in]0,1[$ uma constante dada. A condição para que um passo $t > 0$ seja aceito é:

$$F(x + tv) \leqslant F(x) + \gamma t JF(x)v. \tag{2.6}$$

Como na regra de armijo, começamos com $t = 1$ e enquanto esta condição não estiver satisfeita, fazemos $t := t/2$.

Teorema 2.2.3. *Se F é diferenciável e $JF(x)v < 0$, então existe algum $\epsilon > 0$ tal que*

$$F(x + tv) \leqslant F(x) + \gamma t JF(x)v \tag{2.7}$$

para qualquer $t \in]0, \epsilon[$.

Demonstração. Como assumimos que F é diferenciável, temos

$$F(x + h) = F(x) + JF(x)h + R(h)$$

com

$$\lim_{h \to 0} \frac{|R_i(h)|}{||h||} = 0, \quad i = 1, ..., m.$$

Definimos

$$a := \max_i (JF(x)v),$$

Observe que $a < 0$ e $v \neq 0$. Como $\gamma < 1$, existe um $\epsilon > 0$ tal que

$$0 < t \leqslant \epsilon \Rightarrow \frac{|R(tv)_j|}{||v||} < \frac{(1-\gamma)|a|}{||v||}, \quad i = 1, ..., m.$$

Portanto, para $0 < t \leqslant \epsilon$ obtemos

$$|R_i(tv)| < t(1-\gamma)|a|, \quad i = 1, ..., m,$$

e, como $|a| = -a = -\min_i(JF(x)v)$,

$$R(tv) < -t(1-\gamma)JF(x)v.$$

Portanto,

$$\begin{aligned} F(x+tv) &= F(x) + tJF(x)v + R(tv) \\ &< F(x) + tJF(x)v - t(1-\gamma)JF(x)v \\ &= F(x) + t\gamma JF(x)v \end{aligned}$$

para $0 < t \leqslant \epsilon$. □

Algoritmo 2 Método de Descida Máxima

1. Escolha $\gamma \in]0,1[$, $\sigma \in]0,1[$, $\tau > 0$ e $x_0 \in \mathbb{R}^n$. Defina $k := 1$.
2. Calcule v_k, uma solução aproximada de (2.3) em $x = x_k$ com tolerância σ. Seja α_k o valor da função objetivo de (2.3) em v_k.
3. Se $-\tau < \alpha_k$, pare.
4. Calcule o comprimento de passo $t_k \in]0,1]$ como o máximo em

$$T_k := \{t = 1/2^j | j \in \mathbb{N}, F(x_k + tv_k) \leqslant F(x_k) + \gamma t JF(x_k)v_k\}.$$

5. Defina $x_{k+1} := x_k + t_k v_k$, $k := k+1$ e volte para o passo 2.

Observe que se o Algoritmo 2 termina após um número finito de iterações, este termina em um ponto Pareto crítico. Para o Teorema a seguir, supomos que um sequência infinita é gerada, ou seja, $\alpha(x_k) \neq 0$ para todo $k \in \mathbb{N}$.

Teorema 2.2.4. *Todo ponto de acumulação da sequência (x_k) produzida pelo Algoritmo 2 é Pareto crítico. Se para a função F o conjunto $\{x \in \mathbb{R}^n | F(x) \leqslant F(x_0)\}$ for limitado, então a sequência (x_k) permanece limitada e tem pelo menos um ponto de acumulação.*

Demonstração. Seja y um ponto de acumulação da sequência (x_k), e sejam $v(y)$ e $\alpha(y)$ a solução e o valor ótimo de (2.3) em y:

$$\begin{aligned} v(y) &:= \arg\min_v \left(f_v(v) + \frac{1}{2}\|v\|^2 \right), \\ \alpha(y) &:= \min_v \left(f_v(v) + \frac{1}{2}\|v\|^2 \right), \end{aligned}$$

onde

$$f_v(v) := \max_i (JF(y)v)_i.$$

De acordo com o Teorema 2.2.1, é suficiente provar que $\alpha(y) = 0$. Claramente, a sequência $(F(x_k))$ é estritamente decrescente componente a componente e temos

$$\lim_{k\to\infty} F(x_k) = F(y).$$

Portanto,

$$\lim_{k\to\infty} \|F(x_k) - F(x_{k+1})\| = 0.$$

Mas

$$F(x_k) - F(x_{k+1}) \geqslant -t_k \gamma JF(x_k)v_k \geqslant 0,$$

e, portanto,

$$\lim_{k\to\infty} t_k JF(x_k)v_k = 0. \tag{2.8}$$

Observe que $t_k \in [0,1]$ para todo k. Agora, tome uma subsequência $(x_{k_u}$ convergente para y. Consideraremos duas possibilidades:

$$\limsup_{u\to\infty} t_{k_u} > 0$$

e

$$\limsup_{u\to\infty} t_{k_u} = 0.$$

Primeiro caso. Neste caso, existe uma subsequência $(x_{(k_l)})_l$ convergindo para y e satisfazendo

$$\lim_{l\to\infty} t_{k_l} = \bar{t} > 0.$$

Usando (2.8), concluímos que

$$\lim_{l\to\infty} JF(x_{(k_l)})v_{(k_l)} = 0,$$

o que também implica

$$\lim_{l\to\infty} \alpha(x_{(k_l)}) = 0.$$

Como $x \mapsto \alpha(x)$ é contínua, concluímos que $\alpha(y) = 0$, então y é crítico de Pareto.

Segundo caso. Devido aos resultados do Lema 2.2.1, sabemos que a sequência $(v_{(k_u)})_u$ é limitada. Portanto, podemos tomar uma subsequência $(x_{(k_r)})_r$ de $(x_{(k_u)})_u$ tal que a sequência $(v_{(k_r)})_r$ também converge para algum $\bar{v}$. Observe que para todo r temos

$$\max_i (JF(x_{(k_r)})v_{(k_r)})_i \leqslant \sigma\alpha(x_{(k_r)}) < 0.$$

Portanto, passando para o limite $r \to \infty$, obtemos

$$\frac{1}{\sigma} \max_i (JF(y)\bar{v})_i \leqslant \alpha(y) \leqslant 0. \tag{2.9}$$

Tome algum $q \in \mathbb{N}$. Para r suficientemente grande,

$$t_{k_r} < \frac{1}{2^q},$$

o que significa que a condição de Armijo (2.6) não é satisfeita para $t = 1/2^q$, isto é,

$$F_j(x_{(k_r)} + (1/2^q)\bar{v}) \geqslant F_j(x_{(k_r)}) + \gamma(1/2^q)JF(x_{(k_r)})\bar{v}$$

(para r suficientemente grande). Passando para o limite $r \to \infty$ (ao longo de uma subsequência adequada, se necessário), obtemos

$$F_j(y + (1/2^q)\bar{v}) \geqslant F_j(y) + \gamma(1/2^q)(JF(y)\bar{v})_j,$$

para pelo menos um $j \in \{1, ..., m\}$. Note que esta desigualdade vale para qualquer $q \in \mathbb{N}$. Do Teorema 2.2.3,

$$\max_i (JF(y)\bar{v})_i \geqslant 0,$$

o que, juntamente com (2.9), implica $\alpha(y) = 0$. Então, novamente concluímos que y é crítico de Pareto. □

2.3 Método dos Pesos Com Coleta

Ambas as estratégias anteriores encontram um único ponto Pareto sempre que o algoritmo é executado. Pensando em remediar isto, e inspirados no que foi abordado em [20, 21], propomos o Método dos Pesos com Coleta.

O método consiste em resolver o problema de pesos (2.1), assim como no Método dos Pesos convencional, mas coletando todos os pontos encontrados durante a busca linear de cada iteração e adicionando-os à uma lista temporária de pontos, criada anteriormente. O Algoritmo 3 descreve este processo.

Algoritmo 3 Método dos Pesos com Coleta

1. Escolha $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$, uma tolerância $\epsilon > 0$, um vetor peso $\omega \in \mathbb{R}^m$, e um ponto inicial $x_0 \in \mathbb{R}^n$. Defina $k := 0$. Inicialize uma lista temporária de pontos $\mathcal{L} := \varnothing$.
2. Defina a função objetivo escalar como $F_\omega(x) = \sum_{i=1}^{m} \omega_i f_i(x)$.
3. Calcule o gradiente $\nabla F_\omega(x_k) = \sum_{i=1}^{m} \omega_i \nabla f_i(x_k)$.
4. Se $\|\nabla F_\omega(x_k)\| < \epsilon$, pare.
5. Escolha o comprimento de passo inicial $t := 1$.
6. Execute o procedimento de busca linear com *backtracking* para determinar $t_k \in (0, 1]$ que satisfaça a condição

   $$F_\omega(x_k - t_k \nabla F_\omega(x_k)) \leqslant F_\omega(x_k) - \alpha t_k \|\nabla F_\omega(x_k)\|^2,$$

   coletando, ao longo do processo, os pontos intermediários gerados $x_{temp} := x_k - t \nabla F_\omega(x_k)$ e adicionando-os à lista $\mathcal{L}$, isto é, $\mathcal{L} := \mathcal{L} \cup \{x_{temp}\}$.
7. Atualize $x_{k+1} := x_k - t_k \nabla F_\omega(x_k)$, $k := k + 1$, e volte ao passo 3.
8. Retorne a lista de pontos $\mathcal{L}$.

Esta lista temporária é então filtrada, removendo todos os pontos dominados por algum outro da lista, restando apenas pontos não dominados. Este processo de coleta e filtragem é descrito no Algoritmo 4.

Algoritmo 4 Remover Pontos Dominados

1. **Entrada**: lista L onde cada elemento representa um ponto no espaço objetivo. Defina n como o número de elementos da lista L;
2. Inicie uma lista L_{filtro} para os pontos não dominados;
3. **Para** $i = 1, 2, \ldots, n_pontos$ **faça**:
   a) Defina $dominado := falso$;
   b) **Para** $j = 1, 2, \ldots, n_pontos$ **faça**:
      i. Se $i \neq j$ e o ponto $L[j]$ domina o ponto $L[i]$, então defina $dominado :=$ verdadeiro e **pare** o laço interno.
   c) Se $dominado =$ falso, adicione $L[i]$ à lista L_{filtro}.
4. **Retorne** L_{filtro}.

Desta forma, cada execução do algoritmo pode retornar múltiplos pontos Pareto. Conforme será mostrado no Capítulo 4, o Método dos Pesos com Coleta se mostrou capaz de encontrar pontos Pareto nas regiões de concavidade do espaço objetivo, algo que não é possível com o Método dos Pesos convencional.

