

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

Thomás Henrique Lopes Silva

REDUÇÃO DE OBJETIVOS EM MAOPS POR MEIO DE APRENDIZADO DE MÁQUINA NÃO
SUPERVISIONADO: UMA ABORDAGEM COM SELEÇÃO E EXTRAÇÃO DE ATRIBUTOS



Belo Horizonte
2024



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO CURSO DE MESTRADO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

***“REDUÇÃO DE OBJETIVOS EM MAOPS POR MEIO DE
APRENDIZADO DE MÁQUINA NÃO SUPERVISIONADO: UMA
ABORDAGEM COM SELEÇÃO E EXTRAÇÃO DE ATRIBUTOS”***

Dissertação de Mestrado apresentada por **Thomás Henrique Lopes Silva**, em 30 de outubro de 2024, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:



Documento assinado digitalmente

ELIZABETH FIALHO WANNER

Data: 18/06/2025 05:56:46-0300

Verifique em <https://validar.iti.gov.br>

Prof^ª. Dr^a. Elizabeth Fialho Wanner
Centro Federal de Educação Tecnológica de Minas Gerais



Documento assinado digitalmente

DENIS EMANUEL DA COSTA VARGAS

Data: 23/06/2025 15:26:16-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Dênis Emanuel da Costa Vargas
Centro Federal de Educação Tecnológica de Minas Gerais



Documento assinado digitalmente

CLAUDIO LUCIO DO VAL LOPES

Data: 18/06/2025 09:16:57-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Cláudio Lúcio do Val Lopes
A3Data



Documento assinado digitalmente

FLAVIO VINICIUS CRUZEIRO MARTINS

Data: 23/06/2025 11:07:46-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Flávio Vinícius Cruzeiro Martins
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof. Dr. Thiago de Souza Rodrigues
Coordenador do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

S586r Silva, Thomás Henrique Lopes
Redução de objetivos em MaOPs por meio de aprendizado de máquina não supervisionado: uma abordagem com seleção e extração de atributos / Thomás Henrique Lopes Silva. – 2024.
153 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientadora: Elizabeth Fialho Wanner.
Coorientador: Dênis Emanuel da Costa Vargas.
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Otimização matemática – Teses. 2. Redução de dimensão (Estatística)
3. Aprendizado do computador – Teses. 4. Computação evolutiva – Teses.
5. Análise de algoritmos e complexidade de problemas – Teses. I. Warner, Elizabeth Fialho. II. Vargas, Dênis Emanuel da Costa. III. Centro Federal de Educação Tecnológica de Minas Gerais. IV. Título.

CDD 519.6

Agradecimentos

Minha trajetória na busca por este título começou durante a pandemia de COVID-19, um período caótico em minha vida pessoal. No entanto, sabia que essa escolha representava uma grande oportunidade para adquirir mais conhecimento e conhecer pessoas incríveis e inspiradoras. O caminho foi árduo, como eu já previa ao decidir trilhá-lo. Hoje, sou uma pessoa muito diferente, e tenho certeza de que o Thomás do passado teria orgulho do Thomás que escreve estas palavras. A retomada das atividades presenciais foi um tanto quanto estranha, ainda mais por se tratar de um ambiente totalmente novo, em uma cidade que mal conhecia e tão grande como Belo Horizonte. Nesse ambiente, tive a honra de conhecer pessoas que levarei comigo para sempre, e que continuarão sendo fonte de inspiração. Esta jornada não termina aqui, e desejo que venham mais aprendizados e novas experiências!

Minha gratidão à minha mãe, Tânia, não pode ser medida em palavras. Seu esforço e dedicação para garantir o melhor para mim, meu irmão e meu pai são algo que nem em inúmeras vidas eu conseguiria retribuir. Seu amor permanecerá eternamente em mim, e me esforço diariamente para transmiti-lo às pessoas que cruzam o meu caminho. Ao meu irmão Lucas, que apesar de mais novo, me ensina todos os dias e é uma grande inspiração, visto que alcançar meus sonhos é, de certa forma, ajudá-lo a realizar os dele. Ao meu pai, Carlos, que mesmo com as poucas oportunidades que teve, nunca desistiu e se esforçou ao máximo para entregar tudo que pôde à nossa família. Além das necessidades materiais, também manifesto aqui meu amor e eterna gratidão.

À minha orientadora Beth, meu imensurável agradecimento por ter aceitado me orientar desde o início desta jornada. Ela, além de uma excelente professora, é uma verdadeira inspiração em todos os aspectos da vida profissional. Como ela mesma diz, e como os outros alunos podem confirmar: uma verdadeira mãe no ambiente acadêmico. Ao meu coorientador Dênis, agradeço imensamente por toda a disposição em me ajudar com as questões técnicas e desafios ao longo do percurso, sua orientação foi fundamental para que eu seguisse em

frente. Espero poder continuar colaborando com ambos por muito tempo.

Aos meus amigos Raul, Renê e Wanderson, que mesmo à distância estiveram presentes todos os dias, compartilhando experiências, discussões e muitas risadas pela internet. Ao querido Daniel, com quem tive o prazer de dividir esta caminhada, agradeço pelas conversas e pelas histórias compartilhadas, tenho certeza de que será uma amizade para a vida toda. Ao Vanderci, pela parceria e boas conversas no laboratório de sistemas inteligentes, meu muito obrigado.

À minha amiga Aryane e seu cachorrinho Raul, que me ensinaram o verdadeiro significado de dizer “eu te amo”. Agradeço pela convivência, por me conhecerem profundamente e por acreditarem em mim, mesmo quando lidar com minha intimidade era desafiador. Obrigado por me ajudarem a entender os sentimentos, inclusive os mais sombrios. Minha gratidão por essa experiência, breve, porém de enorme valor. À Raquel, outra amizade valiosa conquistada em Belo Horizonte, que todos os dias me lembra do poder das mudanças positivas que a vida pode trazer, tanto pessoais quanto sociais. À minha psicóloga Marília, que, além de seu trabalho profissional, me tratou com compaixão e me mostrou o caminho para lidar com meus sentimentos. Ainda há um longo percurso no caminho do autoconhecimento, mas o que você me proporcionou até aqui é inestimável. Sem sua ajuda, este trabalho não teria sido possível.

Agradeço também aos colegas de curso, ao CEFET-MG e ao Programa de Pós-Graduação em Modelagem Matemática e Computacional (PPGMMC), por fornecerem o ambiente ideal para a realização deste trabalho. Aos técnicos e funcionários do campus, que sempre estiveram dispostos a auxiliar, e à CAPES pelo apoio financeiro. À banca examinadora, meus sinceros agradecimentos pelas valiosas contribuições.

A todas as pessoas que acreditam na educação inclusiva como uma forma de libertação, e a você, leitor, que me concede seu tempo ao mergulhar neste texto. Este trabalho representa uma pequena parte da minha história, curta diante da imensidão do Cosmos. Porém, essa imensidão se torna menor diante de pessoas tão queridas que fizeram parte da minha jornada. A todos vocês, meu muito obrigado, do fundo do coração.

“Girando com esta parábola familiar
Tecendo em torno de cada nova experiência
Reconheça isso como um presente sagrado e
Celebre esta chance de estar vivo e respi-
rando.”

Parabola, TOOL. Tradução livre.

RESUMO

Este trabalho aborda os desafios de resolver problemas de otimização com muitos objetivos (MaOPs), que envolvem a otimização simultânea de mais de três funções objetivo conflitantes. Para contornar esses desafios, propôs-se uma abordagem baseada em técnicas de redução de dimensionalidade utilizando aprendizado de máquina (AM), ou seja, reduzir o número de objetivos dos MaOPs sem perder informações relevantes. De acordo com trabalhos relacionados da literatura, esse tipo de abordagem mostrou benefícios significativos, assim como a possibilidade de resolver MaOPs difíceis com algoritmos evolutivos multiobjetivo (MOEAs) tradicionais e melhorias na escalabilidade desses algoritmos. Além disso, a redução de dimensionalidade auxilia na tomada de decisões e visualização dos resultados pelos usuários. Para isto, foram aplicadas duas abordagens aos conjuntos de dados extraídos dos problemas de teste *benchmark* da família DTLZ: uma de seleção e outra de extração de atributos, *Laplacian Score* (LS) e Análise de Componentes Principais (PCA), respectivamente. Em seguida, as novas funções reduzidas foram otimizadas pelo MOEA considerado estado da arte: NSGA-III. As soluções obtidas foram então expandidas para o número inicial de objetivos, a fim de compará-las com as soluções dos problemas originais utilizando os indicadores de qualidade (IQs) IGD^+ e $I_{\epsilon+}$, além de seus respectivos *boxplots*. Para avaliar estatisticamente os ganhos de desempenho das abordagens propostas, foi realizado o Teste de Postos Sinalizados de Wilcoxon. Os resultados mostraram que a técnica PCA se destacou como mais eficaz, particularmente em cenários de alta dimensionalidade, enquanto o LS teve desempenho mais limitado, sendo eficiente apenas em alguns DTLZ com menor número de objetivos. Por fim, conclui-se para o conjunto de experimentos que os métodos *offline* possuem limitações e novos estudos devem ser desenvolvidos para garantir a robustez das abordagens propostas.

Palavras-chave: Problemas de otimização com muitos objetivos; Redução de dimensionalidade; Aprendizado de máquina; Seleção de atributos; Extração de atributos; Algoritmos evolutivos multiobjetivo; Problemas de teste *benchmark* da família DTLZ.

ABSTRACT

This work addresses the challenges of solving many-objective optimization problems (MaOPs), which involve the simultaneous optimization of more than three conflicting objective functions. To overcome these challenges, a machine learning (ML)-based dimensionality reduction approach is proposed, aiming to reduce the number of objectives in MaOPs without losing relevant information. According to related works in the literature, this type of approach has shown significant benefits, such as the possibility of solving difficult MaOPs with traditional multi-objective evolutionary algorithms (MOEAs) and improvements in the scalability of these algorithms. Additionally, dimensionality reduction helps in decision-making and in visualizing the results for users. To this end, two approaches were applied to the data sets extracted from the DTLZ family's *benchmark* test problems: one for attribute selection and another for attribute extraction, Laplacian Score (LS) and Principal Component Analysis (PCA), respectively. The reduced functions were then optimized using the state-of-the-art MOEA: NSGA-III. The obtained solutions were subsequently expanded to the original number of objectives in order to compare them with the original problem solutions using quality indicators (QIs) IGD^+ and $I_{\epsilon+}$, along with their respective boxplots. To statistically evaluate the performance improvements of the proposed approaches, the Wilcoxon Signed-Rank Test was performed. The results showed that PCA technique stood out as more effective, particularly in high-dimensional scenarios, while LS had a more limited performance, being efficient only in some DTLZ with fewer objectives. Lastly, it is concluded for the set of experiments that the offline methods have limitations, and further studies are needed to ensure the robustness of the proposed approaches.

Keywords: Many-objective optimization problems; Dimensionality reduction; Machine learning; Feature selection; Feature extraction; Multi-objective evolutionary algorithms; DTLZ benchmark test problems.

Sumário

Lista de Figuras	xiv
Lista de Tabelas	xx
Lista de Abreviaturas e Siglas	xxi
1 Introdução	1
1.1 Descrição do Problema	3
1.2 Motivação	4
1.3 Objetivos	5
1.4 Organização do Documento	6
2 Fundamentação Teórica	7
2.1 Problema de Otimização Multiobjetivo	7
2.1.1 Problemas de Teste <i>Benchmark</i> da Família DTLZ	10
2.2 Algoritmos Evolutivos Multiobjetivo	20
2.2.1 NSGA-III	24
2.3 Indicadores de Qualidade	27
2.3.1 <i>Inverted Generational Distance Plus</i>	27
2.3.2 Épsilon-Indicador Aditivo Binário	30
2.4 Desafios para Solucionar os Problemas com Muitos Objetivos	32
2.4.1 Baixa Escalabilidade	33
2.4.2 Alto Custo Computacional	34

2.4.3	Dificuldade de Visualização das Soluções	34
2.5	Redução de Dimensionalidade	35
2.5.1	Laplacian Score	37
2.5.2	Análise de Componentes Principais	42
2.6	Formulação do Teste de Hipóteses	46
3	Trabalhos Relacionados	47
3.1	Revisão Bibliográfica	47
3.1.1	Baseados em Correlação	48
3.1.2	Baseados em Estrutura de Dominância	49
3.1.3	Baseados em Seleção de Atributos	50
3.1.4	Outras Publicações	52
3.2	Estado da Arte	53
4	Experimentos Computacionais	56
4.1	Parâmetros Atribuídos aos Algoritmos	56
4.2	Laplacian Score: Análise dos Resultados	58
4.3	PCA: Análise dos Resultados	59
4.4	Comparação dos Resultados de LS e PCA	61
5	Conclusões	64
5.1	Trabalhos Futuros	67
	Referências Bibliográficas	69
A	Resultados Detalhados da Seleção de Atributos: LS	78
A.1	DTLZ1	78
A.2	DTLZ2	83
A.3	DTLZ3	88
A.4	DTLZ4	93
A.5	DTLZ5	98
A.6	DTLZ6	104
A.7	DTLZ7	109

B	Resultados Detalhados da Extração de Atributos: PCA	115
B.1	DTLZ1	115
B.2	DTLZ2	120
B.3	DTLZ3	126
B.4	DTLZ4	131
B.5	DTLZ5	136
B.6	DTLZ6	142
B.7	DTLZ7	147