3 Implementação

Desenvolvemos a implementação dos três métodos previamente descritos utilizando a linguagem Julia [22]. Este capítulo detalha os aspectos técnicos relativos à implementação de cada método.

Os códigos desenvolvidos neste trabalho estão disponíveis publicamente no GitHub e podem ser acessados em: https://github.com/lucianobegot/Trabalho-de-mestrado.

Em todos os casos, empregamos a estratégia de multi-start como abordagem para identificar múltiplos pontos ao longo da fronteira de Pareto, buscando uma aproximação robusta e abrangente desta fronteira. Nesse contexto, estabelecemos um número inicial de pontos, denotado por n_0, como ponto de partida para os algoritmos. Esses pontos iniciais são gerados aleatoriamente, de maneira uniforme, dentro de um intervalo pré-definido, $[a, b]$, que é especificado em função do problema analisado.

Nos casos em que o problema envolve múltiplas variáveis de decisão, cada variável é gerada de forma independente, mas sempre respeitando o mesmo intervalo $[a, b]$. Essa abordagem garante diversidade nas condições iniciais, favorecendo uma melhor exploração do espaço de busca e, consequentemente, uma representação mais precisa da região de Pareto.

3.1 Método dos pesos

Neste método, além dos pontos iniciais, geramos aleatoriamente os pesos $\omega_1, \cdots, \omega_m$, onde m é o número de funções objetivo, de modo que $\sum_{i=1}^{m} \omega_i = 1$. Estes pesos são sorteados de novo para cada ponto inicial.

Para discutir o processo de geração do vetor de pesos ω, é importante observar que, conforme destacado por [23], a simples normalização de um vetor arbitrário não resulta em um conjunto de pontos uniformemente distribuídos no simplex unitário. Esse fato ocorre devido à concentração desigual dos pontos no espaço, o que compromete a uniformidade necessária.

Algoritmo 5 Geração de pesos uniformemente distribuídos no simplex unitário.

1. Entrada: Um valor inteiro n, representando o número de pesos a serem gerados.
2. Inicialize $\texttt{sum} := 0$ e uma lista vazia ω para armazenar os pesos gerados.
3. **Para** $i = 1, 2, \ldots, n-1$ **faça**:
   a) Gere $\phi_i$uniformemente em $[0, 1]$;
   b) Calcule $x_i := (1 - \texttt{sum}) \left(1 - \phi_i^{1/(n+1-i)}\right)$;
   c) Atualize $\texttt{sum} := \texttt{sum} + x_i$;
   d) Adicione x_i à lista ω.
4. Após o laço, calcule $x_n := 1 - \texttt{sum}$ e adicione x_n à lista ω.
5. **Retorne** a lista ω.

Para contornar essa limitação, utilizamos o método presente em [23], capaz de gerar pontos uniformemente distribuídos no simplex unitário. Esse método é descrito no Algoritmo (5). Essa abordagem, além de possuir embasamento teórico sólido, apresenta vantagens práticas, como simplicidade na implementação e eficiência computacional.

Para resolver o problema (2.1), aplicamos o método do gradiente com busca linear inexata, utilizando a regra de Armijo com *backtracking*, conforme descrito no Algoritmo 1. Os parâmetros do método foram definidos como tolerância $\epsilon = 10^{-6}$ e constante de Armijo $\alpha = 10^{-2}$. O critério de parada do algoritmo é atingido quando

$$\left\| \sum_{i=1}^{m} \omega_i \nabla f_i(x) \right\|_{\infty} < \epsilon$$

ou quando o número máximo de 500 iterações for atingido.

Desta forma, o método encontra no máximo um ponto Pareto para cada ponto inicial x_n^0.

3.2 Método de descida máxima

Implementamos o algoritmo 2 resolvendo o problema (2.5), ao invés de (2.3). Começamos definindo a função (2.2), para resolvermos o problema (2.5). Desta forma, implementamos o Algoritmo 2, utilizando os parâmetros $\gamma = 10^{-3}$ e $\tau = 10^{-6}$. No passo 2 do algoritmo, os valores de v_k e α_k são obtidos por meio do solver CPLEX, utilizando o método e os parâmetros padrões do solver. Assim, o algoritmo para quando $-\tau < \alpha_k$ ou quando o número máximo de 500 iterações for atingido.

Assim como o método dos pesos, este método encontra no máximo um ponto Pareto para cada ponto inicial x_n^0.

3.3 Método dos pesos com coleta

A implementação do método dos pesos com coleta segue a mesma estrutura básica do método dos pesos convencional, envolvendo o sorteio de pesos e a minimização da soma ponderada. A principal modificação reside na coleta de pontos ao longo do processo de busca linear.

No início do código é criada uma lista vazia L. Durante a execução do algoritmo de minimização de (2.1), todos os pontos visitados ao longo da busca linear são adicionados à lista L. Após a execução do algoritmo para os n_0 pontos iniciais, realiza-se uma filtragem nos pontos acumulados em L, removendo aqueles que são dominados por outros. Esse procedimento resulta em um conjunto final contendo apenas pontos não dominados. O processo de filtragem dos pontos dominados está detalhado no Algoritmo 4.

Com essa abordagem, cada ponto inicial x_n^0 pode produzir múltiplos pontos Pareto em uma única execução do algoritmo.

3.4 Métricas

A avaliação do desempenho dos métodos nos diferentes problemas teste foi conduzida utilizando métricas de pureza e espalhamento, seguindo a abordagem adotada em [20]. A métrica de pureza quantifica a proporção de pontos Pareto presentes nas aproximações das fronteiras de Pareto obtidas por cada método, permitindo avaliar a qualidade da convergência para a fronteira ótima. A métrica de espalhamento, por sua vez, avalia a distribuição espacial dos pontos Pareto na aproximação, fornecendo informações sobre a diversidade das soluções encontradas. Para uma análise comparativa abrangente do desempenho dos métodos, foram construídos perfis de desempenho, conforme proposto em [24], que permitem visualizar e comparar o comportamento dos métodos em todo o conjunto de problemas.

3.4.1 Métrica de Pureza

Seja M o conjunto dos métodos e P o conjunto dos problemas teste. Chamamos de $F_{me,p}$ a aproximação da fronteira de Pareto encontrada pelo método $me \in M$ para o problema $p \in P$. Definimos também F_p como senda a aproximação da fonteira de Pareto real para o problema p, calculando primeiramente $\bigcup_{me \in M} F_{me,p}$ e em seguida removendo todos os pontos dominados.

Assim, a métrica consiste em calcular a razão $\frac{|F_{me,p} \cap F_p|}{|F_p|}$. Assim, a métrica

é representada por um número $\bar{t}_{me,p} = \dfrac{|F_{me,p} \cap F_p|}{|F_p|}$ entre 0 e 1, onde valores maiores de $\bar{t}_{me,p}$ representam uma fronteira de Pareto melhor.

3.4.2 Métrica de Espalhamento

Mais uma vez, denotamos de M o conjunto dos métodos e P o conjunto dos problemas teste. Como estamos interessados em computar um conjunto de pontos que abranja toda a verdadeira fronteira de Pareto, as métricas propostas devem considerar primeiro os "pontos extremos" no espaço das funções objetivo $\mathbb{R}^n$, que serão os mesmos para a aplicação das métricas em qualquer uma das fronteiras obtidas. A descrição do cálculo desses pontos extremos será dada mais adiante nesta seção.

Suponhamos que um método $me \in M$ computou N pontos pareto para um problema $p \in P$. Para cada objetivo j, colocamos os valores das funções objetivo no ponto em ordem crescente. Assim, a estes N pontos computados adicionamos os pontos extremos, indexados por 0 e $N+1$. Assim, a métrica de espalhamento Γ consiste de calcular:

$$\Gamma_{me,p} = \max_{j \in \{1,\cdots,m\}} \left(\max_{i \in \{1,\cdots,N\}} (\delta_{i,j}) \right) \tag{3.1}$$

onde $\delta_{i,j} = f_j(x_{i+1}) - f_j(x_i)$. Desta forma, esta fórmula busca medir o maior "buraco" na fronteira de Pareto, considerando a maior distância pela norma do infinito entre dois pontos consecutivos na fronteira de Pareto aproximada. A Figura 8 ilustra as distâncias $\delta_{i,j}$.

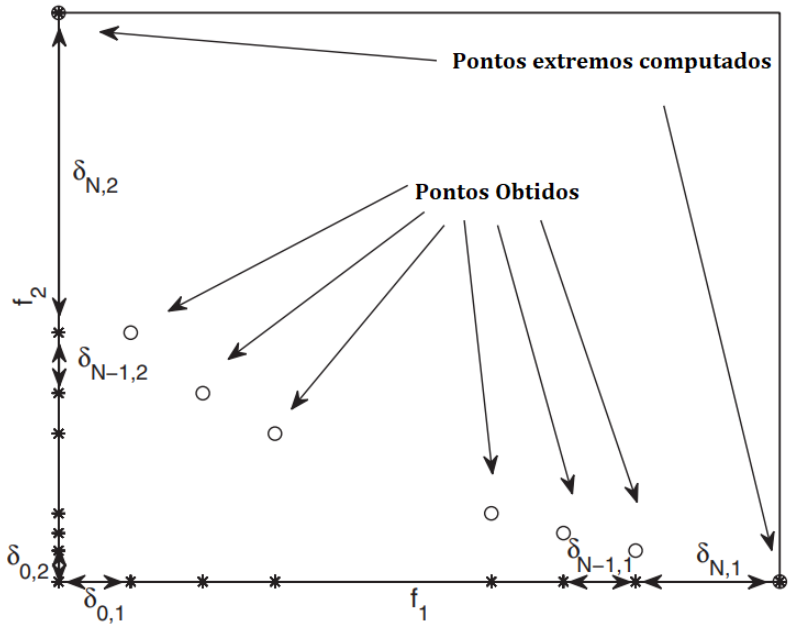

Figura 8 – Distâncias calculadas por $\delta_{i,j}$

Fonte: Custódio, Madeira, Vaz e Vicente [20].

4 Testes Numéricos

Coletamos diversos problemas teste presentes na literatura, todos sem restrição. A lista completa de problemas teste pode ser vista no Apêndice A . Dentre os problemas escolhidos, existem problemas convexos e não convexos.

Estamos interessados em dois pontos principais: a eficiência dos métodos em cada problema e se conseguimos ou não gerar uma boa aproximação da fronteira de Pareto.

4.1 Problemas Ilustrativos

Iniciamos os testes com problemas mais simples, os quais chamamos de problemas ilustrativos, com o objetivo de verificar a capacidade de cada método de encontrar a aproximação da fronteira de Pareto. Estes são problemas teste com apenas uma variável e duas funções objetivo, de modo que conseguimos plotar a aproximação da fronteira de Pareto para termos uma ideia de como cada método se comporta.

Em todos os exemplos abaixo, o número de pontos inciais é $n_0 = 300$, resultando em no máximo 300 pontos pareto para os métodos dos pesos e da descida máxima. Em todos os problemas sorteamos os pontos iniciais dentro de um intervalo $[a, b]$, definido previamente dependendo do problema a ser resolvido. O intervalo utilizado para cada problema é especificado no Apêndice A.

Discutimos primeiro o desempenho do Método dos Pesos e do Método da Descida Máxima nestes problemas teste e em seguida analisamos em que aspectos o Método dos Pesos com Coleta consegue alguma melhora para estes mesmos problemas.

Começamos com o problema DGO1:

Exemplo 4.1.1. *Considere* $f_1 = \text{sen}(x)$ *e* $f_2 = \text{sen}(x + 0.7)$ *e o seguinte problema:*

$$\begin{gathered} \textit{Minimizar } \{f_1(x), f_2(x)\} \\ \textit{sujeito a } x \in \mathbb{R} \end{gathered}$$

Temos neste caso um problema com região objetiva convexa. Tanto o Método dos Pesos quanto o da Descida Máxima conseguiram encontrar uma aproximação da fronteira de Pareto.

Porém, o Método da descida máxima encontra também pontos de 'máximo', que são pontos Pareto locais indesejáveis. Estes pontos de 'máximo' podem ser retirados aplicando o filtro previamente discutido à lista de pontos Pareto. Contudo, isso resulta em uma quantidade menor

de pontos Pareto crítico obtidos pelo Método da Descida Máxima, em comparação ao Método dos Pesos, que encontra um ponto Pareto para cada um dos 300 pontos iniciais. Especificamente, o Método da Descida Máxima identifica 26 pontos de 'máximo', os quais são descartados pelo filtro, reduzindo o total efetivo de pontos Pareto crítico encontrados.

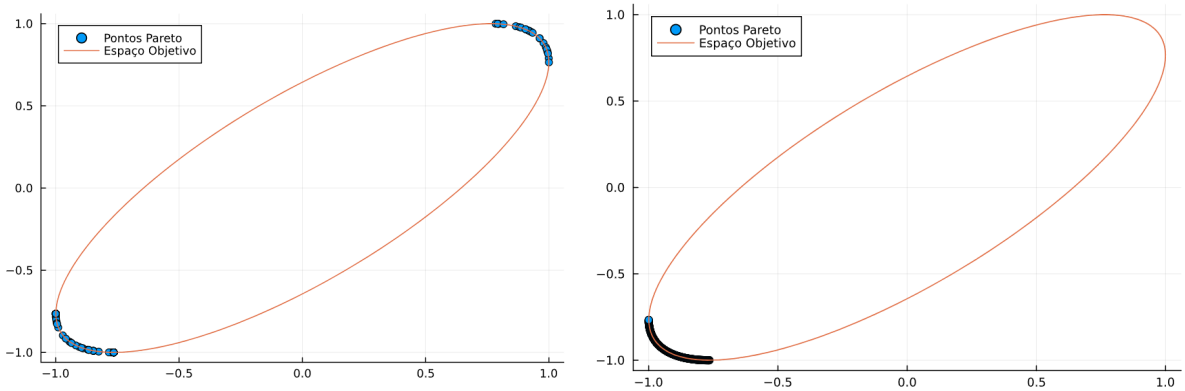

Figura 9 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.1.

Em seguida, analisamos o problema SSFYY2, com concavidades no espaço objetivo.

Exemplo 4.1.2. *Considere:* $f_1 = 10 - 10\cos\left(x\frac{\pi}{2}\right) + x^2$ *e* $f_2 = (x-4)^2$ *e o seguinte problema:*

$$\text{Minimizar } \{f_1(x), f_2(x)\}$$
$$\text{sujeito a } x \in \mathbb{R}$$

Mais uma vez, ambos os métodos são capazes de encontrar uma aproximação da fronteira de Pareto. O método da descida máxima continua encontrando pontos de 'máximo', assim como no exemplo anterior. Além disso, o método dos pesos não consegue encontrar pontos Pareto próximos à região de concavidade.

Por fim, analisamos o problema DGO1 Modificado:

Exemplo 4.1.3. *Considere:*

$$f_1 = \text{sen}(x) + 0.1\cos(x)\text{sen}(x),$$

$$f_2 = \text{sen}(x+0.7) + 0.1\cos(x)\text{sen}(x+0.7)$$

e o seguinte problema:

$$\text{Minimizar } \{f_1(x), f_2(x)\}$$
$$\text{sujeito a } x \in \mathbb{R}$$

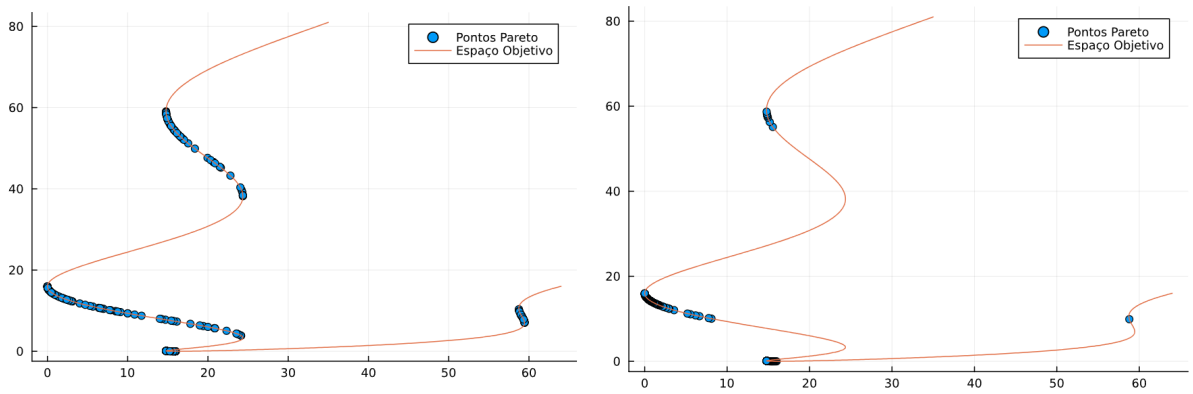

Figura 10 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.2.

Apesar dos dois métodos serem capazes de aproximar a fronteira de Pareto, mais uma vez vemos os mesmos problemas com os dois métodos. O método da descida máxima continua encontrando pontos de 'máximo', enquanto o método dos pesos é incapaz de encontrar Pontos Pareto na concavidade.

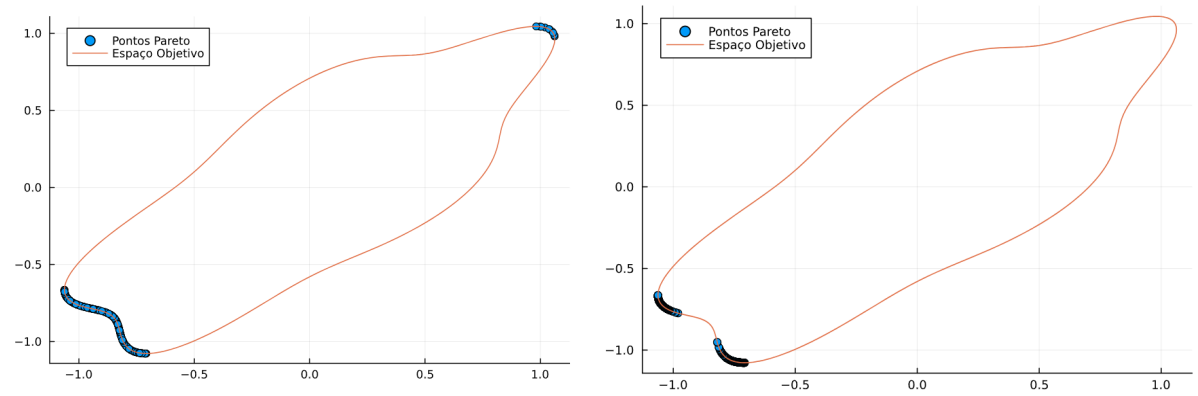

Figura 11 – Aproximação da fronteira de Pareto com método da descida máxima (esquerda) e com método dos pesos (direita) para o Exemplo 4.1.3.

Vemos que mesmo em problemas simples, onde temos apenas uma variável e duas funções objetivos, a não convexidade do problema faz com que a fronteira de Pareto aproximada não seja ideal. Como dito anteriormente, isto nos motivou a verificar o desempenho do método dos pesos com coleta nestes mesmos exemplos.

Nas Figuras 12 e 13 vemos que o método dos pesos com coleta consegue não apenas aproximar a fronteira de Pareto para estes problemas teste, mas também preencher as regiões de concavidade. Além disto, a natureza do processo de filtragem faz com que pontos paretos locais e pontos Pareto de 'máximo' sejam excluídos do processo.

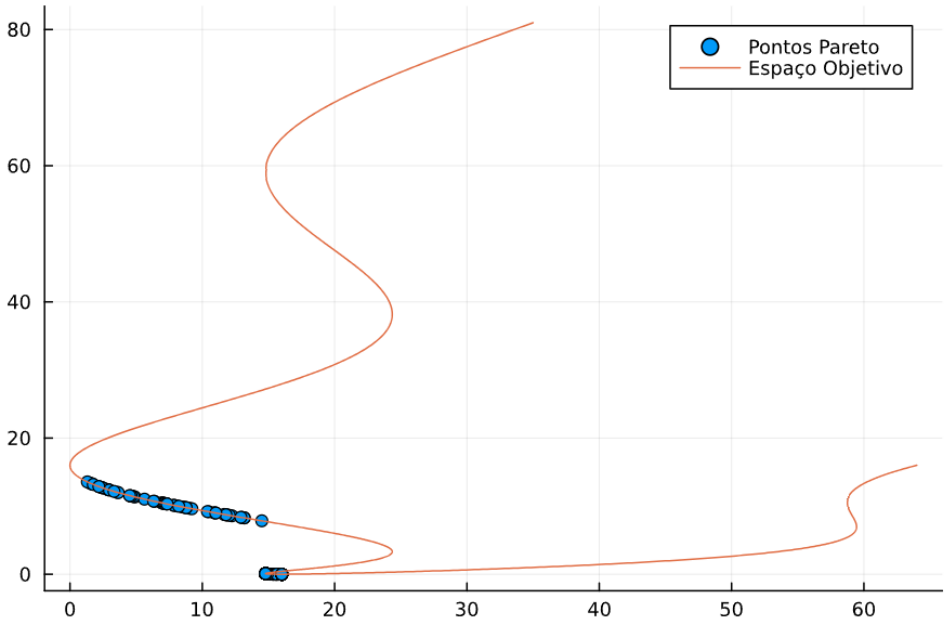

Figura 12 – Aproximação da fronteira de Pareto com método dos pesos com coleta para o Exemplo 4.1.2.