Lista de Figuras

2.1	Mapeamento do espaço de decisão no espaço de objetivos. Adaptado de: Mendes (2022).	8
2.2	Abstração da ideia de dominância. Fonte: Mendes (2022).	9
2.3	Visualização da PF gerada para o DTLZ1 com 3 funções objetivo.	12
2.4	Visualização da PF gerada para o DTLZ2 com 3 funções objetivo.	13
2.5	Visualização da PF gerada para o DTLZ3 com 3 funções objetivo.	14
2.6	Visualização da PF gerada para o DTLZ4 com 3 funções objetivo.	15
2.7	Visualização da PF gerada para o DTLZ5 com 3 funções objetivo.	17
2.8	Visualização da PF gerada para o DTLZ6 com 3 funções objetivo.	18
2.9	Visualização da PF gerada para o DTLZ7 com 3 funções objetivo.	19
2.10	Exemplo de um par de conjuntos de solução A e B onde A é melhor que B em termos de dominância de Pareto. Adaptado de: Ishibuchi et al. (2016). .	29
2.11	ϵ -Dominância. Fonte: Zitzler et al. (2003).	31
2.12	Aumento das regiões não dominadas. Fonte: Mendes (2022).	33
2.13	Um exemplo ilustrativo de atributos relevantes, redundantes e irrelevantes. Adaptado de: Li et al. (2017).	36
2.14	Grafo correspondente ao exemplo.	39
2.15	Exemplo PCA.	44
A.1	<i>Borplots</i> dos IQs: DTLZ1 por seleção com $m = 5$	79
A.2	<i>Borplots</i> dos IQs: DTLZ1 por seleção com $m = 10$	80
A.3	<i>Borplots</i> dos IQs: DTLZ1 por seleção com $m = 15$	81

A.4	<i>Boxplots</i> dos IQs: DTLZ1 por seleção com $m = 20$.	82
A.5	<i>Boxplots</i> dos IQs: DTLZ1 por seleção com $m = 30$.	83
A.6	<i>Boxplots</i> dos IQs: DTLZ2 por seleção com $m = 5$.	84
A.7	<i>Boxplots</i> dos IQs: DTLZ2 por seleção com $m = 10$.	85
A.8	<i>Boxplots</i> dos IQs: DTLZ2 por seleção com $m = 15$.	86
A.9	<i>Boxplots</i> dos IQs: DTLZ2 por seleção com $m = 20$.	87
A.10	<i>Boxplots</i> dos IQs: DTLZ2 por seleção com $m = 30$.	88
A.11	<i>Boxplots</i> dos IQs: DTLZ3 por seleção com $m = 5$.	89
A.12	<i>Boxplots</i> dos IQs: DTLZ3 por seleção com $m = 10$.	90
A.13	<i>Boxplots</i> dos IQs: DTLZ3 por seleção com $m = 15$.	91
A.14	<i>Boxplots</i> dos IQs: DTLZ3 por seleção com $m = 20$.	92
A.15	<i>Boxplots</i> dos IQs: DTLZ3 por seleção com $m = 30$.	93
A.16	<i>Boxplots</i> dos IQs: DTLZ4 por seleção com $m = 5$.	94
A.17	<i>Boxplots</i> dos IQs: DTLZ4 por seleção com $m = 10$.	95
A.18	<i>Boxplots</i> dos IQs: DTLZ4 por seleção com $m = 15$.	96
A.19	<i>Boxplots</i> dos IQs: DTLZ4 por seleção com $m = 20$.	97
A.20	<i>Boxplots</i> dos IQs: DTLZ4 por seleção com $m = 30$.	98
A.21	<i>Boxplots</i> dos IQs: DTLZ5 por seleção com $m = 5$.	99
A.22	<i>Boxplots</i> dos IQs: DTLZ5 por seleção com $m = 10$.	100
A.23	<i>Boxplots</i> dos IQs: DTLZ5 por seleção com $m = 15$.	101
A.24	<i>Boxplots</i> dos IQs: DTLZ5 por seleção com $m = 20$.	102
A.25	<i>Boxplots</i> dos IQs: DTLZ5 por seleção com $m = 30$.	103
A.26	<i>Boxplots</i> dos IQs: DTLZ6 por seleção com $m = 5$.	104
A.27	<i>Boxplots</i> dos IQs: DTLZ6 por seleção com $m = 10$.	105
A.28	<i>Boxplots</i> dos IQs: DTLZ6 por seleção com $m = 15$.	106
A.29	<i>Boxplots</i> dos IQs: DTLZ6 por seleção com $m = 20$.	107
A.30	<i>Boxplots</i> dos IQs: DTLZ6 por seleção com $m = 30$.	108
A.31	<i>Boxplots</i> dos IQs: DTLZ7 por seleção com $m = 5$.	110
A.32	<i>Boxplots</i> dos IQs: DTLZ7 por seleção com $m = 10$.	111
A.33	<i>Boxplots</i> dos IQs: DTLZ7 por seleção com $m = 15$.	112

A.34	<i>Boxplots</i> dos IQs: DTLZ7 por seleção com $m = 20$.	113
A.35	<i>Boxplots</i> dos IQs: DTLZ7 por seleção com $m = 30$.	114
B.1	<i>Boxplots</i> dos IQs: DTLZ1 por extração com $m = 5$.	116
B.2	<i>Boxplots</i> dos IQs: DTLZ1 por extração com $m = 10$.	117
B.3	<i>Boxplots</i> dos IQs: DTLZ1 por extração com $m = 15$.	118
B.4	<i>Boxplots</i> dos IQs: DTLZ1 por extração com $m = 20$.	119
B.5	<i>Boxplots</i> dos IQs: DTLZ1 por extração com $m = 30$.	120
B.6	<i>Boxplots</i> dos IQs: DTLZ2 por extração com $m = 5$.	121
B.7	<i>Boxplots</i> dos IQs: DTLZ2 por extração com $m = 10$.	122
B.8	<i>Boxplots</i> dos IQs: DTLZ2 por extração com $m = 15$.	123
B.9	<i>Boxplots</i> dos IQs: DTLZ2 por extração com $m = 20$.	124
B.10	<i>Boxplots</i> dos IQs: DTLZ2 por extração com $m = 30$.	125
B.11	<i>Boxplots</i> dos IQs: DTLZ3 por extração com $m = 5$.	126
B.12	<i>Boxplots</i> dos IQs: DTLZ3 por extração com $m = 10$.	127
B.13	<i>Boxplots</i> dos IQs: DTLZ3 por extração com $m = 15$.	128
B.14	<i>Boxplots</i> dos IQs: DTLZ3 por extração com $m = 20$.	129
B.15	<i>Boxplots</i> dos IQs: DTLZ3 por extração com $m = 30$.	130
B.16	<i>Boxplots</i> dos IQs: DTLZ4 por extração com $m = 5$.	132
B.17	<i>Boxplots</i> dos IQs: DTLZ4 por extração com $m = 10$.	133
B.18	<i>Boxplots</i> dos IQs: DTLZ4 por extração com $m = 15$.	134
B.19	<i>Boxplots</i> dos IQs: DTLZ4 por extração com $m = 20$.	135
B.20	<i>Boxplots</i> dos IQs: DTLZ4 por extração com $m = 30$.	136
B.21	<i>Boxplots</i> dos IQs: DTLZ5 por extração com $m = 5$.	137
B.22	<i>Boxplots</i> dos IQs: DTLZ5 por extração com $m = 10$.	138
B.23	<i>Boxplots</i> dos IQs: DTLZ5 por extração com $m = 15$.	139
B.24	<i>Boxplots</i> dos IQs: DTLZ5 por extração com $m = 20$.	140
B.25	<i>Boxplots</i> dos IQs: DTLZ5 por extração com $m = 30$.	141
B.26	<i>Boxplots</i> dos IQs: DTLZ6 por extração com $m = 5$.	143
B.27	<i>Boxplots</i> dos IQs: DTLZ6 por extração com $m = 10$.	144
B.28	<i>Boxplots</i> dos IQs: DTLZ6 por extração com $m = 15$.	145

B.29 <i>Boxplots</i> dos IQs: DTLZ6 por extração com $m = 20$	146
B.30 <i>Boxplots</i> dos IQs: DTLZ6 por extração com $m = 30$	147
B.31 <i>Boxplots</i> dos IQs: DTLZ7 por extração com $m = 5$	148
B.32 <i>Boxplots</i> dos IQs: DTLZ7 por extração com $m = 10$	149
B.33 <i>Boxplots</i> dos IQs: DTLZ7 por extração com $m = 15$	150
B.34 <i>Boxplots</i> dos IQs: DTLZ7 por extração com $m = 20$	151
B.35 <i>Boxplots</i> dos IQs: DTLZ7 por extração com $m = 30$	152

Lista de Tabelas

2.1	$I_{\epsilon+}(A, B)$ para todas as combinações dos conjuntos A_1, A_2, A_3 e P	32
4.1	Resultados comparativos de LS e PCA - Parte 1 (DTLZ1 a DTLZ4)	62
4.2	Resultados comparativos de LS e PCA - Parte 2 (DTLZ5 a DTLZ7)	62
A.1	Escore: DTLZ1 com $m = 5$	78
A.2	Resultados da seleção de objetivos: DTLZ1 com $m = 5$	79
A.3	Escore: DTLZ1 com $m = 10$	79
A.4	Resultados da seleção de objetivos: DTLZ1 com $m = 10$	80
A.5	Escore: DTLZ1 com $m = 15$	80
A.6	Resultados da seleção de objetivos: DTLZ1 com $m = 15$	81
A.7	Escore: DTLZ1 com $m = 20$	81
A.8	Resultados da seleção de objetivos: DTLZ1 com $m = 20$	82
A.9	Escore: DTLZ1 com $m = 30$	82
A.10	Resultados da seleção de objetivos: DTLZ1 com $m = 30$	83
A.11	Escore: DTLZ2 com $m = 5$	84
A.12	Resultados da seleção de objetivos: DTLZ2 com $m = 5$	84
A.13	Escore: DTLZ2 com $m = 10$	84
A.14	Resultados da seleção de objetivos: DTLZ2 com $m = 10$	85
A.15	Escore: DTLZ2 com $m = 15$	85
A.16	Resultados da seleção de objetivos: DTLZ2 com $m = 15$	86
A.17	Escore: DTLZ2 com $m = 20$	86
A.18	Resultados da seleção de objetivos: DTLZ2 com $m = 20$	87

A.19	Escores: DTLZ2 com $m = 30$	87
A.20	Resultados da seleção de objetivos: DTLZ2 com $m = 30$	88
A.21	Escores: DTLZ3 com $m = 5$	89
A.22	Resultados da seleção de objetivos: DTLZ3 com $m = 5$	89
A.23	Escores: DTLZ3 com $m = 10$	89
A.24	Resultados da seleção de objetivos: DTLZ3 com $m = 10$	90
A.25	Escores: DTLZ3 com $m = 15$	90
A.26	Resultados da seleção de objetivos: DTLZ3 com $m = 15$	91
A.27	Escores: DTLZ3 com $m = 20$	91
A.28	Resultados da seleção de objetivos: DTLZ3 com $m = 20$	92
A.29	Escores: DTLZ3 com $m = 30$	92
A.30	Resultados da seleção de objetivos: DTLZ3 com $m = 30$	93
A.31	Escores: DTLZ4 com $m = 5$	94
A.32	Resultados da seleção de objetivos: DTLZ4 com $m = 5$	94
A.33	Escores: DTLZ4 com $m = 10$	94
A.34	Resultados da seleção de objetivos: DTLZ4 com $m = 10$	95
A.35	Escores: DTLZ4 com $m = 15$	95
A.36	Resultados da seleção de objetivos: DTLZ4 com $m = 15$	96
A.37	Escores: DTLZ4 com $m = 20$	96
A.38	Resultados da seleção de objetivos: DTLZ4 com $m = 20$	97
A.39	Escores: DTLZ4 com $m = 30$	97
A.40	Resultados da seleção de objetivos: DTLZ4 com $m = 30$	98
A.41	Escores: DTLZ5 com $m = 5$	99
A.42	Resultados da seleção de objetivos: DTLZ5 com $m = 5$	99
A.43	Escores: DTLZ5 com $m = 10$	100
A.44	Resultados da seleção de objetivos: DTLZ5 com $m = 10$	100
A.45	Escores: DTLZ5 com $m = 15$	101
A.46	Resultados da seleção de objetivos: DTLZ5 com $m = 15$	101
A.47	Escores: DTLZ5 com $m = 20$	102
A.48	Resultados da seleção de objetivos: DTLZ5 com $m = 20$	102

A.49	Escores: DTLZ5 com $m = 30$	103
A.50	Resultados da seleção de objetivos: DTLZ5 com $m = 30$	103
A.51	Escores: DTLZ6 com $m = 5$	104
A.52	Resultados da seleção de objetivos: DTLZ6 com $m = 5$	104
A.53	Escores: DTLZ6 com $m = 10$	105
A.54	Resultados da seleção de objetivos: DTLZ6 com $m = 10$	105
A.55	Escores: DTLZ6 com $m = 15$	106
A.56	Resultados da seleção de objetivos: DTLZ6 com $m = 15$	106
A.57	Escores: DTLZ6 com $m = 20$	107
A.58	Resultados da seleção de objetivos: DTLZ6 com $m = 20$	107
A.59	Escores: DTLZ6 com $m = 30$	108
A.60	Resultados da seleção de objetivos: DTLZ6 com $m = 30$	108
A.61	Escores: DTLZ7 com $m = 5$	109
A.62	Resultados da seleção de objetivos: DTLZ7 com $m = 5$	109
A.63	Escores: DTLZ7 com $m = 10$	110
A.64	Resultados da seleção de objetivos: DTLZ7 com $m = 10$	110
A.65	Escores: DTLZ7 com $m = 15$	111
A.66	Resultados da seleção de objetivos: DTLZ7 com $m = 15$	111
A.67	Escores: DTLZ7 com $m = 20$	112
A.68	Resultados da seleção de objetivos: DTLZ7 com $m = 20$	112
A.69	Escores: DTLZ7 com $m = 30$	113
A.70	Resultados da seleção de objetivos: DTLZ7 com $m = 30$	113
B.1	Porcentagens de variância total explicada: DTLZ1 com $m = 5$	115
B.2	Resultados da extração de objetivos: DTLZ1 com $m = 5$	116
B.3	Porcentagens de variância total explicada: DTLZ1 com $m = 10$	116
B.4	Resultados da extração de objetivos: DTLZ1 com $m = 10$	117
B.5	Porcentagens de variância total explicada: DTLZ1 com $m = 15$	117
B.6	Resultados da extração de objetivos: DTLZ1 com $m = 15$	118
B.7	Porcentagens de variância total explicada: DTLZ1 com $m = 20$	118
B.8	Resultados da extração de objetivos: DTLZ1 com $m = 20$	119