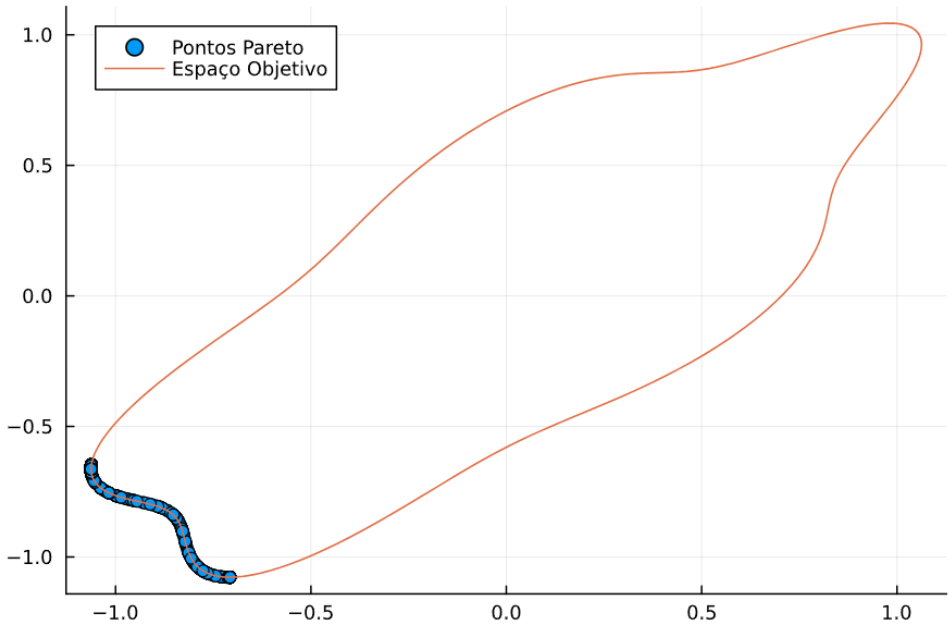

Figura 13 – Aproximação da fronteira de Pareto com método dos pesos com coleta para o Exemplo 4.1.3.

4.2 Checagem de Pareto Crítico

Apesar do método dos pesos com coleta mostrar melhores resultados nos problemas ilustrativos, nos falta uma garantia teórica de que os pontos não dominados encontrados sejam de fato pontos Pareto.

Logo, na falta de uma garantia teórica, nos voltamos ao conceito de ponto Pareto crítico introduzido na Seção 2.2. Nos utilizando da ideia introduzida no método da Descida Máxima, podemos verificar computacionalmente se os pontos encontrados pelo método dos pesos com coleta são Pareto críticos.

Desta forma, para cada ponto retornado pelo método dos pesos com coleta, aplicamos a checagem de Pareto crítico descrita no Algoritmo 6. Se o ponto for Pareto crítico o mantemos, caso contrário ele é descartado. Usamos $\sigma = 10^{-3}$ e $\tau = 10^{-6}$.

Algoritmo 6 Checagem de Pareto Crítico

1. **Entrada**: $x_k \in \mathbb{R}^n$.
2. Escolha $\sigma \in]0, 1[$, $\tau > 0$ e $x \in \mathbb{R}^n$.
3. Calcule v_k, uma solução aproximada de (2.5) em $x = x_k$ com tolerância σ. Seja α_k o valor da função objetivo de (2.5) em v_k.
4. Se $-\tau < \alpha$, conclua que x **não é Pareto crítico**. Caso contrário, x é **Pareto crítico** com tolerância τ.

Assim, a Tabela 1 mostra a proporção de pontos encontrados pelo método dos pesos com coleta que é de fato Pareto crítico para alguns problemas selecionados. A coluna PND representa a quantidade de pontos não dominados retornados pelo método, a coluna PPC representa a quantidade de pontos que passaram pela checagem de Pareto crítico e a coluna Razão nos dá a razão

$$r = \frac{PPC}{PND},$$

ou seja, a proporção dos pontos encontrados que são de fato Pareto crítico. Os problemas selecionados foram escolhidos para ilustrar como a razão se comporta em função de diferentes características, como a quantidade de variáveis, o número de funções objetivo e a convexidade do problema.

Problema	n	m	Conv	PND	PPC	Razão
LOV1	2	2	Sim	1519	1458	0.9598
SLCDT	10	3	Sim	19211	13589	0.7073
DGO1	1	2	Sim	1053	1053	1
DGO1 M.	1	2	Não	5437	5437	1
SSFFY 2	1	2	Não	15250	15250	1
MOP 2	15	2	Não	14331	12329	0.8603
VU1	2	2	Não	11441	9599	0.8390
DGO1 M. 2	100	2	Não	5095	3814	0.7485
DGO1 M. 3	1	100	Não	15611	15611	1

Tabela 1 – Razão de pontos Pareto Crítico.

Analisando a Tabela 1, percebemos que para alguns problemas, como o *DGO 1* e *DGO 1 M. 3*, a razão é exatamente 1, indicando que todos os pontos não dominados

retornados pelo método são Pareto críticos. Ao mesmo tempo, para problemas como *DGO1 M. 2*, a razão é significativamente menor, indicando que uma grande parte dos pontos não dominados não atende ao critério de Pareto crítico.

Neste caso, não é possível identificar de imediato uma relação entre a proporção de pontos Pareto crítico e a convexidade dos problemas em questão. Porém, percebemos que existe uma relação entre o número de variáveis e a razão. Mais especificamente, problemas com um número muito alto de variáveis tendem a ter uma proporção de pontos Pareto crítico baixa. Além disso, é interessante notar que para problemas com poucas variáveis e muitos objetivos a razão não é muito afetada.

Portanto, é necessário desenvolver estratégias para melhorar a razão r, já que valores baixos indicam que o método dos pesos com coleta retorna um grande número de pontos que não são, de fato, Pareto críticos. Isso reduz sua utilidade prática, especialmente para tomadores de decisão que dependem de soluções mais confiáveis. Além disso, uma razão mais próxima de 1 implica menor desperdício de recursos computacionais, evitando o processamento desnecessário de pontos irrelevantes.

Para resolver isto, a estratégia adotada foi a de restringir a maneira como os pontos são coletados durante a busca linear. Observamos experimentalmente que os pontos que eram aceitos como pontos Pareto crítico tendiam a ser aqueles que estavam mais próximos do fim da busca linear. Intuitivamente, isto faz sentido pois esses pontos tendem a estar mais próximos dos pontos que, segundo o Teorema 2.1.2, têm garantia teórica de serem Pareto críticos.

Assim, implementamos esta restrição impondo uma condição de apenas coletar pontos na busca linear quando

$$||\nabla F_\omega(x_k)|| \leqslant \lambda \tag{4.1}$$

onde λ é uma tolerância definida previamente.

Desta forma construímos a Tabela 2, refazendo todos os testes da Tabela 1, mas utilizando a restrição (4.1) com $\lambda = 10^{-1}$ para limitar os pontos coletados. Cada coluna da Tabela 2 representa o mesmo que representava na Tabela 1.

Problema	n	m	Conv	PND	PPC	Razão
LOV1	2	2	Sim	1455	1450	0.9965
SLCDT	10	3	Sim	11829	11706	0.9896
DGO1	1	2	Sim	997	997	1
DGO1 M.	1	2	Não	5179	5179	1
SSFFY 2	1	2	Não	14602	14602	1
MOP 2	15	2	Não	13357	12325	0.9227
VU1	2	2	Não	9382	9382	1
DGO1 M. 2	100	2	Não	4013	3567	0.8888
DGO1 M. 3	1	100	Não	15273	15273	1

Tabela 2 – Razão de pontos Pareto Crítico após a implementação da restrição.

Desta vez, podemos perceber uma melhora considerável na razão r para todos os problemas listados. Vale destacar que, nos casos em que alcançamos a razão igual a 1, alguns pontos críticos de Pareto podem ter sido descartados. No entanto, essa restrição continua sendo vantajosa, pois a quantidade de pontos críticos descartados é pequena em relação ao total de pontos encontrados. Portanto, para os testes a seguir utilizaremos sempre esta adaptação do método dos pesos com coleta.

4.2.1 Escolha de λ

A escolha de λ é feita de modo evitar a coleta de pontos não Pareto críticos e, assim, diminuir o gasto computacional. Porém, uma escolha inadequada de λ pode levar o algoritmo à exclusão de pontos Pareto crítico, fazendo com que o método não encontre uma fronteira de Pareto adequada.

Iremos ilustrar isto através do problema MOP 2, apresentado a seguir:

Exemplo 4.2.1. *Considere as funções objetivo:* $f_1(x) = 1 - \exp\left(-\frac{1}{15}\sum_{j=1}^{15}(x_j - 1)^2\right)$ *e* $f_2(x) = 1 - \exp\left(-\frac{1}{15}\sum_{j=1}^{15}(x_j + 1)^2\right)$, *e o seguinte problema:*

$$\begin{gathered} \textit{Minimizar } \{f_1(x), f_2(x)\}, \\ \textit{sujeito a } x \in \mathbb{R}^{15}. \end{gathered}$$

Este é um problema não convexo, com 15 variáveis e 2 funções objetivo. Portanto, é possível visualizar o seu espaço objetivo e sua respectiva fronteira de Pareto.

Inicialmente, comparamos a fronteira de Pareto obtida pelo método dos pesos com coleta, utilizando $\lambda = 10^{-3}$, com as obtidas pelos métodos de descida máxima e dos pesos convencionais. A Figura 14 apresenta os resultados.

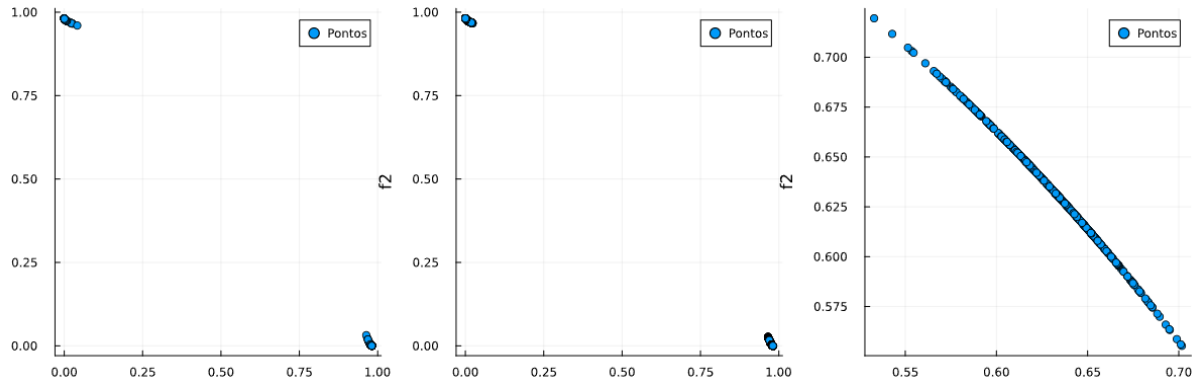

Figura 14 – Aproximação da fronteira de Pareto para o problema MOP 2 para o método dos pesos (esquerda), método dos pesos com coleta utilizando $\lambda = 10^{-3}$ (centro) e o método da descida máxima (direita).

Observa-se que, para $\lambda = 10^{-3}$, o método dos pesos com coleta não apresenta uma aproximação significativamente superior à do método dos pesos convencional. Esse resultado sugere que o valor de λ foi restritivo a ponto de descartar pontos Pareto críticos relevantes.

Com isso, realizamos uma nova análise, aumentando λ para 10^{-1}. Os resultados obtidos são apresentados na Figura 15.

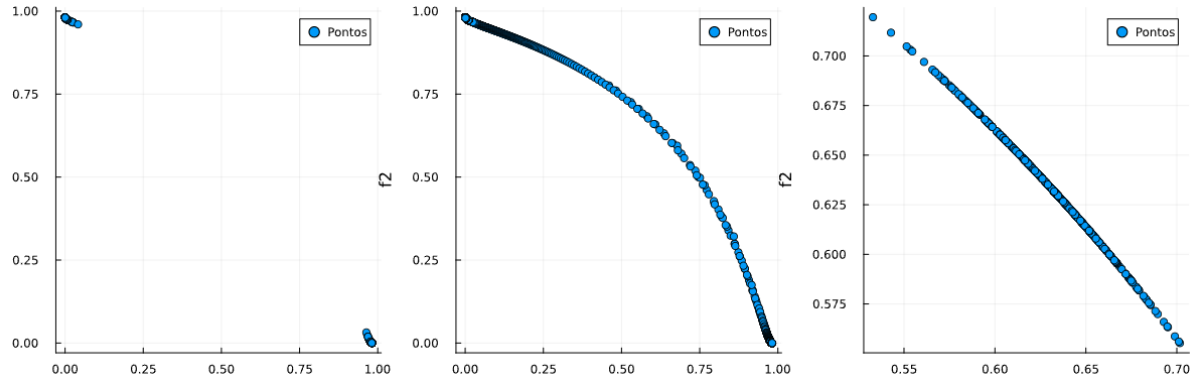

Figura 15 – Aproximação da fronteira de Pareto para o problema MOP 2 para o método dos pesos (esquerda), método dos pesos com coleta utilizando $\lambda = 10^{-1}$ (centro) e o método da descida máxima (direita).

Com $\lambda = 10^{-1}$, o método dos pesos com coleta consegue uma aproximação muito mais abrangente da fronteira de Pareto, comparável à obtida em [9]. Esse resultado destaca a relevância de um ajuste apropriado de λ para garantir a eficácia do método.

Por fim, ressaltamos que a escolha de λ deve equilibrar dois objetivos: evitar a coleta de pontos não Pareto críticos, o que reduz o custo computacional, e

minimizar o risco de descartar pontos Pareto críticos, o que comprometeria a qualidade da aproximação da fronteira de Pareto. Uma restrição excessiva em λ pode levar a resultados insatisfatórios, como evidenciado no caso de $\lambda = 10^{-3}$. Portanto, para todos os próximos testes realizados neste capítulo, utilizaremos $\lambda = 10^{-1}$.

4.3 Tempo de Execução

Nesta seção, comparamos o tempo de execução do método dos pesos com coleta em relação ao método da descida máxima. O objetivo é avaliar a eficiência de cada abordagem na identificação de pontos Pareto dentro de um intervalo de tempo específico, destacando o desempenho relativo dos métodos na geração de soluções eficientes. Não comparamos ambas as estratégias com o método dos pesos convencional devido à sua incapacidade de encontrar pontos Pareto nas regiões de concavidade Como a filtragem dos pontos Pareto no método dos pesos é dependente de condições internas que dificultam o controle direto do tempo de execução, foi necessário desenvolver uma abordagem alternativa para medir e comparar sua eficiência frente aos demais métodos. A ideia central foi medir a eficiência relativa dos métodos em termos da quantidade de pontos Pareto gerados e o tempo necessário para obter essa quantidade. A abordagem adotada pode ser descrita em duas etapas principais, com seus resultados destacados na Tabela 3:

- **Execução limitada do método da descida máxima (D.M.):** O método D.M. foi executado por um intervalo de tempo fixo de 10 segundos para cada problema. O número de pontos Pareto encontrados nesse intervalo foi registrado na coluna *10s c/ D.M.*.
- **Execução do método dos pesos com coleta de pontos:** Para o método dos pesos com coleta, foi mensurado o tempo necessário para coletar aproximadamente o mesmo número de pontos Pareto encontrados pelo método da descida máxima em 10 segundos. Permitimos uma margem de erro de até 5% na quantidade de pontos coletados para facilitar a comparação. Esse tempo é registrado na coluna *Tempo coleta*.

Problema	n	m	Convexo?	10s c/ D.M.	Tempo coleta (s)
LOV1	2	2	Sim	5944	4.99
MHHM1	1	3	Sim	14546	17.32
MHHM2	2	3	Sim	10689	9.09
SLCDT 2	10	3	Sim	86	0.79
DGO 1	1	2	Sim	17746	14.84
JOS1	100	2	Sim	95	2.01
SSFFY 1	2	2	Sim	1123	3.51
FAR1	2	2	Não	194	3.04
DGO1 M.	1	2	Não	22811	25.33
SSFFY 2	1	2	Não	2041	3.86
MOP 2	15	2	Não	1153	3.52
VU1	2	2	Não	312	0.54
DGO1 M. 2	100	2	Não	18	0.65
DGO1 M. 3	1	100	Não	2331	12.01
QV1	10	2	Não	1	1.52
MOP 5	2	3	Não	38	1.70
MOP 5 Modificado	200	10	Não	10	0.14
Toi9	4	4	Não	758	0.867
MGH26	4	4	Não	2014	5.92
MGH26 Modificado	50	50	Não	7	0.96
FF1	2	2	Não	418	1.98

Tabela 3 – Resultados dos testes de tempo.

Para a maioria dos problemas, o tempo necessário para o método dos pesos com coleta é inferior ao intervalo fixo de 10 segundos utilizado pelo método da descida máxima. No entanto, em problemas com poucas variáveis ($n = 1$ ou $n = 2$), essa tendência pode se inverter. Um exemplo disso é o problema $DGO1M.$, onde o método dos pesos com coleta levou 25.33 segundos para coletar 22811 pontos, superando o tempo limite do método da descida máxima.

Isto sugere que o método dos pesos com coleta apresenta melhor desempenho relativo em problemas de maior dimensionalidade. Em geral, à medida que o tamanho do problema aumenta, o método dos pesos com coleta tende a ser mais eficiente do que o método da descida máxima. Vale destacar que, no caso extremo do problema $DGO1M.3$, que apresenta apenas uma variável e um elevado número de objetivos (100), o método dos pesos com coleta mostrou-se ligeiramente inferior ao método da descida máxima. Essa diferença pode ser atribuída ao fato de que a filtragem ocorre no espaço dos objetivos. Assim, quando o número de objetivos é significativamente maior que o número de variáveis, o tempo requerido para a filtragem impacta negativamente a eficiência temporal do método dos pesos com coleta.

Para ilustrar a eficiência relativa dos métodos na geração de soluções em relação ao tempo, definimos a seguinte métrica, denominada *Taxa de Pontos Pareto*

(TPP):

$$\text{TPP} = \frac{\text{Número total de Pontos Pareto encontrados}}{\text{Tempo de execução (em segundos)}}.$$

Essa razão indica quantos pontos Pareto, em média, cada método produz por segundo de processamento, sendo portanto um critério de eficiência temporal na geração de soluções.

Tendo calculado a TPP de cada método para os problemas listados na Tabela 3, construímos um perfil de desempenho baseado nesses valores. A Figura 16 apresenta o perfil de desempenho para os métodos analisados, considerando todos os problemas da Tabela 3.

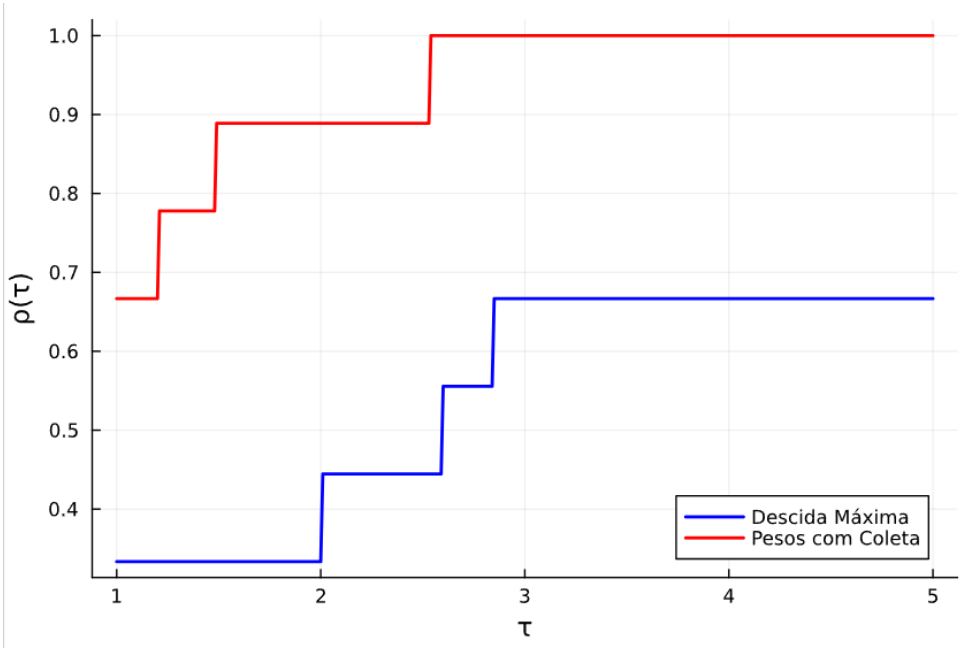

Figura 16 – Perfil de desempenho comparando os métodos com base na métrica de tempo.