B.9	Porcentagens de variância total explicada: DTLZ1 com $m = 30$.	119
B.10	Resultados da extração de objetivos: DTLZ1 com $m = 30$.	120
B.11	Porcentagens de variância total explicada: DTLZ2 com $m = 5$.	121
B.12	Resultados da extração de objetivos: DTLZ2 com $m = 5$.	121
B.13	Porcentagens de variância total explicada: DTLZ2 com $m = 10$.	122
B.14	Resultados da extração de objetivos: DTLZ2 com $m = 10$.	122
B.15	Porcentagens de variância total explicada: DTLZ2 com $m = 15$.	123
B.16	Resultados da extração de objetivos: DTLZ2 com $m = 15$.	123
B.17	Porcentagens de variância total explicada: DTLZ2 com $m = 20$.	124
B.18	Resultados da extração de objetivos: DTLZ2 com $m = 20$.	124
B.19	Porcentagens de variância total explicada: DTLZ2 com $m = 30$.	125
B.20	Resultados da extração de objetivos: DTLZ2 com $m = 30$.	125
B.21	Porcentagens de variância total explicada: DTLZ3 com $m = 5$.	126
B.22	Resultados da extração de objetivos: DTLZ3 com $m = 5$.	126
B.23	Porcentagens de variância total explicada: DTLZ3 com $m = 10$.	127
B.24	Resultados da extração de objetivos: DTLZ3 com $m = 10$.	127
B.25	Porcentagens de variância total explicada: DTLZ3 com $m = 15$.	128
B.26	Resultados da extração de objetivos: DTLZ3 com $m = 15$.	128
B.27	Porcentagens de variância total explicada: DTLZ3 com $m = 20$.	129
B.28	Resultados da extração de objetivos: DTLZ3 com $m = 20$.	129
B.29	Porcentagens de variância total explicada: DTLZ3 com $m = 30$.	130
B.30	Resultados da extração de objetivos: DTLZ3 com $m = 30$.	130
B.31	Porcentagens de variância total explicada: DTLZ4 com $m = 5$.	131
B.32	Resultados da extração de objetivos: DTLZ4 com $m = 5$.	131
B.33	Porcentagens de variância total explicada: DTLZ4 com $m = 10$.	132
B.34	Resultados da extração de objetivos: DTLZ4 com $m = 10$.	132
B.35	Porcentagens de variância total explicada: DTLZ4 com $m = 15$.	133
B.36	Resultados da extração de objetivos: DTLZ4 com $m = 15$.	133
B.37	Porcentagens de variância total explicada: DTLZ4 com $m = 20$.	134
B.38	Resultados da extração de objetivos: DTLZ4 com $m = 20$.	134

B.39	Porcentagens de variância total explicada: DTLZ4 com $m = 30$.	135
B.40	Resultados da extração de objetivos: DTLZ4 com $m = 30$.	135
B.41	Porcentagens de variância total explicada: DTLZ5 com $m = 5$.	136
B.42	Resultados da extração de objetivos: DTLZ5 com $m = 5$.	137
B.43	Porcentagens de variância total explicada: DTLZ5 com $m = 10$.	137
B.44	Resultados da extração de objetivos: DTLZ5 com $m = 10$.	138
B.45	Porcentagens de variância total explicada: DTLZ5 com $m = 15$.	138
B.46	Resultados da extração de objetivos: DTLZ5 com $m = 15$.	139
B.47	Porcentagens de variância total explicada: DTLZ5 com $m = 20$.	140
B.48	Resultados da extração de objetivos: DTLZ5 com $m = 20$.	140
B.49	Porcentagens de variância total explicada: DTLZ5 com $m = 30$.	141
B.50	Resultados da extração de objetivos: DTLZ5 com $m = 30$.	141
B.51	Porcentagens de variância total explicada: DTLZ6 com $m = 5$.	142
B.52	Resultados da extração de objetivos: DTLZ6 com $m = 5$.	142
B.53	Porcentagens de variância total explicada: DTLZ6 com $m = 10$.	143
B.54	Resultados da extração de objetivos: DTLZ6 com $m = 10$.	143
B.55	Porcentagens de variância total explicada: DTLZ6 com $m = 15$.	144
B.56	Resultados da extração de objetivos: DTLZ6 com $m = 15$.	144
B.57	Porcentagens de variância total explicada: DTLZ6 com $m = 20$.	145
B.58	Resultados da extração de objetivos: DTLZ6 com $m = 20$.	145
B.59	Porcentagens de variância total explicada: DTLZ6 com $m = 30$.	146
B.60	Resultados da extração de objetivos: DTLZ6 com $m = 30$.	146
B.61	Porcentagens de variância total explicada: DTLZ7 com $m = 5$.	147
B.62	Resultados da extração de objetivos: DTLZ7 com $m = 5$.	148
B.63	Porcentagens de variância total explicada: DTLZ7 com $m = 10$.	148
B.64	Resultados da extração de objetivos: DTLZ7 com $m = 10$.	149
B.65	Porcentagens de variância total explicada: DTLZ7 com $m = 15$.	149
B.66	Resultados da extração de objetivos: DTLZ7 com $m = 15$.	150
B.67	Porcentagens de variância total explicada: DTLZ7 com $m = 20$.	150
B.68	Resultados da extração de objetivos: DTLZ7 com $m = 20$.	151

B.69 Porcentagens de variância total explicada: DTLZ7 com $m = 30$	152
B.70 Resultados da extração de objetivos: DTLZ7 com $m = 30$	152

Lista de Abreviaturas e Siglas

$I_{\epsilon+}$	ϵ -Indicador Aditivo Binário
AM	Aprendizado de Máquina
BI	Interseção de Limite
DC	Diagrama de Cordas
DM	Tomador de Decisão
DM1	Mecanismo de Gerenciamento de Diversidade 1
EA	Algoritmo Evolutivo
EP	População Externa
GD	Generational Distance
IBEA	<i>Indicator-Based Evolutionary Algorithm</i>
IGD	Inverted Generational Distance
IGD ⁺	Inverted Generational Distance Plus
IQ	Indicador de Qualidade
k-NN	<i>k-Nearest Neighbours Algorithm</i>
LS	<i>Laplacian Score</i>
MaOP	Problema de Otimização com Muitos Objetivos
MDMOEA	MOEA Dinâmico Baseado em L-dominância
MOEA	Algoritmo Evolutivo Multiobjetivo
MOEA/D	MOEA Baseado em Decomposição
MOP	Problema de Otimização Multiobjetivo
MOSS	Subconjunto Mínimo de Objetivos
MVU	<i>Maximum Variance Unfolding</i>
NSGA-II	<i>Non-Dominating Sorting Genetic Algorithm II</i>
NSGA-III	<i>Non-Dominating Sorting Genetic Algorithm III</i>
PCA	Análise de Componentes Principais
PCSEA	<i>Pareto Corner Search</i>
PCP	Gráfico de Coordenadas Paralelas
PBEA	<i>Preference-Based Evolutionary Algorithm</i>

PBI	Interseção de Limite Baseada em Penalidade
PF	Fronteira Pareto-ótima
RBF	Função de Base Radial
SPEA2	<i>Strength Pareto Evolutionary Algorithm 2</i>
SPEA2+SDE	<i>SPEA2 with Shift-Based Density Estimation</i>
VRPDRT	Problema de Roteamento de Veículos com Transporte Reativo à Demanda

Introdução

Problemas de otimização multiobjetivo (do inglês, *multiobjective optimization problems* – MOPs) buscam otimizar simultaneamente mais de uma função objetivo, geralmente conflitantes entre si. Isso significa que não se pode melhorar o valor de uma dessas funções objetivo sem piorar o valor de alguma outra. Resolver um MOP significa obter um conjunto de soluções ótimas – denominadas soluções Pareto-ótimas – que, quando avaliadas, produzem vetores cujas componentes representam os melhores acordos entre as funções objetivo (*trade-offs*). Ao final desta etapa, um tomador de decisão (do inglês, *decision maker* – DM) escolhe uma delas, de acordo com suas preferências, para ser a solução do problema em questão.

Devido às suas características, os MOPs são comumente resolvidos utilizando algoritmos evolutivos multiobjetivo (do inglês, *multiobjective evolutionary algorithms* – MOEAs), que se inspiram em processos biológicos da evolução. Os MOEAs são algoritmos baseados em população e não requerem conhecimento sobre as características das funções presentes no MOP, como diferenciabilidade, multimodalidade ou continuidade. Além disso, são capazes de fornecer uma boa aproximação das soluções Pareto-ótimas em uma única execução, ou seja, em um único processo de otimização, sem a necessidade de múltiplas execuções ou ajustes manuais para explorar o espaço de soluções (Deb, 2001).

A dimensionalidade de um MOP pode ser caracterizada pelo número de funções objetivo que o constitui (dimensionalidade do espaço de objetivos) ou pelo número de variáveis de decisão (dimensionalidade do espaço de decisão) que compõem o problema. MOPs com mais de três funções objetivo são denominados problemas de otimização com muitos objetivos (do inglês, *many-objective optimization problems* – MaOPs) (Farina e Amato, 2002). Os MaOPs representam uma classe muito importante desses problemas, pois à medida que aumenta-se a

dimensão do espaço de busca de um MOP, mais complexo ele fica, limitando a eficiência dos MOEAs. Ao trabalhar com MaOPs, visto que o número de objetivos é elevado, os MOEAs convencionais encontram sérias dificuldades (Li et al., 2015). A seguir, as três principais dificuldades são articuladas segundo os autores Coello Coello et al. (2020) e Nguyen et al. (2020).

A primeira é a baixa escalabilidade da maioria dos MOEAs baseados em dominância de Pareto, comprometendo a convergência para a fronteira Pareto-ótima (do inglês, *Pareto-optimal front* – PF). Quando há um grande número de objetivos, quase toda a população se torna não dominada entre si, dificultando o processo de seleção das gerações seguintes dos MOEAs. A segunda dificuldade é o alto custo computacional, uma vez que o tamanho da população necessária para aproximar aumenta exponencialmente com o número de objetivos e, conseqüentemente, mais operações a serem processadas. A terceira é a visualização das soluções em dimensões superiores, que pode dificultar o processo de escolha da solução final do MaOP pelo DM.

Estas dificuldades são frequentemente associadas à “maldição da dimensionalidade”, expressão cunhada por Bellman (1957) ao considerar problemas em programação dinâmica. Sob o contexto dos MaOPs, é importante diferenciar a maldição da dimensionalidade entre: a causada pelas dificuldades intrínsecas aos problemas, como, por exemplo, um conjunto Pareto-ótimo de alta dimensão, cujas soluções podem ter formas geométricas complexas e difíceis de caracterizar, com uma estrutura altamente não linear ou com regiões de soluções concentradas em áreas pequenas e de difícil análise; e a causada pela baixa escalabilidade da maioria dos MOEAs existentes, mesmo quando o conjunto Pareto-ótimo correspondente tem formas geométricas mais simples. Por exemplo, um MOP com menos de quatro funções objetivo pode representar um caso em que o problema em si não sofre da maldição da dimensionalidade (Saxena et al., 2012).

Ao analisar as publicações de Coello Coello et al. (2020) e de Nguyen et al. (2020), observa-se que a primeira apresenta três abordagens diferentes para lidar adequadamente com MaOPs, enquanto a segunda aborda duas delas, que também estão presentes no primeiro trabalho. Essa sobreposição indica uma área de convergência entre as publicações, onde as duas abordagens compartilhadas são baseadas em ordenação de preferências e em redução de dimensionalidade. Além dessas, o primeiro artigo também introduz uma terceira abordagem baseada em indicadores de qualidade (IQs), que são métricas utilizadas para avaliar a diversidade e a distribuição das soluções em um conjunto de soluções Pareto-ótimas aproximadas. Tais indicadores são frequentemente usados para comparar e selecionar

soluções em MOEAs, como descrito em Zitzler et al. (2003).

As abordagens de ordenação de preferências têm como suposição a ausência de objetivos redundantes no MaOP em questão. Elas assumem que nenhum objetivo no problema dado é redundante e visam contrariar a baixa pressão de seleção para convergência, induzindo uma ordem de preferência sobre as soluções não dominadas. Por outro lado, as abordagens de redução de dimensionalidade partem do pressuposto de que existem objetivos redundantes, buscando identificar um subconjunto menor de objetivos conflitantes que produza uma PF muito aproximada ou até igual à do MaOP original. Por fim, as abordagens baseadas em IQs incorporam métricas de desempenho no mecanismo de seleção de um MOEA, permitindo uma avaliação mais precisa da diversidade e da distribuição das soluções. No entanto, essas abordagens também possuem limitações, como a dependência de uma boa definição dos indicadores, a sensibilidade a parâmetros específicos e o custo computacional elevado para a avaliação de grandes populações, o que pode impactar a eficiência do algoritmo (Saxena et al., 2012).

As abordagens de redução de dimensionalidade oferecem alguns benefícios significativos. Por exemplo, se o número de objetivos não redundantes puder ser reduzido para menos de quatro, isso transformaria um MaOP – comumente difícil de se resolver – em um problema solucionável pelos MOEAs tradicionais. Mesmo quando o número de objetivos permanece em quatro ou mais, a redução de dimensionalidade melhora a escalabilidade do MOEA, resultando em menor custo computacional. No entanto, é importante notar que a redução de dimensionalidade pode levar a uma perda de informação, uma vez que, mesmo com uma correlação linear alta entre os objetivos (por exemplo, 0.9), ainda há um comprometimento na quantidade de informações representadas pelos objetivos restantes. Além disso, apesar dessa perda, há benefícios tanto na etapa de tomada de decisão quanto na visualização dos resultados. Esses ganhos adicionais reforçam a relevância das abordagens de redução de dimensionalidade na área de otimização multiobjetivo.