4.4 Qualidade da Fronteira de Pareto

Nesta seção, iremos avaliar os métodos estudados quanto à qualidade das suas fronteiras de Pareto. Isto é, para todos os métodos apresentados no Apêndice A, calcularemos uma aproximação da fronteira de Pareto utilizando os três métodos descritos no trabalho e as compararemos quanto à métrica de espalhamento apresentada no capítulo anterior e quanto a quantidade de pontos Pareto crítico encontrados.

Para os métodos dos pesos e da descida máxima, utilizamos 500 pontos iniciais nos testes a seguir. Para o método dos pesos com coleta foram utilizados 150 pontos iniciais. O número de pontos iniciais para o método dos pesos com coleta é menor pois, dependendo do problema, podemos coletar um número extremamente alto

de pontos durante a busca linear caso o número de iterações na solução do problema ponderado seja alto demais, o que aumenta o tempo de execução do algoritmo de filtragem. Para os métodos baseados em pesos, cada ponto inicial foi associado a um peso distinto, totalizando 500 pesos para o método dos pesos convencional e 150 pesos para o método com coleta.

Para garantir a reprodutibilidade dos experimentos e assegurar que cada método fosse testado sob as mesmas condições, utilizamos uma seed fixa para o gerador de números aleatórios. Dessa forma, os pontos iniciais e os vetores de pesos sorteados para um determinado problema permanecem idênticos em todos os testes realizados, independente do método, permitindo uma comparação mais consistente dos resultados entre diferentes problemas.

Para avaliar o desempenho dos métodos em cada problema, utilizamos a métrica Γ, definida em (3.1), e contabilizamos o número de pontos Pareto crítico identificados. A Tabela 4 apresenta os resultados obtidos e contém informações detalhadas sobre cada problema analisado. Nela, a coluna **Problema** indica o nome do problema testado, enquanto as colunas n e m correspondem, respectivamente, ao número de variáveis e ao número de funções objetivo do problema. A coluna **Convexo?** informa se o problema é convexo.

As colunas relacionadas aos métodos apresentam os valores da métrica Γ e o número de pontos Pareto crítico encontrados, representado por P.C. em cada coluna. Γ **Pesos** e **P.C. Pesos** referem-se ao método dos pesos, Γ **D.M.** e **P.C. D.M.** ao método da descida máxima, enquanto Γ **Coleta** e **P.C. Coleta** correspondem ao método dos pesos com coleta.

Problema	n	m	Convexo?	Γ Pesos	P.C. Pesos	Γ D.M.	P.C. D.M.	Γ Coleta	P.C. Coleta
LOV1	2	2	Sim	0,4144	500	0,8576	500	0,6558	1450
MHHM1	1	3	Sim	0,0006	500	0,0003	500	0,0012	482
MHHM2	2	3	Sim	0.0006	500	0,0003	500	0,0016	479
SLCDT 2	10	3	Sim	88.219	500	17,16888	500	88,218	11706
DGO 1	1	2	Sim	0.0054	500	0,007596	454	0,00997	997
JOS1	100	2	Sim	0.0818	500	0,105095	500	0,19165	49408
SSFFY 1	2	2	Sim	0.1023	500	0,638398	500	0,24052	521
FAR1	2	2	Não	0.0233	164	0,019581	174	0,02739	6998
DGO1 M	1	2	Não	0.1966	500	0,012034	492	0,00953	5179
SSFFY 2	1	2	Não	9.1590	479	11,37266	60	8,90854	14602
MOP 2	15	2	Não	0.7809	409	0,555288	500	0,00993	12325
VU1	2	2	Não	10,6170	343	0,692647	429	9,83904	9382
DGO1 M 2	100	2	Não	1.6171	367	21,0101	211	0.8420	3567
DGO1 M 3	1	100	Não	0.1750	500	0,065409	500	0,00596	15273
QV1	10	2	Não	1.6868	16	1,666456	14	1,55471	629
MOP 5	2	3	Não	9.423	417	96,83	9	8,743	19005
MOP 5 M.	200	10	Não	31442	500	32103	491	25945	66183
Toi9	4	4	Não	24.72	206	2.999	489	2.845	30289
MGH26	4	4	Não	0.0202	500	0.0166	500	0.0065	7104
MGH26 M.	50	50	Não	0.0378	500	0,0292	500	0,0213	13554
FF1	2	2	Não	0.9995	500	0.0404	500	0.0390	10556

Tabela 4 – Resultados dos testes.

Podemos ver pela Tabela 4 que para os problemas convexos os métodos são, no geral, competitivos. Não há um claro vencedor quando analisamos a métrica de espalhamento. Quando analisamos a quantidade de pontos Pareto crítico encontrados, percebemos uma superioridade do método dos pesos com coleta. Apesar de termos casos como o do problema MHHM1, onde o método dos pesos com coleta encontra menos pontos Pareto crítico que os outros dois, devemos lembrar que para este método utilizamos apenas 150 pontos iniciais, desta forma estando sujeito a coletar poucos pontos quando o problema ponderado converge rápido demais. Porém, ainda encontramos em média mais de 3 pontos Pareto crítico por execução do algoritmo.

Para problemas não convexos, porém, é possível perceber uma superioridade do método dos pesos com coleta, tanto pela métrica de espalhamento quanto pelo número de pontos Pareto crítico encontrados. Vemos até casos extremos como do problema QV1, onde os métodos dos pesos e de descida máxima falham em convergir para pontos não dominados na maioria das execuções do algoritmo, encontrando apenas 16 e 14 pontos Pareto, respectivamente.

Outro caso interessante é o MOP 5, em que a métrica de espalhamento Γ apresenta valores muito distintos para cada método. O método de descida máxima, por exemplo, atinge Γ de 96,83 e encontra apenas 9 pontos Pareto críticos, enquanto o método com coleta encontra mais de 19 mil pontos críticos e um valor de Γ consideravelmente menor (8,743). A menor métrica de espalhamento indica uma fronteira mais "preenchida", logo, potencialmente com maior qualidade.

Os problemas de elaboração própria DGO1 Modificado 2 e DGO1 Modificado 3 foram criados para analisar como os métodos se comportam em situações "extremas", isto é, muitas variáveis e poucos objetivos ou poucas variáveis e muitos objetivos. Estes problemas foram elaborados utilizando DGO1 Modificado como base pois, como visto nos problemas ilustrativos, este apresenta características interessantes para cada um dos métodos estudados.

Assim, no problema DGO1 Modificado 2, vemos uma dificuldade de convergência para o método dos pesos e para o método da descida máxima, resultando em números piores tanto para a métrica de espalhamento quanto para o número de pontos Pareto crítico encontrados quando comparados ao método dos pesos com coleta. Isto sugere uma vantagem do método dos pesos com coleta para problemas com muitas variáveis e poucos objetivos.

No problema DGO1 Modificado 3, com 100 objetivos, a alta dimensionalidade do espaço objetivo influencia fortemente o desempenho dos métodos. Enquanto o método dos pesos com coleta gera um conjunto muito maior de pontos, o custo de filtragem cresce à medida que mais funções objetivo aumentam o custo de verificação de dominância. Apesar de proporcionar vantagens notáveis na métrica de espalhamento, é preciso considerar o ônus computacional adicional decorrente do processamento de um número significativamente maior de pontos.

O desempenho dos métodos é também ilustrado nas Figuras 17 e 18, com perfis de desempenho. A Figura 17 mostra o desempenho dos métodos para todos os problemas exibidos na Tabela 4. A Figura 18 se limita a mostrar o desempenho para os problemas não convexos da tabela.

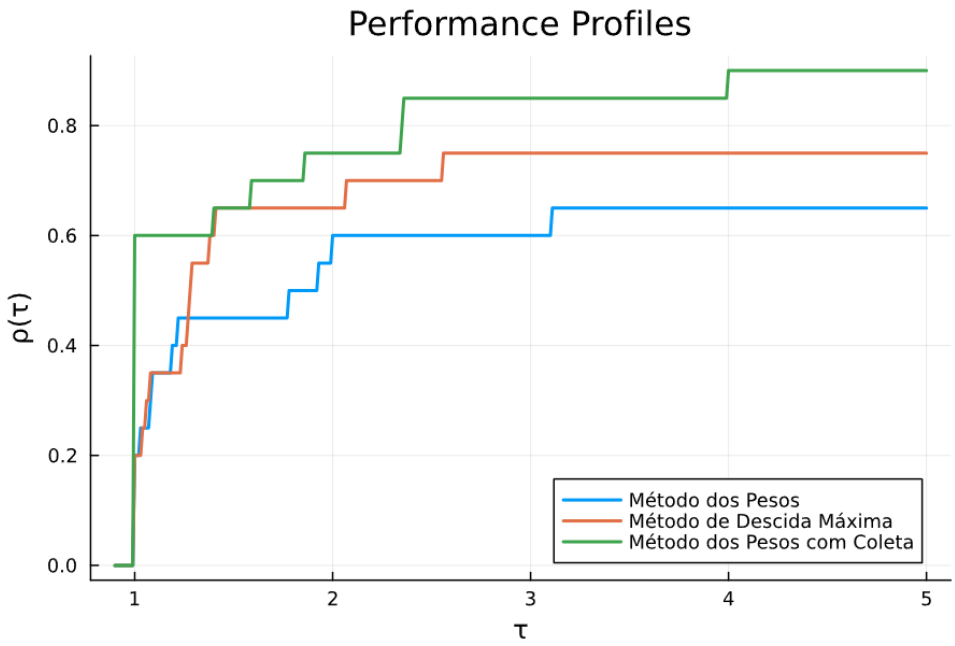

Figura 17 – Perfil de desempenho comparando os métodos com base na métrica de espalhamento Γ.