1.1 Descrição do Problema

Tendo em vista os desafios importantes encontrados na literatura envolvendo MOEAs (Brockhoff e Zitzler, 2006a,b; Li et al., 2015; Nguyen et al., 2020; Coello Coello et al., 2020), este trabalho propõe uma metodologia para redução do número de objetivos dos MaOPs, buscando minimizar a perda de informação, por meio de técnicas de aprendizado de máquina (AM, do inglês, *machine learning*) não supervisionado. Ou seja, a abordagem proposta utiliza

uma estrutura baseada em AM que opera nos vetores objetivos das soluções não dominadas obtidas de um MOEA para o modelo de MaOP fornecido, com o objetivo de aprender a estrutura de preferência das funções objetivo, procurando preservar a correlação entre as soluções na medida do possível.

De maneira geral, o AM pode ser definido como métodos computacionais usando a experiência para melhorar o desempenho ou fazer previsões precisas das tarefas. Neste caso, a experiência é estabelecida com as informações anteriores disponíveis para a máquina, normalmente em forma de dados coletados e disponibilizados para análise (Bishop e Nasrabadi, 2006). Uma vez que o AM utiliza técnicas de otimização para ajustar os parâmetros dos modelos e alcançar um melhor desempenho, o inverso também pode ser atribuído. Sendo assim, já que o AM e a otimização são intrínsecos, é possível adotar seus métodos para desenvolver a linha de pesquisa de MaOPs (Song et al., 2019).

Especificamente, uma abordagem de seleção e uma de extração de atributos são escolhidas nesse trabalho: *Laplacian Score* (LS) (He et al., 2005) e Análise de Componentes Principais (do inglês, *Principal Component Analysis* – PCA) (Jolliffe, 2002), respectivamente. Elas são utilizadas para reduzir o número de objetivos de problemas de teste *benchmark* da família DTLZ (Deb et al., 2002b) e depois otimizados pelo MOEA estado da arte NSGA-III (Deb e Jain, 2013; Jain e Deb, 2013). Feito isso, as novas funções são expandidas para o número de objetivos original, a fim de comparação com as soluções atingidas pelos mesmos MOEAs para os casos primitivos, usando os IQs: IGD^+ (Ishibuchi et al., 2015) e I_{ϵ^+} (Zitzler et al., 2003) com seus respectivos *boxplots* (em português, diagramas de caixa). Por fim, é feita uma análise estatística utilizando o Teste de Postos Sinalizados de Wilcoxon.

1.2 Motivação

Os MaOPs aparecem amplamente em aplicações do mundo real, por exemplo: controle de tráfego aéreo (Herrero et al., 2008), calibração de motores automotivos (Fleming et al., 2005), controlador de carro híbrido (Narukawa e Rodemann, 2012), planejamento do portfólio de abastecimento de água (Kasprzyk et al., 2012), alinhamento de sequência de proteínas (DeRonne e Karypis, 2013), otimização estrutural (Carvalho et al., 2021), além do Problema de Roteamento de Veículos com Transporte Reativo a Demanda (Mendes et al., 2017), e muitos outros modelos. Conforme mencionado por Nguyen et al. (2020), a modelagem e a solução dos MaOPs são cada vez mais reconhecidas como um tópico de pesquisa chave na comunidade de otimização multiobjetivo.

Além disso, na publicação Saxena et al. (2012), os autores destacam que a produção de bens e serviços enfrenta diversas demandas conflitantes e é crucial estabelecer projetos que atendam simultaneamente a uma ampla gama de objetivos de desempenho. Embora essa necessidade esteja cada vez mais evidente, é preocupante constatar que a capacidade de busca de alguns dos MOEAs mais reconhecidos se deteriora significativamente quando mais de três objetivos estão envolvidos.

Segundo evidenciado no artigo de revisão dos autores Coello Coello et al. (2020), apesar dos MOEAs terem alcançado um alto nível de maturidade como área de pesquisa, ainda há uma série de desafios importantes a serem enfrentados. Entre eles, destaca-se o desenvolvimento de MOEAs capazes de lidar com MaOPs que sejam mais consistentes com problemas do mundo real. Nesse contexto, a abordagem de redução de dimensionalidade surge como uma alternativa para enfrentar esses desafios. Essa abordagem pode ser aplicada de diversas maneiras, inclusive explorando técnicas provenientes de outras áreas da computação, a fim de fornecer soluções alternativas para a resolução desses desafios complexos.

1.3 Objetivos

O propósito geral desta pesquisa é integrar um algoritmo de redução de dimensionalidade à um MOEA, que reduza para dois ou três o número de objetivos de um problema de otimização *benchmark* com muitos objetivos, obtendo resultados semelhantes ou melhores quando comparados aos MOEAs estado da arte aplicados no problema original. Especificamente, o trabalho se propõe a:

- Investigar os principais trabalhos relacionados que empregam a redução da dimensionalidade em MaOPs;
- Produzir conjuntos de dados dos problemas de teste *benchmark* contínuos da família DTLZ, que são amplamente utilizados na literatura por sua capacidade de representar problemas com diferentes números de objetivos e estruturas de interdependência entre as funções;
- Aplicar nestes conjuntos de dados duas técnicas de AM não supervisionado para redução de dimensionalidade: o LS, com seleção dos objetivos, e a PCA, com extração dos objetivos;
- Resolver estes MaOPs reduzidos utilizando o NSGA-III e retomar o espaço de objetivos

dos resultados à dimensão original para comparação com os resultados dos problemas primários;

- Efetuar uma comparação qualitativa entre as soluções obtidas e as soluções dos problemas originais, utilizando o IGD^+ e o I_{ϵ^+} , assim como a ferramenta de visualização *boxplot*;
- Realizar uma análise estatística por meio do Teste de Postos Sinalizados de Wilcoxon para discutir os resultados, visando identificar e mensurar os ganhos de desempenho.

1.4 Organização do Documento

Este trabalho está organizado em cinco capítulos. O primeiro capítulo apresenta a introdução, a definição do problema, a motivação da pesquisa e os objetivos a serem alcançados. O Capítulo 2 trata da fundamentação teórica, abordando conceitos de otimização multiobjetivo, problemas de teste *benchmark* e as técnicas de aprendizado de máquina não supervisionado empregadas. O Capítulo 3 apresenta uma revisão bibliográfica sobre o estado da arte na área de otimização multiobjetivo e redução de dimensionalidade. No Capítulo 4, são discutidos os resultados dos experimentos computacionais realizados para a redução do número de objetivos. Por fim, o Capítulo 5 apresenta as conclusões do trabalho, as propostas de continuidade da pesquisa, e os apêndices, que contêm a descrição dos experimentos, organizados por técnicas de redução de dimensionalidade e problemas DTLZ.

Fundamentação Teórica

Neste capítulo, são revisados os conceitos fundamentais relacionados aos MaOPs e seus desafios, aos MOEAs e IQs que são abordados neste trabalho e à redução de dimensionalidade por seleção e extração de atributos.

2.1 Problema de Otimização Multiobjetivo

Otimização é o processo de encontrar e comparar soluções candidatas viáveis até que nenhuma solução candidata melhor possa ser encontrada. Elas são caracterizadas como boas ou ruins em termos dos objetivos, que podem ser, por exemplo, as dimensões de uma peça, o custo de fabricação da mesma e a sua resistência mecânica. No campo da otimização, o modelo para descrever um problema como este pode considerar apenas um objetivo – tal como o custo de fabricação – e as outras metas como restrições de projeto que não devem ser violadas. Desta maneira, existe apenas uma única solução. Alternativamente, este problema pode ser modelado contendo vários objetivos que devem ser otimizados ao mesmo tempo e, se for o caso, ainda possuir restrições. Assim, a solução para o problema busca gerar um conjunto de soluções que carregam uma série de *trade-offs* entre os objetivos, onde um não pode ser melhorado sem piorar o(s) outro(s).

Primeiramente, informa-se que a notação matemática em **negrito** representa vetores. Dito isso, o primeiro modelo é denominado problema de otimização mono-objetivo e é formulado da seguinte maneira:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} f(\mathbf{x}), \\ \text{sujeito a: } &\begin{cases} g_j(\mathbf{x}) \leq 0, & j = 1, 2, \dots, r, \\ h_k(\mathbf{x}) = 0, & k = 1, 2, \dots, p, \end{cases} \end{aligned} \quad (2.1)$$

em que $\mathbf{x} \in \mathbb{R}^n$, $f(\cdot) \in \mathbb{R}^n$, $\mathbf{g}(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^r$ e $\mathbf{h}(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^p$. O vetor n -dimensional $\mathbf{x}$ é chamado vetor de variáveis de decisão, $f(\mathbf{x})$ é denominado função objetivo e $g_j(\mathbf{x})$ e $h_k(\mathbf{x})$ são conhecidas como restrições de desigualdade e igualdade, respectivamente. O número de variáveis n e o número de restrições r e/ou p não precisam estar relacionadas de alguma forma.

O segundo modelo é definido como MOP e se trata de uma ampliação na dimensão da Equação (2.1). Este problema é formulado a seguir:

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} \mathbf{f}(\mathbf{x}), \\ \text{sujeito a: } &\begin{cases} g_j(\mathbf{x}) \leq 0, & j = 1, 2, \dots, r, \\ h_k(\mathbf{x}) = 0, & k = 1, 2, \dots, p, \end{cases} \end{aligned} \quad (2.2)$$

em que $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{f}(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \geq 2$, $\mathbf{g}(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^r$ e $\mathbf{h}(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^p$. Neste caso, temos mais de um vetor de variáveis de decisão $\mathbf{x} \in \mathbb{R}^n$ e encontram-se em um espaço vetorial chamado espaço de decisão. Os vetores de funções objetivo $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$ situam-se em um espaço vetorial denominado espaço de objetivos (Deb, 2001). A Figura 2.1 apresenta um exemplo de mapeamento entre $\mathbf{x}$ e sua imagem $\mathbf{f}(\mathbf{x})$.

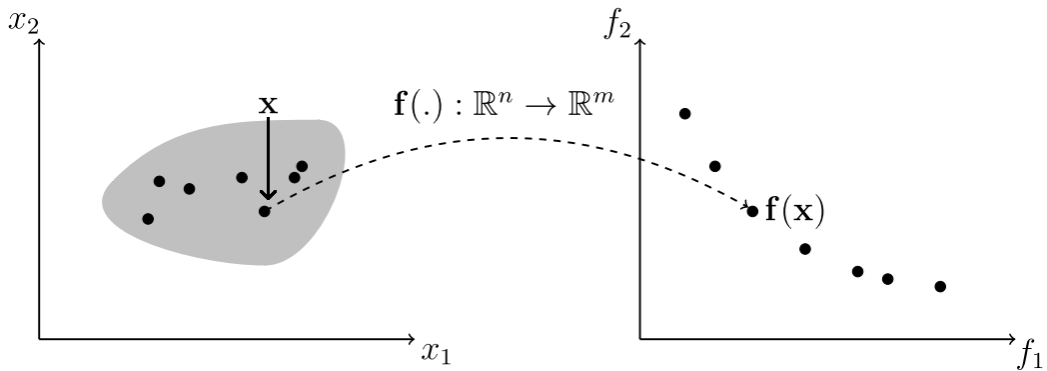


Figura 2.1: Mapeamento do espaço de decisão no espaço de objetivos. Adaptado de: Mendes (2022).

Algumas definições importantes do campo da otimização multiobjetivo são apresentadas a seguir. Maiores detalhes podem ser obtidos em referências clássicas da literatura tais como

(Deb, 2001) e (Coello, 2007).

Definição 2.1.1 (Solução Factível). Uma solução $\mathbf{x}$ que não satisfaça todas as restrições $(r + p)$ é chamada de solução infactível. Por outro lado, se qualquer solução $\mathbf{x}$ satisfizer todas as restrições, ela é conhecida como uma solução factível.

Definição 2.1.2 (Região Factível). Uma vez que o MOP apresente restrições, todo o espaço de decisão não é necessariamente factível. Então, o conjunto de soluções factíveis é chamado de região factível, ou simplesmente espaço de busca Ω .

Definição 2.1.3 (Dominância). Uma solução $\mathbf{x}_1$ domina outra solução $\mathbf{x}_2$ se $\mathbf{f}(\mathbf{x}_1) \leq \mathbf{f}(\mathbf{x}_2)$ e $\mathbf{f}(\mathbf{x}_1) \neq \mathbf{f}(\mathbf{x}_2)$. Nestas mesmas condições, diz-se que $\mathbf{f}(\mathbf{x}_1)$ domina $\mathbf{f}(\mathbf{x}_2)$ e esta relação de dominância é denotada por $\mathbf{f}(\mathbf{x}_1) \preceq \mathbf{f}(\mathbf{x}_2)$.

Definição 2.1.4 (Dominância Estrita). Se uma solução $\mathbf{x}_1$ é estritamente melhor que a solução $\mathbf{x}_2$ para todos os objetivos, isto é, $\mathbf{f}(\mathbf{x}_1) < \mathbf{f}(\mathbf{x}_2)$, diz-se que $\mathbf{x}_1$ domina estritamente $\mathbf{x}_2$, ou, equivalentemente, $\mathbf{f}(\mathbf{x}_1)$ domina estritamente $\mathbf{f}(\mathbf{x}_2)$. Esta relação pode ser denotada por $\mathbf{f}(\mathbf{x}_1) \prec \mathbf{f}(\mathbf{x}_2)$.

Para exemplificar as relações de dominância nas Definições 2.1.3 e 2.1.4, consideraremos a Figura 2.2 que apresenta cinco soluções diferentes no espaço de objetivos. Também consideraremos que ambas as funções f_1 e f_2 devem ser minimizadas.

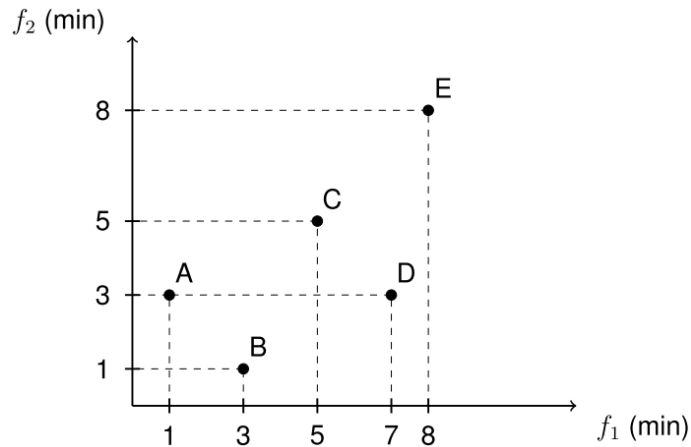


Figura 2.2: Abstração da ideia de dominância. Fonte: Mendes (2022).

Comparando-se a solução A com a solução D, nota-se que para o valor de f_1 a solução A é menor que D e, para f_2 , as duas soluções possuem o mesmo valor, ou seja, a solução A domina a solução D ($A \preceq D$). Agora, comparando-se as soluções A e C, percebe-se que A possui menor valor de f_1 e f_2 que C, logo, A domina estritamente C ($A \prec C$).

Outro cenário acontece comparando-se as soluções A e B. Nota-se que a solução A possui menor valor de f_1 em relação a B, porém, possui maior valor de f_2 . Neste caso, como as relações de dominância $A \prec B$ e $B \prec A$ ou $A \preceq B$ e $B \preceq A$ não puderam ser verificadas, as soluções são denominadas incomparáveis ou não dominadas entre si.

Definição 2.1.5 (Solução Pareto-Ótima). $\mathbf{x}^*$ é uma solução Pareto-ótima se não existe qualquer outra solução $\mathbf{x}$ tal que $\mathbf{f}(\mathbf{x}) \preceq \mathbf{f}(\mathbf{x}^*)$, ou seja, se $\mathbf{x}^*$ não é dominada por nenhum outro ponto factível.

Existe uma outra nomenclatura para solução Pareto-ótima, sendo referida como solução não dominada ou solução eficiente. Mais uma vez analisando a Figura 2.2, podemos dizer que as soluções A e B são soluções não dominadas.

Definição 2.1.6 (Conjunto Pareto-Ótimo). O conjunto de soluções não dominadas de todo o espaço de busca factível é o conjunto Pareto-ótimo. O conjunto-imagem associado ao conjunto Pareto-ótimo é chamado de PF.

Por último, um MOP com mais de três objetivos pode ser denominado como um MaOP, cuja principal motivação é destacar os desafios impostos por um grande número de objetivos aos MOEAs existentes. Sua formulação é dada pela Equação (2.2) com $m \geq 4$. Conforme escrito por (Li et al., 2015, pág. 4):

“[...] Embora atualmente utiliza-se quatro objetivos como critério divisor entre MOPs e MaOPs, esse número tende a aumentar à medida que MOEAs mais avançados forem desenvolvidos no futuro. Portanto, a definição apresentada aqui representa uma evolução que serve a um propósito mais prático do que teórico.”

2.1.1 Problemas de Teste *Benchmark* da Família DTLZ

No artigo de Deb et al. (2002b), os autores introduzem os problemas de teste DTLZ, concebidos com o objetivo de serem úteis em diversas atividades de pesquisa em MOEAs. Esses conjuntos de funções foram especialmente elaborados para testar o desempenho de novos algoritmos, possibilitar a comparação entre diferentes abordagens e aprofundar a compreensão dos princípios subjacentes aos MOEAs. Além disso, a capacidade de expansão do número de objetivos desses problemas permite que suas formulações possam ser adaptadas para os MaOPs.

Essa escolha se fundamenta na revisão bibliográfica realizada no Capítulo 3, que revelou o uso recorrente destes problemas em estudos como Saxena et al. (2012); Brockhoff e Zitzler

(2006a,b); López Jaimes et al. (2008). A meta é empregar as mesmas técnicas de LS e PCA para estes MaOPs. A seguir, são brevemente apresentados os sete problemas de teste desta família, em ordem crescente.

O DTLZ1 se apresenta como única dificuldade a convergência para o hiperplano Pareto-ótimo. O espaço de busca contém $(11^k - 1)$ PFs locais, para as quais um MOEA pode ser atraído antes de atingir a PF global. Sua formulação é apresentada a seguir:

$$\arg \min_{\mathbf{x}} \begin{cases} f_1(\mathbf{x}) = \frac{1}{2}x_1x_2 \cdots x_{m-1} (1 + g(\mathbf{x}_m)), \\ f_2(\mathbf{x}) = \frac{1}{2}x_1x_2 \cdots (1 - x_{m-1}) (1 + g(\mathbf{x}_m)), \\ \vdots \\ f_{m-1}(\mathbf{x}) = \frac{1}{2}x_1 (1 - x_2) (1 + g(\mathbf{x}_m)), \\ f_m(\mathbf{x}) = \frac{1}{2} (1 - x_1) (1 + g(\mathbf{x}_m)), \end{cases} \quad \text{sujeito a: } \begin{cases} 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \quad (2.3)$$

o funcional $g(\mathbf{x}_m)$ requer variáveis $|\mathbf{x}_m| = k$ e deve aceitar qualquer função com $g \geq 0$. Os autores sugerem igual à função Rastrigin (Mühlenbein et al., 1991):

$$g(\mathbf{x}_m) = 100 \left[|\mathbf{x}_m| + \sum_{x_i \in \mathbf{x}_m} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right], \quad (2.4)$$

sendo $n = m + k - 1$ o número total de variáveis e $k = 5$, conforme aconselhado pelos autores. O conjunto Pareto-ótimo corresponde a $x_i^* = 0.5 \forall x_i^* \in \mathbf{x}_m$ e a PF está no hiperplano linear: $\sum_{j=1}^m f_j^* = 0.5$. Para três funções objetivo, a PF forma um triângulo no espaço tridimensional, conforme a Figura 2.3.

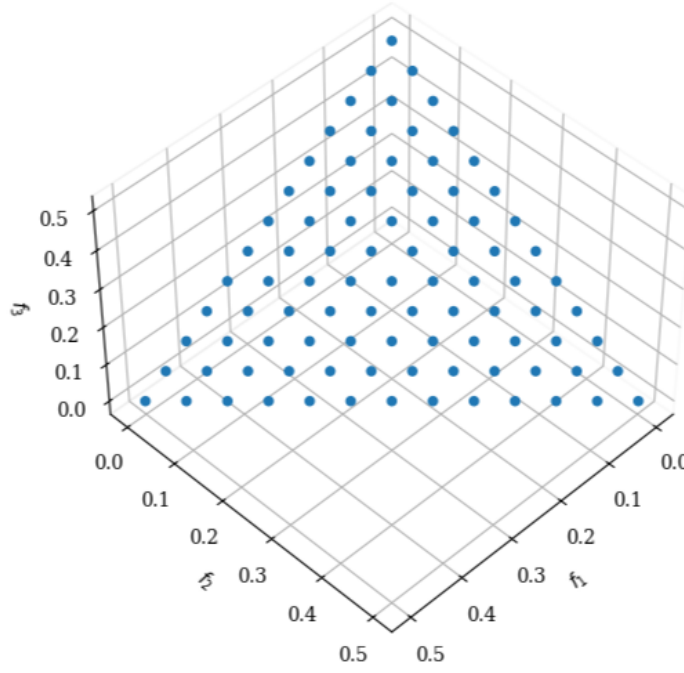


Figura 2.3: Visualização da PF gerada para o DTLZ1 com 3 funções objetivo.

Os pontos nesta fronteira satisfazem a condição $\sum_{j=1}^m f_j^* = 0.5$, refletindo um *trade-off* equilibrado entre os objetivos. Este formato é ideal para estudar problemas com soma constante e restrições lineares.

Seguindo para o DTLZ2, ele pode ser usado para avaliar a capacidade de um MOEA em melhorar seu desempenho quando lida com um grande número de objetivos. Este problema é formulado da seguinte maneira:

$$\arg \min_{\mathbf{x}} \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \cos(x_{m-2} \pi/2) \cos(x_{m-1} \pi/2), \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \cos(x_{m-2} \pi/2) \sin(x_{m-1} \pi/2), \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \sin(x_{m-2} \pi/2), \\ \vdots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \sin(x_1 \pi/2), \end{cases}$$

$$\text{sujeito a: } \begin{cases} g(\mathbf{x}_m) = \sum_{x_i \in \mathbf{x}_m} (x_i - 0.5)^2, \\ 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \quad (2.5)$$

e mais uma vez, o vetor $\mathbf{x}_m$ é construído sendo $n = m + k - 1$ e $k = 5$, assim como o conjunto Pareto-ótimo, que corresponde a $x_i^* = 0.5 \forall x_i^* \in \mathbf{x}_m$. De maneira similar ao DTLZ1 com $m = 3$, no DTLZ2 a PF está localizada no primeiro octante da hipersfera unitária em um gráfico tridimensional representado na Figura 2.4, onde f_m é um dos eixos.

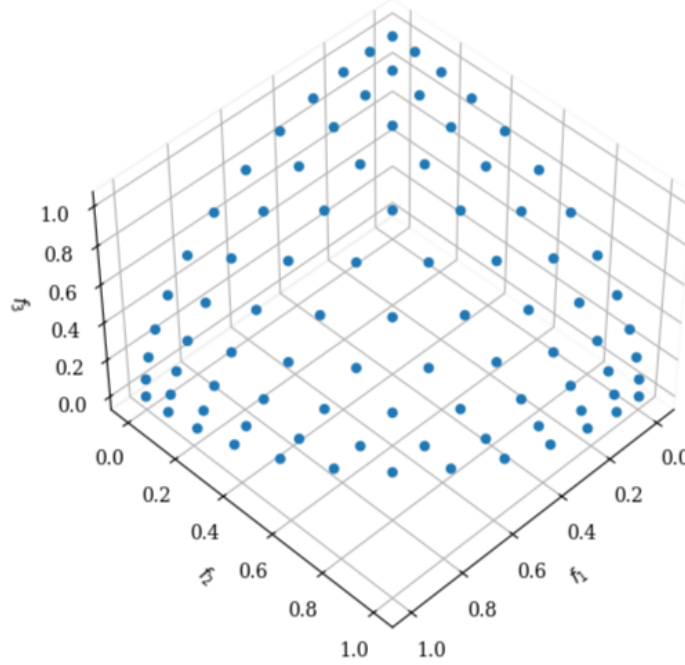


Figura 2.4: Visualização da PF gerada para o DTLZ2 com 3 funções objetivo.

Portanto, todas as soluções Pareto-ótimas devem satisfazer $\sum_{j=1}^m f_j^* = 1$, onde os objetivos competem de maneira simétrica. Além disso, a diferença entre o valor do termo esquerdo nas soluções obtidas e 1 também pode ser usada como uma métrica de convergência.

Já o DTLZ3 trata a potencialidade de um MOEA de convergir para a PF global, ele possui funções objetivo idênticas às do DTLZ2, mas com o funcional $g(\mathbf{x}_m)$ igual à função Rastrigin em 2.4. Dessa forma, $g(\mathbf{x}_m)$ introduz $(3^k - 1)$ PFs locais, além de uma PF global. Todas as PFs locais são paralelas à PF global, e um MOEA pode ficar preso em qualquer uma dessas PFs locais antes de convergir para a PF global (em $g^* = 0$). Formulado a seguir:

$$\arg \min_{\mathbf{x}} \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \cos(x_{m-2} \pi/2) \cos(x_{m-1} \pi/2), \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \cos(x_{m-2} \pi/2) \sin(x_{m-1} \pi/2), \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1 \pi/2) \cdots \sin(x_{m-2} \pi/2), \\ \vdots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \sin(x_1 \pi/2), \end{cases}$$

$$\text{sujeito a: } \begin{cases} g(\mathbf{x}_m) = 100 \left[|\mathbf{x}_m| + \sum_{x_i \in \mathbf{x}_m} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right], \\ 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \quad (2.6)$$

outra vez, a PF global corresponde ao conjunto Pareto-ótimo $x_i^* = 0.5 \forall x_i^* \in \mathbf{x}_m$ e a sua forma também é uma hipersfera, como pode ser vista na Figura 2.5.

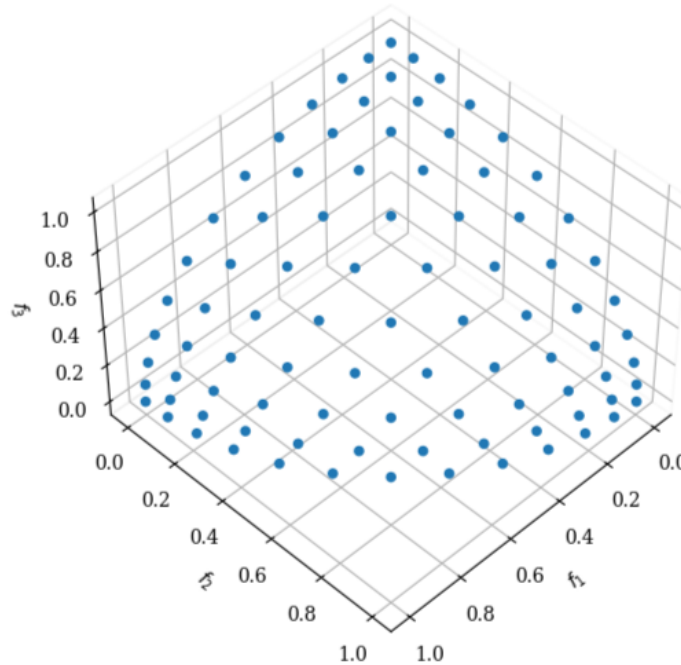


Figura 2.5: Visualização da PF gerada para o DTLZ3 com 3 funções objetivo.