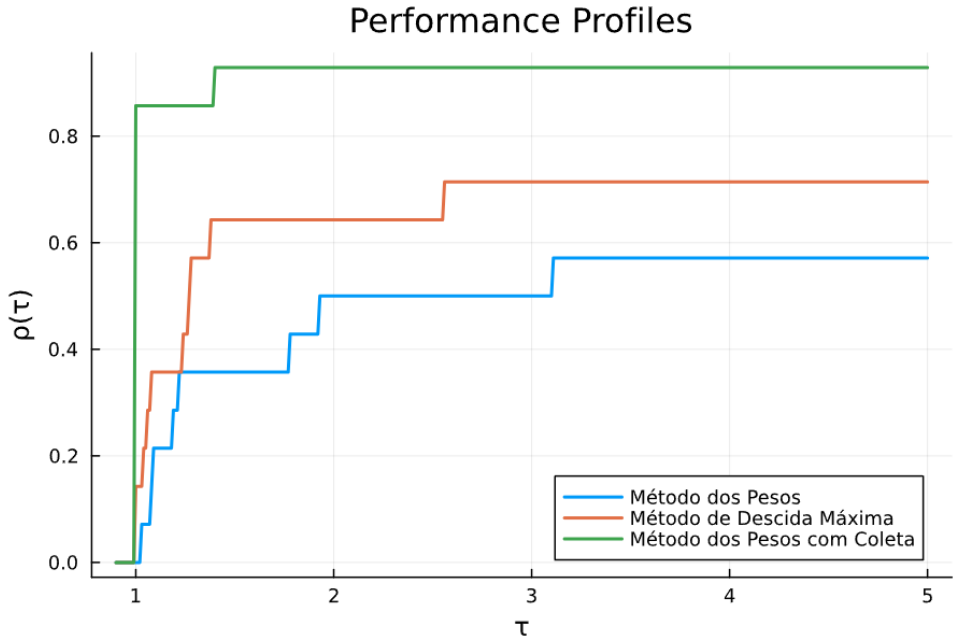

Figura 18 – Perfil de desempenho comparando os métodos com base na métrica de espalhamento Γ para os problemas não convexos.

5 Considerações Finais

O presente trabalho teve como objetivo contribuir para o estudo teórico e computacional em *otimização multiobjetivo*, cobrindo desde a formalização rigorosa do problema até a implementação de métodos numéricos em linguagem Julia. Inicialmente, definiram-se conceitos fundamentais como *otimalidade de Pareto* (em suas versões fraca, forte e própria) e *Pareto criticidade*, destacando sua importância para caracterizar pontos "ótimos" no sentido multiobjetivo.

Do ponto de vista prático, foram investigados três métodos principais. O **método dos pesos** mostrou-se eficaz para problemas convexos, embora seja limitado quando a fronteira de Pareto apresenta regiões côncavas. Já o **método de descida máxima** trouxe a vantagem de uma interpretação direta de ponto estacionário no contexto multiobjetivo, mas as vezes convergiu para pontos dominados e pontos de máximo. Por fim, o **método dos pesos com coleta** surgiu como proposta autoral ao aliar a flexibilidade do método dos pesos à coleta de pontos adicionais durante a busca linear, mitigando os problemas de convergência em zonas côncavas e aumentando a cobertura na fronteira de Pareto.

Os experimentos numéricos realizados abordaram problemas de diferentes dimensões, número de objetivos e graus de convexidade. As métricas de espalhamento, pureza e custo computacional forneceram elementos quantitativos para comparar os métodos, evidenciando que a estratégia de coleta de pontos e posterior filtragem resulta em maior diversidade de soluções, em especial para problemas não convexos. Apesar de demandar maior cuidado na implementação, o *método dos pesos com coleta* mostrou-se competitivo também em instâncias convexas.

Como principais **limitações** do trabalho, destaca-se que, à medida que a dimensão do problema aumenta, os critérios de parada adotados podem não ser suficientemente eficientes para garantir boa cobertura da fronteira de Pareto sem incorrer em alto custo computacional, exigindo, portanto, maior investigação. Além disso, ressalta-se o impacto de parâmetros como a tolerância de coleta no desempenho do método dos pesos com coleta. Reconhecemos também que a fundamentação teórica desse método ainda requer um exame mais aprofundado, uma vez que não há prova formal de que ele consiga, em todas as situações, coletar um ponto Pareto crítico em regiões de concavidade.

Conclui-se, portanto, que os métodos abordados contribuem de forma significativa para a resolução de problemas multiobjetivo, sobretudo em cenários onde a não convexidade impede o uso eficaz de técnicas mais clássicas. A abordagem

proposta, ao incorporar um mecanismo de coleta de pontos, abre caminho para futuras melhorias e pesquisas, fomentando o desenvolvimento de métodos cada vez mais robustos e abrangentes na otimização multiobjetivo.

Referências

1 MARLER, R.; ARORA, J. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, v. 26, n. 6, p. 369–395, apr 2004. ISSN 1615-1488. Disponível em: https://doi.org/10.1007/s00158-003-0368-6. Citado na página 12.

2 KAISA, M. *Nonlinear Multiobjective Optimization*. Boston, USA: Kluwer Academic Publishers, 1999. v. 12. (International Series in Operations Research & Management Science, v. 12). Citado 4 vezes nas páginas 15, 18, 34 e 35.

3 EHRGOTT, M. *Multicriteria Optimization*. Berlin, Heidelberg: Springer-Verlag, 2005. ISBN 3540213988. Citado na página 15.

4 ANDREANI, R.; RAMIREZ, V. A.; SANTOS, S. A.; SECCHIN, L. D. Bilevel optimization with a multiobjective problem in the lower level. *Numerical Algorithms*, v. 81, n. 3, p. 915–946, 2019. Citado na página 22.

5 KUHN, H. W.; TUCKER, A. W. Nonlinear programming. In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. Berkeley and Los Angeles: University of California Press, 1951. p. 481–492. Citado 2 vezes nas páginas 25 e 29.

6 GEOFFRION, A. M. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, v. 22, n. 3, p. 618–630, 1968. ISSN 0022-247X. Disponível em: https://www.sciencedirect.com/science/article/pii/0022247X68902011. Citado na página 25.

7 GASS, S.; SAATY, T. The computational algorithm for the parametric objective function. *Naval Research Logistics Quarterly*, v. 2, n. 1-2, p. 39–45, 1955. Disponível em: https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.3800020106. Citado na página 32.

8 ZADEH, L. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, v. 8, n. 1, p. 59–60, 1963. Citado na página 32.

9 PRUDENTE, L. F.; FERREIRA, O. P.; ASSUNçãO, P. B. Conditional gradient method for multiobjective optimization. *Computational Optimization and Applications*, v. 78, 2021. Citado 2 vezes nas páginas 32 e 54.

10 CARRIZO, G. A.; FAZZIO, N. S.; SCHUVERDT, M. L. A nonmonotone projected gradient method for multiobjective problems on convex sets. *Journal of the Operations Research Society of China*, Springer Science and Business Media LLC, v. 12, n. 2, p. 410–427, abr. 2022. ISSN 2194-6698. Disponível em: http://dx.doi.org/10.1007/s40305-022-00410-y. Citado na página 32.

11 FLIEGE, J.; DRUMMOND, L. M. G.; SVAITER, B. F. Newton's method for multiobjective optimization. *SIAM Journal on Optimization*, Society for Industrial Applied Mathematics (SIAM), v. 20, n. 2, p. 602–626, jan. 2009. ISSN 1095-7189. Disponível em: http://dx.doi.org/10.1137/08071692X. Citado na página 32.

12 GONcALVES, M.; LIMA, F.; PRUDENTE, L. A study of liu-storey conjugate gradient methods for vector optimization. *Applied Mathematics and Computation*, Elsevier BV, v. 425, p. 127099, jul. 2022. ISSN 0096-3003. Disponível em: http://dx.doi.org/10.1016/j.amc.2022.127099. Citado na página 32.

13 GONcALVES, M. L. N.; LIMA, F. S.; PRUDENTE, L. F. Globally convergent newton-type methods for multiobjective optimization. *Computational Optimization and Applications*, Springer Science and Business Media LLC, v. 83, n. 2, p. 403–434, set. 2022. ISSN 1573-2894. Disponível em: http://dx.doi.org/10.1007/s10589-022-00414-7. Citado na página 32.

14 GONcALVES, M. L. N.; PRUDENTE, L. F. On the extension of the hager–zhang conjugate gradient method for vector optimization. *Computational Optimization and Applications*, Springer Science and Business Media LLC, v. 76, n. 3, p. 889–916, out. 2019. ISSN 1573-2894. Disponível em: http://dx.doi.org/10.1007/s10589-019-00146-1. Citado na página 32.

15 PéREZ, L. R. L.; PRUDENTE, L. F. Nonlinear conjugate gradient methods for vector optimization. *SIAM Journal on Optimization*, Society for Industrial Applied Mathematics (SIAM), v. 28, n. 3, p. 2690–2720, jan. 2018. ISSN 1095-7189. Disponível em: http://dx.doi.org/10.1137/17M1126588. Citado na página 32.

16 MOROVATI, V.; POURKARIMI, L.; BASIRZADEH, H. Barzilai and borwein's method for multiobjective optimization problems. *Numerical Algorithms*, Springer Science and Business Media LLC, v. 72, n. 3, p. 539–604, nov. 2015. ISSN 1572-9265. Disponível em: http://dx.doi.org/10.1007/s11075-015-0058-7. Citado na página 32.

17 PRUDENTE, L. F.; SOUZA, D. R. A quasi-newton method with wolfe line searches for multiobjective optimization. *Journal of Optimization Theory and Applications*, Springer Science and Business Media LLC, v. 194, n. 3, p. 1107–1140, jul. 2022. ISSN 1573-2878. Disponível em: http://dx.doi.org/10.1007/s10957-022-02072-5. Citado na página 32.

18 RAMIREZ, V. A.; SOTTOSANTO, G. N. Nonmonotone trust region algorithm for solving the unconstrained multiobjective optimization problems. *Computational Optimization and Applications*, Springer Science and Business Media LLC, v. 81, n. 3, p. 769–788, jan. 2022. ISSN 1573-2894. Disponível em: http://dx.doi.org/10.1007/s10589-021-00346-8. Citado na página 32.

19 FLIEGE, J.; SVAITER, B. Steepest descent methods for multicriteria optimization. *Math Methods Oper Res*, v. 51, p. 479–494, 08 2000. Citado 2 vezes nas páginas 32 e 36.