O problema de teste DTLZ4 foi proposto a fim de investigar a capacidade de um MOEA de manter uma boa distribuição de soluções, para isso os autores modificaram o problema

DTLZ2 com um mapeamento de variável paramétrica diferente, assim segue:

$$\arg \min_{\mathbf{x}} \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1^\alpha \pi/2) \cdots \cos(x_{m-2}^\alpha \pi/2) \cos(x_{m-1}^\alpha \pi/2), \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1^\alpha \pi/2) \cdots \cos(x_{m-2}^\alpha \pi/2) \sin(x_{m-1}^\alpha \pi/2), \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(x_1^\alpha \pi/2) \cdots \sin(x_{m-2}^\alpha \pi/2), \\ \vdots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \sin(x_1^\alpha \pi/2), \end{cases}$$

$$\text{sujeito a: } \begin{cases} g(\mathbf{x}_m) = \sum_{x_i \in \mathbf{x}_m} (x_i - 0.5)^2, \\ 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \quad (2.7)$$

o parâmetro $\alpha = 100$ é sugerido pelos autores e adotado neste trabalho. Esta modificação permite que exista um conjunto denso de soluções próximo ao plano $f_m - f_1$. Apesar desta diferença fundamental, a forma da PF de DTLZ4 também é uma hiperesfera como em DTLZ2 e DTLZ3, apresentada na Figura 2.6.

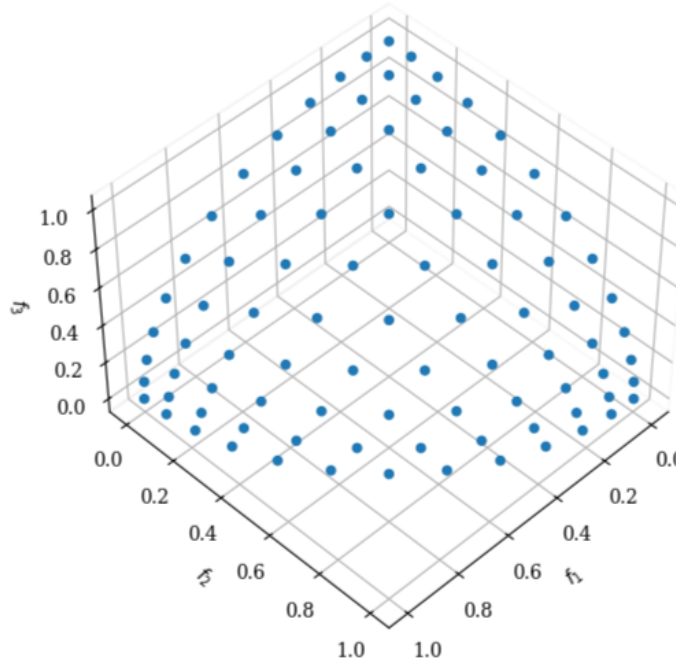


Figura 2.6: Visualização da PF gerada para o DTLZ4 com 3 funções objetivo.

Este viés exige que o algoritmo de otimização explore regiões de menor densidade de

soluções, desafiando sua capacidade de encontrar uma cobertura uniforme.

Seguindo para o DTLZ5, ele foi desenvolvido com o objetivo de testar o potencial de convergência de um MOEA para uma determinada curva, ou seja, possui uma PF que gradualmente reduz a dimensionalidade efetiva. Além disso, ele facilita a visualização do desempenho do MOEA, permitindo que se plote f_m em relação a qualquer outra função objetivo. Formulado da seguinte maneira:

$$\begin{aligned} \arg \min_{\mathbf{x}} \quad & \begin{cases} f_1(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(\theta_1 \pi/2) \cdots \cos(\theta_{m-2} \pi/2) \cos(\theta_{m-1} \pi/2), \\ f_2(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(\theta_1 \pi/2) \cdots \cos(\theta_{m-2} \pi/2) \sin(\theta_{m-1} \pi/2), \\ f_3(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \cos(\theta_1 \pi/2) \cdots \sin(\theta_{m-2} \pi/2), \\ \vdots \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_m)) \sin(\theta_1 \pi/2), \end{cases} \\ \text{sujeito a:} \quad & \begin{cases} \theta_i = \frac{1}{2(1+g(\mathbf{x}_m))} (1 + 2g(\mathbf{x}_m)x_i), \quad \text{para } i = 2, 3, \dots, (m-1), \\ g(\mathbf{x}_m) = \sum_{x_i \in \mathbf{x}_m} (x_i - 0.5)^2, \\ 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \end{aligned} \quad (2.8)$$

o conjunto de soluções Pareto-ótimo corresponde a $x_i^* = 0.5 \forall x_i^* \in \mathbf{x}_m$ e os valores da PF satisfazem $\sum_{j=1}^m f_j^* = 1$. Para três objetivos, a fronteira forma uma curva hiperbólica em um espaço bidimensional, conforme representado na Figura 2.7.

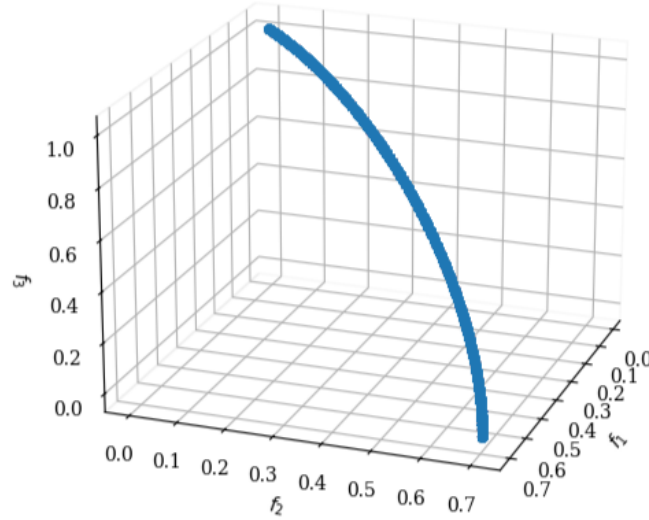


Figura 2.7: Visualização da PF gerada para o DTLZ5 com 3 funções objetivo.

O DTLZ5 é particularmente interessante para avaliar algoritmos em contextos de redução da dimensionalidade do espaço de soluções. Além disso, devido ao viés natural em direção a soluções próximas à PF, esse problema pode ser mais facilmente resolvido por um algoritmo.

O problema de teste DTLZ5 pode ser dificultado fazendo uma modificação semelhante à função $g(\mathbf{x}_m)$, como feito no DTLZ3 ao utilizar a função Rastrigin. Sendo assim, para a formulação do DTLZ6 os autores optaram por um funcional $g(\mathbf{x}_m)$ diferente no DTLZ5, sendo:

$$g(\mathbf{x}_m) = \sum_{x_i \in \mathbf{x}_m} x_i^{0.1}, \quad (2.9)$$

a dimensão do vetor $\mathbf{x}_m$ é escolhida como 10 e o número total de variáveis é idêntico ao do DTLZ5. De acordo com Blank e Deb (2020), esta mudança no problema dificulta a convergência do NSGA-II e do SPEA2 para a verdadeira PF, diferente de como ocorre no DTLZ5. Conforme a Figura 2.8, a PF de DTLZ6 é semelhante à de DTLZ5, mas com maior complexidade devido à introdução de efeitos de escalabilidade diferenciada nos objetivos.

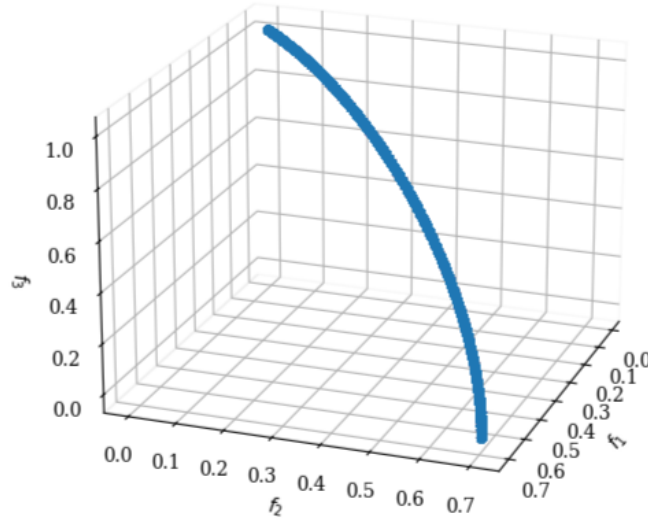


Figura 2.8: Visualização da PF gerada para o DTLZ6 com 3 funções objetivo.

Aqui também o conjunto de soluções Pareto-ótimo corresponde a $x_i^* = 0.5 \forall x_i^* \in \mathbf{x}_m$. Os objetivos têm pesos não uniformes, o que resulta em regiões da fronteira com densidades variáveis de soluções, destacando a necessidade de equilíbrio entre exploração e convergência.

Por último, o problema de teste DTLZ7 possui regiões Pareto-ótimas desconectadas em 2^{m-1} no espaço de busca. O funcional $g(\mathbf{x}_m)$ requer $k = |\mathbf{x}_m| = n - m + 1$ variáveis de decisão. Este problema testa a capacidade de um MOEA de manter subpopulações em diferentes regiões Pareto-ótimas, formulado a seguir:

$$\arg \min_{\mathbf{x}} \begin{cases} f_1(\mathbf{x}_1) = x_1, \\ f_2(\mathbf{x}_2) = x_2, \\ \vdots \\ f_{m-1}(\mathbf{x}_{m-1}) = x_{m-1}, \\ f_m(\mathbf{x}) = (1 + g(\mathbf{x}_m))h(f_1, f_2, \dots, f_{m-1}, g), \end{cases}$$

$$\text{sujeito a: } \begin{cases} g(\mathbf{x}_m) = 1 + \frac{9}{|\mathbf{x}_m|} \sum_{x_i \in \mathbf{x}_m} x_i, \\ h(f_1, f_2, \dots, f_{m-1}, g) = m - \sum_{i=1}^{m-1} \left[\frac{f_i}{1+g} (1 + \sin(3\pi f_i)) \right], \\ 0 \leq x_i \leq 1, \\ i = 1, 2, \dots, n, \end{cases} \quad (2.10)$$

o conjunto de soluções Pareto-ótimo corresponde a $x_i^* = 0 \forall x_i^* \in \mathbf{x}_m$. A PF de DTLZ7 é disforme e composta por uma região mista com características descontínuas. Para três objetivos, os dois primeiros formam uma região bidimensional, enquanto o terceiro objetivo é determinado por uma equação condicional que resulta em uma forma de “corte”, como pode ser visto na Figura 2.9.

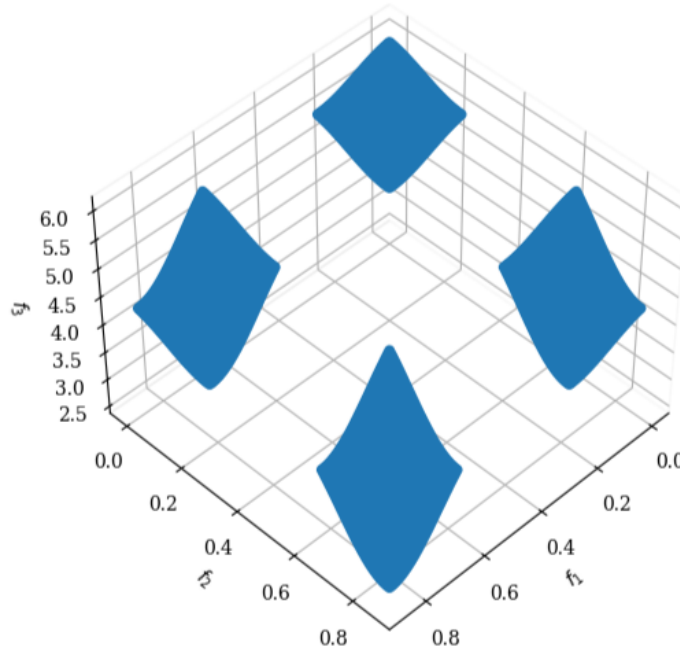


Figura 2.9: Visualização da PF gerada para o DTLZ7 com 3 funções objetivo.

Em suma, os sete problemas de teste DTLZ são fundamentais para avaliar MOEAs, oferecendo funções diversificadas e escaláveis que testam o desempenho e comparam diferentes abordagens. Eles complementam a formulação do MaOP discutida nesta seção, fornecendo uma base prática para a subsequente análise dos MOEAs. Assim, esses problemas de teste ligam a teoria dos MaOPs à prática dos MOEAs, facilitando uma avaliação robusta dos algoritmos apresentados na sequência.

2.2 Algoritmos Evolutivos Multiobjetivo

O texto desta seção foi fundamentado no livro de Deb (2001). Buscando encontrar soluções para os MOPs, primeiramente os pesquisadores adotaram duas metodologias bem fundamentadas na literatura: os métodos clássicos de busca e otimização e os Algoritmos Evolutivos (do inglês, Evolutionary Algorithms – EAs), afinal, ambos resolviam a maioria dos problemas de otimização mono-objetivo.

A maior parte dos algoritmos clássicos de busca e otimização usa um procedimento determinístico para se aproximar da solução ótima. Esses algoritmos partem de uma solução de estimativa aleatória. A partir daí, com base em uma regra de transição pré-especificada, os algoritmos sugerem uma direção de busca, que muitas vezes é determinada considerando informações locais. Uma busca unidirecional é então realizada ao longo da direção de busca para encontrar a melhor solução. Esta melhor solução torna-se a nova solução, e o procedimento acima é continuado várias vezes. Os algoritmos variam principalmente na forma como as direções de busca são definidas em cada solução intermediária.

Já os EAs são uma família de algoritmos de otimização inspirados no processo de evolução biológica, imitando os princípios da seleção natural, reprodução e mutação. Os EAs melhoram iterativamente uma população de soluções candidatas ao longo de várias gerações para procurar a melhor solução possível. Diferentemente dos métodos clássicos, não precisam de suposições particulares como continuidade ou diferenciabilidade das funções.

Outra diferença importante entre um método clássico de busca e otimização e um EA é que, neste último, uma população de soluções é processada em cada iteração ou geração. Esse atributo por si só confere ao EA uma enorme vantagem para uso na resolução de MOPs.

Por meio de modificações no EA básico, é possível capturar uma população de soluções Pareto-ótimas em uma única simulação do EA. Com isso, elimina-se a necessidade de recorrer repetidamente a um método de otimização mono-objetivo para encontrar uma solução Pareto-ótima diferente em cada execução a fim de formar a PF, como é feito pelos métodos clássicos.

Assim, a abordagem populacional dos EAs pode ser explorada para enfatizar todas as soluções não dominadas em uma população de forma igualitária e, simultaneamente, preservar um conjunto diversificado de múltiplas soluções não dominadas por meio do uso de um operador de preservação de nicho. Dessa forma, é possível encontrar e manter várias boas soluções em uma população. Ao longo de algumas gerações, esse processo pode levar a população a convergir em direção à PF, apresentando um bom afastamento entre as soluções

encontradas. Por esses motivos, os EAs são altamente adequados para lidar com MOPs e, como resultado, receberam a denominação de MOEAs. A seguir é apresentado brevemente o passo a passo simplificado de como eles funcionam:

1. **Inicialização:** os MOEAs começam criando uma população inicial de soluções candidatas. Cada solução é representada como um conjunto de variáveis de decisão.
2. **Avaliação:** A adequação de cada solução candidata é avaliada com base em múltiplas funções objetivo. Essas funções objetivo quantificam diferentes aspectos do problema que está sendo resolvido e podem ser conflitantes.
3. **Seleção:** MOEAs empregam mecanismos de seleção para determinar quais soluções devem ser permitidas para reproduzir e criar descendentes. Vários métodos de seleção, como seleção de torneio, seleção de roleta ou seleção baseada em dominância de Pareto, podem ser usados. O objetivo é favorecer soluções que sejam melhores em pelo menos um objetivo sem desconsiderar completamente os demais.
4. **Reprodução:** As soluções selecionadas são usadas para criar descendentes por meio de operadores como cruzamento e mutação. O cruzamento combina partes de duas ou mais soluções pais para criar uma nova prole, enquanto a mutação introduz mudanças aleatórias na prole. A fim de introduzir diversidade e explorar o espaço de busca de forma eficaz.
5. **Avaliação e Seleção:** A aptidão das soluções descendentes é avaliada usando as funções objetivo. A população descendente é combinada com a população parental e um mecanismo de seleção é aplicado para determinar as soluções que avançarão para a próxima geração. O processo de seleção geralmente favorece soluções que não são apenas competitivas em termos de objetivos individuais, mas também diversificadas e representativas de diferentes *trade-offs*.
6. **Término:** O algoritmo continua evoluindo a população através de gerações sucessivas até que uma condição de término seja atendida. As condições comuns de terminação incluem atingir um número máximo de gerações, convergência da população para um certo nível ou diversidade suficiente na PF.
7. **Análise de resultados:** Uma vez que o algoritmo termina, a população final de soluções representa uma aproximação da PF. Essas soluções oferecem diferentes com-

pensações entre objetivos conflitantes, permitindo que os DMs explorem o espaço da solução e tomem decisões informadas com base em suas preferências.

Segundo Li et al. (2015), os primeiros MOEAs foram classificados em quatro categorias distintas, baseadas em diferentes abordagens: dominância de Pareto, agregação, IQs e preferências.

Entre os MOEAs tradicionais baseados em dominância de Pareto, como o NSGA-II (Deb et al., 2002a) e o SPEA2 (Zitzler et al., 2001), há uma tendência de apresentarem desempenho significativamente inferior em MaOPs. Isso ocorre, em grande parte, devido à perda de pressão seletiva, um fenômeno em que a capacidade de diferenciar soluções com base na dominância é reduzida em problemas de alta dimensionalidade, dificultando a convergência para a PF. A fim de contornar essa limitação, surgiram duas abordagens principais para melhorar a escalabilidade dos algoritmos desta categoria: a baseada em dominância relaxada, que reduz os critérios de dominância para ampliar a pressão seletiva, e a baseada em diversidade, que promove a manutenção de soluções espalhadas pelo espaço de objetivos.

Além dessas cinco abordagens, outras duas também foram desenvolvidas: a baseada em conjunto de referência e, de maneira diferente das outras, a baseada em redução de dimensionalidade.

Em resumo, um total de sete categorias foram adotadas pela literatura. A seguir, são apresentadas breves definições de cada uma dessas classes. Para o leitor interessado, recomenda-se consultar Li et al. (2015) que apresenta uma exploração mais aprofundada sobre o assunto e outros exemplos de MOEAs em cada categoria:

1. **Baseados em dominância relaxada:** A dominância relaxada é uma abordagem que permite diferenciar soluções não dominadas, aumentando assim a pressão seletiva. Essa técnica visa contornar o problema do aumento das regiões não dominadas, este desafio é detalhado na Seção 2.4. Existem duas formas principais de realizar a relaxação da dominância:
 - **Relaxação baseada no valor das funções objetivo:** Essa abordagem modifica a dominância de Pareto, alterando os valores de cada objetivo. O objetivo é ampliar a região de soluções dominadas pelas soluções não dominadas. Um exemplo é o algoritmo ϵ -NSGA-II (Kollat e Reed, 2006) que aplica no NSGA-II o conceito de ϵ -dominância de Laumanns et al. (2002);
 - **Relaxação baseada no número de funções objetivo dominantes:** Nesse método, duas soluções são comparadas com base na contagem de objetivos em

que uma solução é maior, igual ou menor que a outra. Por exemplo, o MOEA dinâmico baseado em L-dominância, denominado MDMOEA (Zou et al., 2008).

2. **Baseados em diversidade:** Esses algoritmos empregam mecanismos para controlar a diversidade da população. Na literatura, muitos estudos buscam aumentar a pressão seletiva dos algoritmos em direção à PF. No entanto, essa abordagem pode levar a uma pressão de seleção excessivamente alta, resultando na concentração da população em uma região específica do espaço de busca. Um exemplo é o NSGA-II/DM1 (Adra e Fleming, 2010) que introduz no NSGA-II um mecanismo de gerenciamento de diversidade chamado DM1.
3. **Baseados em indicadores de qualidade:** Os algoritmos avaliam conjuntos de soluções obtidos usando IQs que guiam o processo de busca. Por exemplo, o *Indicator-Based Evolutionary Algorithm* (IBEA) (Zitzler e Künzli, 2004), que usa um ϵ -Indicador Aditivo Binário ($I_{\epsilon+}$).
4. **Baseados em conjunto de referência:** Os algoritmos baseados em conjunto de referência utilizam um conjunto de soluções como referência para avaliar a qualidade de outras soluções. No entanto, esses métodos enfrentam algumas dificuldades, como a definição do conjunto de soluções de referência e a forma de medir a qualidade das soluções em relação a esse conjunto. Por exemplo, o *Non-dominated Sorting Genetic Algorithm III* (NSGA-III) (Deb e Jain, 2013; Jain e Deb, 2013), que é detalhado na Subseção 2.2.1.
5. **Baseados em preferências:** Esses algoritmos buscam reduzir o número de soluções da PF levando em consideração as preferências dos usuários. Um dos critérios importantes a serem definidos nesse tipo de algoritmo é quando inserir a preferência. Esse processo pode ocorrer antes, durante ou após o processo de otimização. Por exemplo, o *Preference-Based Evolutionary Algorithm* (PBEA) (Thiele et al., 2009), que leva em consideração as informações de preferência integrando uma função de realização que depende do ponto de referência.
6. **Baseados em agregação:** Esses algoritmos utilizam funções de agregação que podem ser baseadas informações individuais dos objetivos, como a soma ponderada, ou em comparações em pares, como valores máximos e mínimos. Por exemplo, o *Multiobjective Evolutionary Algorithm Based on Decomposition* (MOEA/D) (Zhang e Li,

2007), que é baseado em decomposição, isto é, fundamentado nos métodos que comumente usam as informações individuais agregadas para comparar as soluções, neste caso, agregação de valores dos objetivos (Li et al., 2015).

7. **Baseados em redução de dimensionalidade:** Esses algoritmos visam resolver os MaOPs reduzindo o número de funções objetivo do problema. Isso pode ser feito selecionando um conjunto representativo de funções ou por meio da combinação de objetivos redundantes. Exemplos foram discutidos no Capítulo 3.

Apresentadas as sete categorias dos MOEAs, segue-se para a definição do algoritmo utilizado neste trabalho: o NSGA-III. A escolha desse algoritmo foi motivada pelo seu amplo uso como estado da arte e, apesar das dificuldades inerentes, ele tem demonstrado resultados satisfatórios na resolução de MaOPs (Coello Coello et al., 2020).

2.2.1 NSGA-III

O algoritmo NSGA-III, proposto por Deb e Jain (2013); Jain e Deb (2013), compartilha uma estrutura básica semelhante ao seu antecessor, o NSGA-II (Deb et al., 2002a), mas incorpora mudanças significativas em seu mecanismo de seleção. Isto é, a manutenção da diversidade entre os membros da população no NSGA-III é auxiliada pelo fornecimento e atualização adaptativa de vários pontos de referência bem espalhados.

A principal ideia por trás do NSGA-III é a utilização de um processo de classificação em duas etapas. Na primeira etapa, todas as soluções da população são classificadas com base em sua dominância, criando assim diferentes PFs. As soluções não dominadas compõem a primeira PF, as soluções dominadas apenas por essas soluções formam a segunda PF, e assim por diante. Na segunda etapa, o NSGA-III atribui a cada solução um nível de não dominância (ou aptidão); PFs de menor nível têm prioridade sobre PFs de níveis mais altos no processo de seleção.

Matematicamente, considere a t -ésima geração do algoritmo NSGA-III. Suponha que a população pai nesta geração seja P_t e seu tamanho seja N , enquanto a população descendente, criada a partir de P_t , é Q_t com mesmo tamanho. O primeiro passo é escolher os melhores N membros da população combinada de pais e filhos $R_t = P_t \cup Q_t$, de tamanho $2N$, permitindo assim preservar os membros de elite da população-mãe. Para conseguir isso, primeiro, a população combinada R_t é classificada de acordo com diferentes níveis de não dominância (F_1, F_2 e assim por diante). Em seguida, cada nível é selecionado um de cada vez para construir uma nova população S_t , começando pela primeira PF F_1 , até que o tamanho

de S_t seja igual a N ou pela primeira vez ultrapasse N . Sendo o l o último nível incluído, todas as soluções do nível $(l + 1)$ em diante são rejeitadas da população combinada R_t . Na maioria das situações, o último nível aceito (l -ésimo nível) é aceito apenas parcialmente.

Além disso, apenas as soluções que maximizarão a diversidade da l -ésima PF são escolhidas. No NSGA-II, isso é obtido por meio de um operador de preservação de nicho computacionalmente eficiente, porém aproximado, que calcula a distância de aglomeração para cada membro do último nível como a soma da distância objetivamente normalizada entre duas soluções vizinhas. Com isso, são escolhidas as soluções com maiores valores de distância de aglomeração. Para o NSGA-III, substitui-se o operador de distância de aglomeração por outras abordagens no processo de seleção.

Assim como o NSGA-II, a classificação da população em níveis não dominados do NSGA-III também usa o mesmo procedimento de identificar PFs não dominadas usando o princípio de dominação usual. Além disso, conforme mencionado anteriormente, o NSGA-III usa um conjunto predefinido de pontos de referência para garantir a diversidade nas soluções obtidas. Os pontos de referência escolhidos podem ser predefinidos de forma estruturada ou fornecidos preferencialmente pelo usuário, neste trabalho utilizou-se o Riesz s-Energy (Blank et al., 2020), uma abordagem sistemática que coloca pontos em um hiperplano normalizado – um *unit simplex* $(m - 1)$ dimensional – que é igualmente inclinado a todos os eixos objetivos e tem uma interceptação de um em cada eixo.

No NSGA-III, além de enfatizar soluções não dominadas, também enfatiza-se os membros da população que estão de alguma forma associados a cada um desses pontos de referência. Uma vez que os pontos de referência criados acima são amplamente distribuídos em todo o hiperplano normalizado, é provável que as soluções obtidas também sejam amplamente distribuídas na PF ou próximo a ela. O pseudocódigo (Algoritmo 1) deste MOEA é apresentado a seguir:

Algoritmo 1: Geração t do NSGA-III

Entrada: H : pontos de referência estruturados, Z^s ou pontos de aspiração
fornecidos Z^a , P_t : população parental

Saída: P_{t+1}

1 **início**

2 $S_t = \emptyset, i = 1$

3 $Q_t = \text{Recombinação} + \text{Muta\c{c}\~ao}(P_t)$

4 $R_t = P_t \cup Q_t$

5 $F_1, F_2, \dots = \text{Ordena\c{c}\~ao n\~ao dominada}(R_t)$

6 **repita**

7 $S_t = S_t \cup F_i$

8 $i = i + 1$

9 **at\~e** $|S_t| \geq N$

10 \c{U}ltima frente a ser inclu\~ida: $F_l = F_t$

11 **se** $|S_t| = N$ **ent\~ao**

12 $P_{t+1} = S_t$, pare

13 **sen\~ao**

14 $P_{t+1} = \cup_{j=1}^{l-1} F_j$

15 Pontos a serem escolhidos F_l : $K = N - |P_{t+1}|$

16 Normalize os objetivos e crie um conjunto de refer\~encia Z^r : **Normalizar**
 $(\mathbf{f}^n, S_t, Z^r, Z^s, Z^a)$

17 Associe cada membro s de S_t com um ponto de refer\~encia: $[\pi(s), d(s)] =$
Associar (S_t, Z^r) , sendo $\pi(s)$: ponto de refer\~encia mais pr\~oximo e d :
dist\~ancia entre s e $\pi(s)$

18 Calcule a contagem de nicho do ponto de refer\~encia $j \in Z^r$:

$$\rho_j = \sum_{\mathbf{s} \in S_t / F_t} ((\pi(s) = j) ? 1 : 0)$$

19 Escolha K membros um de cada vez de F_l para construir P_{t+1} : **Nicho**

$$(K, \rho_j, \pi, d, Z^r, F_l, P_{t+1})$$

Vale ressaltar que cada uma das fun\c{c}\~oes **Normalizar**, **Associar** e **Nicho** do processo de

seleção no Algoritmo 1, trata-se de um procedimento específico apresentado na publicação do NSGA-III. Para o leitor interessado em maiores detalhes de cada uma, recomenda-se consultar Deb e Jain (2013); Jain e Deb (2013).