20 CUSTÓDIO, A. L.; MADEIRA, J. F. A.; VAZ, A. I. F.; VICENTE, L. N. Direct multisearch for multiobjective optimization. *SIAM Journal on Optimization*, v. 21, n. 3, p. 1109–1140, 2011. Disponível em: https://doi.org/10.1137/10079731X. Citado 3 vezes nas páginas 41, 45 e 46.

21 ANDREANI, R.; CUSTóDIO, A. L.; RAYDAN, M. Using first-order information in direct multisearch for multiobjective optimization. *Optimization Methods and Software*, Taylor & Francis, v. 37, n. 6, p. 2135–2156, 2022. Disponível em: https://doi.org/10.1080/10556788.2022.2060971. Citado na página 41.

22 BEZANSON, J.; EDELMAN, A.; KARPINSKI, S.; SHAH, V. B. Julia: A fresh approach to numerical computing. *SIAM Review*, v. 59, n. 1, p. 65–98, 2017. Disponível em: https://doi.org/10.1137/141000671. Citado na página 43.

23 WILLMS, A. R. Uniform sampling on the standard simplex. *Missouri Journal of Mathematical Sciences*, University of Central Missouri, Department of Mathematics and Computer Science, v. 33, n. 1, maio 2021. ISSN 0899-6180. Disponível em: http://dx.doi.org/10.35834/2021/3301119. Citado 2 vezes nas páginas 43 e 44.

24 DOLAN, E. D.; MORÉ, J. J. Benchmarking optimization software with performance profiles. *Mathematical Programming*, v. 91, n. 2, p. 201–213, 2002. Citado na página 45.

25 LOVISON, A. Singular continuation: Generating piecewise linear approximations to pareto sets via global analysis. *SIAM Journal on Optimization*, v. 21, n. 2, p. 463–490, 2011. Disponível em: https://doi.org/10.1137/100784746. Citado na página 68.

26 HUBAND, S.; HINGSTON, P.; BARONE, L.; WHILE, L. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE Transactions on Evolutionary Computation*, v. 10, n. 5, p. 477–506, 2006. Citado 2 vezes nas páginas 68 e 69.

27 SCHüTZE, O.; LAUMANNS, M.; COELLO, C. A. C.; DELLNITZ, M.; TALBI, E.-G. Convergence of stochastic search algorithms to finite size pareto set approximations. *Journal of Global Optimization*, Springer Science+Business Media, v. 41, n. 4, p. 559–577, 2008. ISSN 1573-2916. Disponível em: https://doi.org/10.1007/s10898-007-9265-7. Citado na página 68.

28 JIN, Y.; OLHOFER, M.; SENDHOFF, B. Dynamic weighted aggregation for evolutionary multi-objective optimization: Why does it work and how? 04 2001. Citado na página 68.

29 TOINT, P. L. *Test problems for partially separable optimization and results for the routine pspmin*. [S.l.], 1983. Citado na página 69.

30 MORé, J. J.; GARBOW, B. S.; HILLSTROM, K. E. Testing unconstrained optimization software. *ACM Trans. Math. Softw.*, Association for Computing Machinery, New York, NY, USA, v. 7, n. 1, p. 17–41, mar. 1981. ISSN 0098-3500. Disponível em: https://doi.org/10.1145/355934.355936. Citado na página 69.

Apêndices

APÊNDICE A – Problemas Teste

Tabela 5 – Problemas teste (Parte 1)

Nome	Problema	$[a,b]$	Referência
LOV1	$f_1(x) = 1.05x_1^2 + 0.98x_2^2$ $f_2(x) = 0.99(x_1-3)^2 + 1.03(x_2-2.5)^2$	$[-10,10]$	[25]
MHHM1	$f_i(x) = (x_1 - 0.8 + 0.05i)^2$ para $i \in \{1,2,3\}$	$[0,2]$	[26]
MHHM2	$f_i(x) = (x_1 - 0.8 + 0.05i)^2 + (x_2 - 0.6 + 0.1i)^2$ para $i \in \{1,2,3\}$	$[0,2]$	[26]
SLCDT2	$f_i(x) = \sum_{\substack{j=1 \\ j\neq i}}^{10} (x_j - a_j^i)^2 + (x_i - a_i^i)^4$ para $i \in \{1,\cdots,10\}$ $a^1 = (1,1,1,1,\dots) \in \mathbb{R}^{10}$, $a^2 = (-1,-1,-1,-1,\dots) \in \mathbb{R}^{10}$, $a^3 = (1,-1,1,-1,\dots) \in \mathbb{R}^{10}$.	$[-1,1]$	[27]
DGO1	$f_1(x) = \text{sen}(x_1)$ $f_2(x) = \text{sen}(x_1 + 0.7)$	$[-10,13]$	[26]
JOS1	$f_1(x) = \sum_{i=1}^{100} \frac{(x_i)^2}{100}$ $f_2(x) = \sum_{i=1}^{100} \frac{(x_i-2)^2}{100}$	$[-100,100]$	[28]
SSFFY1	$f_1(x) = x_1^2 + x_2^2$ $f_2(x) = (x_1-1)^2 + (x_2-2)^2$	$[-100,100]$	[26]
FAR1	$f_1(x_1,x_2) = -2\exp\left(15(-(x_1-0.1)^2 - x_2^2)\right)$ $-\exp\left(20(-(x_1-0.4)^2 - (x_2-0.6)^2)\right)$ $+\exp\left(20(-(x_1+0.6)^2 - (x_2-0.6)^2)\right)$ $+\exp\left(20(-(x_1-0.6)^2 - (x_2+0.6)^2)\right)$ $+\exp\left(20(-(x_1+0.6)^2 - (x_2+0.6)^2)\right)$ $f_2(x_1,x_2) = 2\exp\left(20(-x_1^2 - x_2^2)\right)$ $+\exp\left(20(-(x_1-0.4)^2 - (x_2-0.6)^2)\right)$ $-\exp\left(20(-(x_1+0.5)^2 - (x_2-0.7)^2)\right)$ $-\exp\left(20(-(x_1-0.5)^2 - (x_2+0.7)^2)\right)$ $+\exp\left(20(-(x_1+0.4)^2 - (x_2+0.8)^2)\right)$	$[-1,1]$	[26]
DGO1 M.	$f_1(x) = \text{sen}(x_1) + 0.1 \cdot \cos(5x_1) \cdot \text{sen}(x_1)$ $f_2(x) = \text{sen}(x_1+0.7) + 0.1 \cdot \cos(5x_1) \cdot \text{sen}(x_1+0.7)$	$[-10,13]$	Autoral
SSFFY2	$f_1(x) = 10 - 10\cos\left(\frac{x_1\pi}{2}\right) + x_1^2$ $f_2(x) = (x_1-4)^2$	$[-100,100]$	[26]
MOP2	$f_1(x) = 1 - \exp\left(-\frac{\sum_{j=1}^{15}(x_j-1)^2}{15}\right)$ $f_2(x) = 1 - \exp\left(-\frac{\sum_{j=1}^{15}(x_j+1)^2}{15}\right)$	$[-4,4]$	[26]
VU1	$f_1(x) = \frac{1}{x_1^2 + x_2^2 + 1}$ $f_2(x) = x_1^2 + 3x_2^2 + 1$	$[-3,3]$	[26]

Tabela 6 – Problemas teste (Parte 2)

Nome	**Problema**	$[a, b]$	**Referência**
DGO1 M. 2	$f_1(x) = \sum_{i=1}^{100} [\text{sen}(x_i) + 0.1\cos(5x_i)\text{sen}(x_i)]$ $f_2(x) = \sum_{i=1}^{100} [\text{sen}(x_i + 0.7) + 0.1\cos(5x_i)\text{sen}(x_i + 0.7)]$	$[-1, 1]$	Autoral
DGO1 M. 3	$f_i(x) = \text{sen}(x_1 + 0.7i) + 0.1\cos(5x_1)\text{sen}(x_1 + 0.7i)$ para $i \in \{1, \ldots, 100\}$	$[-1, 1]$	Autoral
QV1	$f_1(x) = \left[\frac{1}{10}\sum_{i=1}^{10}\left(x_i^2 - 10\cos(2\pi x_i) + 10\right)\right]^{0.25}$ $f_2(x) = \left[\frac{1}{10}\sum_{i=1}^{10}\left((x_i - 1.5)^2 - 10\cos(2\pi(x_i - 1.5)) + 10\right)\right]^{0.25}$	$[-5.12, 5.12]$	[26]
MOP5	$f_1(x) = 0.5(x_1^2 + x_2^2) + \text{sen}(x_1^2 + x_2^2)$ $f_2(x) = \frac{(3x_1 - 2x_2 + 4)^2}{8} + \frac{(x_1 - x_2 + 1)^2}{27} + 15$ $f_3(x) = \frac{1}{x_1^2 + x_2^2 + 1} - 1.1\exp(-x_1^2 - x_2^2)$	$[-30, 30]$	[26]
MOP5 M.	$f_1(x) = 0.5\sum_{i=1}^{200} x_i^2 + \text{sen}\left(\sum_{i=1}^{200} x_i^2\right)$ $f_2(x) = \frac{(3x_1 - 2x_2 + 4)^2}{8} + \frac{(x_1 - x_2 + 1)^2}{27} + 15$ $f_3(x) = \frac{1}{\sum_{i=1}^{200} x_i^2 + 1} - 1.1\exp\left(-\sum_{i=1}^{200} x_i^2\right)$	$[-30, 30]$	Autoral
Toi9	$f_1(x) = (2x_1 - 1)^2 + x_2^2$ $f_i(x) = i(2x_{i-1} - x_i)^2 - (i - 1)x_{i-1}^2 + ix_i^2, \quad i = 2, 3$ $f_4(x) = 4(2x_3 - x_4)^2 - 3x_3^2$	$[-1, 1]$	[29]
mgh26	$f_i(x) = 4 - \sum_{j=1}^{4}\cos(x_j) + i(1 - \cos(x_i)) - \text{sen}(x_i)$ para $i \in \{1, \ldots, 4\}$	$[-1, 1]$	[30]
mgh26 M.	$f_i(x) = 50 - \sum_{j=1}^{50}\cos(x_j) + i(1 - \cos(x_i)) - \text{sen}(x_i)$ para $i \in \{1, \ldots, 50\}$	$[-1, 1]$	Autoral
FF1	$f_1(x) = 1 - \exp(-((x_1 - 1)^2 + (x_2 + 1)^2))$ $f_2(x) = 1 - \exp(-((x_1 + 1)^2 + (x_2 - 1)^2))$	$[-1, 1]$	[26]