2.3 Indicadores de Qualidade

Sabendo que existe uma vasta variedade de MOEAs na literatura, tornam-se necessários métodos de comparação para medir a qualidade dos conjuntos de soluções obtidos por diferentes algoritmos. Em geral, o desempenho de um MOEA é avaliado por meio de testes experimentais e, por consequência, diversos IQs foram definidos para esse fim, bem como para comparar diferentes abordagens para o mesmo problema (Riquelme et al., 2015).

Neste trabalho, adotaram-se dois IQs conhecidos como *Inverted Generational Distance Plus* (IGD⁺) (Ishibuchi et al., 2015) e o ϵ -Indicador Aditivo Binário (I _{ϵ +}) (Zitzler et al., 2003). Optou-se pelo primeiro devido à sua capacidade de manter um desempenho robusto mesmo diante do aumento do número de objetivos, uma vez que se baseia exclusivamente na distância entre os pontos no espaço de objetivos (Nguyen et al., 2020). O segundo, de forma similar, preserva sua eficácia mesmo diante da dimensão do espaço de objetivos, utilizando apenas a relação de ϵ -dominância. Ambos são preferidos também devido ao seu baixo custo computacional, tornando-os eficientes para análises em MaOPs.

2.3.1 *Inverted Generational Distance Plus*

Para formular o IGD⁺ de maneira clara, é importante descrever os dois IQs dos quais ele se baseia. O primeiro é o GD, proposto por Van Veldhuizen (1999), que mede a distância entre a PF retornada pelo MOEA e a PF de referência definida pelas funções que compõem o MOP. Sendo assim, o GD é um valor que reflete o quão distante a PF encontrada está em relação à PF de referência.

Assumindo que os pontos encontrados pelo algoritmo são o conjunto de vetores objetivo $A = \{a_1, a_2, \dots, a_{|A|}\}$ e o conjunto de pontos de referência é $Z = \{z_1, z_2, \dots, z_{|Z|}\}$, define-se:

$$\text{GD}(A) = \frac{1}{|A|} \left(\sum_{i=1}^{|A|} d_i^p \right)^{1/p} \quad (2.11)$$

dado que d_i representa a distância euclidiana ($p = 2$) entre a_i e o seu ponto de referência mais próximo em Z . Pode-se afirmar que o resultado desta equação representa a média das

distâncias entre qualquer ponto A e o ponto mais próximo na PF. Um resultado igual a zero indica que a solução é idêntica à referência, enquanto qualquer outro valor indica que as soluções são distintas entre si.

A partir do GD, foi proposto o IGD, conforme descrito por Coello Coello e Reyes Sierra (2004). Ele se baseia no conceito do GD, porém com uma abordagem invertida, neste caso utiliza a PF de referência como base e compara cada elemento dela com a PF produzida pelo MOEA. Essa abordagem visa mitigar alguns dos problemas encontrados com o uso do GD quando um MOEA gera apenas um pequeno número de soluções não dominadas. Baseando-se na Equação (2.11), define-se:

$$\text{IGD}(A) = \frac{1}{|Z|} \left(\sum_{i=1}^{|Z|} \hat{d}_i^p \right)^{1/p} \quad (2.12)$$

em que $\hat{d}_i$ representa a distância euclidiana ($p = 2$) entre z_i e o seu ponto de referência mais próximo em A . Visando uma melhora nesta abordagem, o IGD^+ leva em consideração a relação de dominância entre uma solução e um ponto de referência quando calculamos sua distância.

Isto é, caso uma solução seja dominada por um ponto de referência, a distância euclidiana é utilizada para calcular a distância sem alterações. No entanto, quando as soluções não são dominadas entre si, a distância mínima é calculada do ponto de referência até a região dominada pela solução. Essa distância pode ser interpretada como um indicador da inferioridade da solução, ou seja, a insuficiência de seus valores objetivos em relação ao ponto de referência. Partindo da Equação (2.12), obtém-se:

$$\text{IGD}^+(A) = \frac{1}{|Z|} \left(\sum_{i=1}^{|Z|} d_i^{+2} \right)^{1/2} \quad (2.13)$$

dado que $d_i^+ = \max \{a_i - z_i, 0\}$ representa a distância modificada de z_i para a solução mais próxima em A com o valor correspondente a_i .

Em conclusão, os três IQs apresentados são focados na medição da convergência para a PF. No entanto, outras propriedades importantes também devem ser consideradas na avaliação de MOEAs, como a uniformidade, que reflete a distribuição equilibrada das soluções no conjunto; a abrangência (*spread*), que mede a extensão do espaço objetivo coberta pelas soluções; e a cardinalidade, que indica a quantidade de soluções presentes no conjunto. Embora o GD, o IGD e o IGD^+ sejam úteis para avaliar a convergência, eles não capturam diretamente essas propriedades, o que pode limitar sua aplicabilidade em análises mais com-

pletas. Além disso, é necessário ter o conjunto Pareto de referência do MOP para calcular esses IQs. Assim, para problemas em que não se tem acesso ao conjunto Pareto de referência, o uso destes indicadores não é viável.

Outra propriedade crítica dessas métricas é a compatibilidade com Pareto. Segundo descrito por Falcón-Cardona et al. (2021), esta propriedade permite que um IQ não contradiga a ordem imposta pela relação de dominância de Pareto. Portanto, IQs compatíveis com Pareto são notáveis ao comparar conjuntos de aproximação de MOEAs, uma vez que não produzem conclusões enganosas.

Conforme validado por Ishibuchi et al. (2015), o IGD^+ é fracamente compatível com Pareto, enquanto o GD e o IGD original não são compatíveis com Pareto. A Figura 2.10 ilustra um exemplo de um par de conjuntos de soluções A e B para um problema de minimização de dois objetivos.

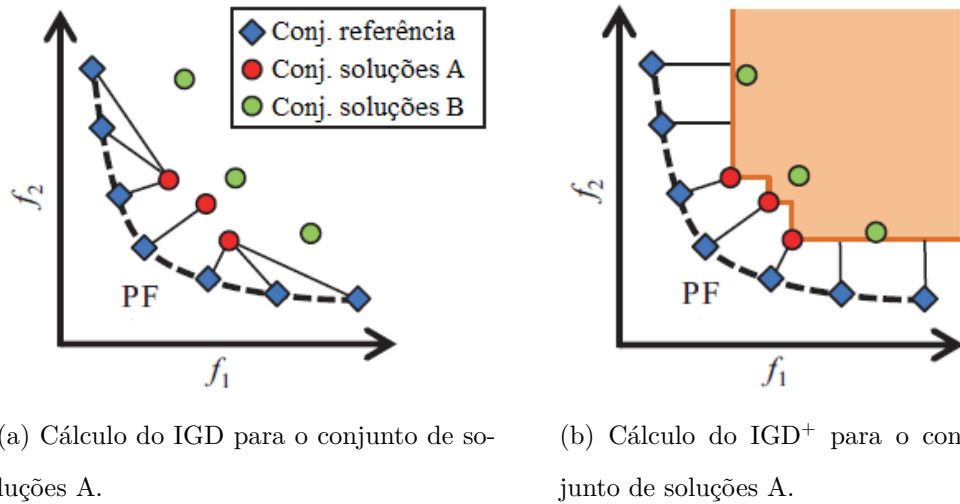


Figura 2.10: Exemplo de um par de conjuntos de solução A e B onde A é melhor que B em termos de dominância de Pareto. Adaptado de: Ishibuchi et al. (2016).

Na Figura 2.10, podemos observar que o conjunto de soluções A domina o conjunto B em termos de dominância de Pareto. Isso significa que, para cada solução em B , existe pelo menos uma solução em A que é igual ou melhor em todos os objetivos e estritamente melhor em pelo menos um. Essa relação é evidenciada pela posição relativa dos pontos na figura, em que os pontos de A estão mais próximos da fronteira de Pareto em comparação aos pontos de B .

No entanto, como mostrado na Figura 2.10a, o IGD atribui um valor pior (maior) ao conjunto A em relação ao conjunto B . Isso ocorre porque o cálculo do IGD considera a distância média dos pontos da PF de referência aos pontos mais próximos na PF aproximada, sem levar em conta a dominância de Pareto. Assim, mesmo que A seja superior a B segundo a

relação de dominância, o IGD pode resultar em uma avaliação contrária devido à distribuição espacial dos pontos.

Na Figura 2.10b, vemos como o IGD^+ corrige essa limitação ao incorporar a dominância de Pareto no cálculo de distância. Com isso, o conjunto A apresenta um valor de IGD^+ melhor (menor) do que B , refletindo sua superioridade em relação à dominância de Pareto. Essa diferença é essencial para comparações mais confiáveis entre conjuntos de soluções.

2.3.2 Épsilon-Indicador Aditivo Binário

O indicador Epsilon Zitzler et al. (2003) é amplamente utilizado para comparar o desempenho de soluções em MOPs. Sua principal característica é ser um indicador binário, ou seja, ele sempre requer a comparação entre dois conjuntos de soluções. Esse requisito adiciona um parâmetro extra ao processo de avaliação, uma vez que é necessário especificar os dois conjuntos que serão analisados. Esse aspecto merece ser ressaltado, pois pode influenciar tanto a interpretação quanto a aplicação do indicador em diferentes contextos.

Definição 2.3.1 (ϵ -Dominância). Para um problema de minimização com m objetivos positivos, isto é, $\mathbf{f}(\mathbf{x}) \subseteq \mathbb{R}^m$, considera-se dois vetores objetivos $\mathbf{f}^1(\mathbf{x}) = (f_1^1, f_2^1, \dots, f_m^1) \in \mathbf{f}(\mathbf{x})$ e $\mathbf{f}^2(\mathbf{x}) = (f_1^2, f_2^2, \dots, f_m^2) \in \mathbf{f}(\mathbf{x})$. Dizemos que $\mathbf{f}^1(\mathbf{x})$ ϵ -domina $\mathbf{f}^2(\mathbf{x})$, escrito como $\mathbf{f}^1(\mathbf{x}) \succeq_\epsilon \mathbf{f}^2(\mathbf{x})$, se e somente se

$$\forall 1 \leq i \leq m : f_i^1 \leq \epsilon + f_i^2 \quad (2.14)$$

para um determinado $\epsilon > 0$.

Definição 2.3.2 (ϵ -Indicador Aditivo Binário). $I_{\epsilon+}(A, B)$ é igual ao menor valor de ϵ tal que qualquer vetor objetivo em B seja ϵ -dominado por pelo menos um vetor objetivo em A . No caso mono-objetivo, $I_{\epsilon+}(A, B)$ é simplesmente a razão entre os dois valores objetivos representados por A e B .

$$I_{\epsilon+}(A, B) = \inf_{\epsilon \in \mathbb{R}} \{ \forall \mathbf{f}^2(\mathbf{x}) \in B \exists \mathbf{f}^1(\mathbf{x}) \in A : \mathbf{f}^1(\mathbf{x}) \succeq_{\epsilon+} \mathbf{f}^2(\mathbf{x}) \} \quad (2.15)$$

para quaisquer dois conjuntos de aproximação $A, B \in \Omega$.

Em outras palavras, diz-se que um vetor $\mathbf{f}^1(\mathbf{x})$ ϵ -domina outro vetor $\mathbf{f}^2(\mathbf{x})$ se, ao somar um fator ϵ a cada valor objetivo em $\mathbf{f}^2(\mathbf{x})$, o vetor resultante ainda for fracamente dominado por $\mathbf{f}^1(\mathbf{x})$. Portanto, $\mathbf{f}^1(\mathbf{x}) \succ \mathbf{f}^2(\mathbf{x})$ implica que existe um $\epsilon < 1$ tal que $\mathbf{f}^1(\mathbf{x})$ ϵ -domina $\mathbf{f}^2(\mathbf{x})$. Assim, o ϵ -indicador fornece o fator pelo qual um conjunto de soluções de aproximação é inferior a outro em relação a todos os objetivos.

Em experimentos computacionais, $I_{\epsilon+}(A, B)$ pode ser calculado no tempo $\mathcal{O}(m \cdot |A| \cdot |B|)$ da seguinte maneira:

$$\begin{aligned}\epsilon_{\mathbf{f}^1(\mathbf{x}), \mathbf{f}^2(\mathbf{x})} &= \max_{1 \leq i \leq m} (f_i^1 - f_i^2) \quad \forall \mathbf{f}^1(\mathbf{x}) \in A, \mathbf{f}^2(\mathbf{x}) \in B \\ \epsilon_{\mathbf{f}^2(\mathbf{x})} &= \min_{\mathbf{f}^1(\mathbf{x}) \in A} \epsilon_{\mathbf{f}^1(\mathbf{x}), \mathbf{f}^2(\mathbf{x})} \quad \forall \mathbf{f}^2(\mathbf{x}) \in B \\ I_{\epsilon+}(A, B) &= \max_{\mathbf{f}^2(\mathbf{x}) \in B} \epsilon_{\mathbf{f}^2(\mathbf{x})}\end{aligned}$$

ou equivalente

$$I_{\epsilon+}(A, B) = \max_{\mathbf{f}^2(\mathbf{x}) \in B} \min_{\mathbf{f}^1(\mathbf{x}) \in A} \max_{1 \leq i \leq m} (f_i^1 - f_i^2) \quad (2.16)$$

A Figura 2.11 é um exemplo retirado da bibliografia, onde $I_{\epsilon+}(A_1, A_2) = 1$, $I_{\epsilon+}(A_1, A_3) = \frac{9}{10}$ e $I_{\epsilon+}(A_1, P) = 4$. A área sombreada representa o subespaço que é ϵ -dominado pelas soluções em A_1 para $\epsilon = \frac{9}{10}$; a área sombreada média representa o subespaço fracamente dominado por A_1 (equivalente a $\epsilon = 1$). A área sombreada clara refere-se ao subespaço não ϵ -dominado pelas soluções em A_1 para $\epsilon = 4$. Observa-se que as áreas estão sobrepostas, ou seja, a área sombreada média inclui a área sombreada escura, e a área sombreada clara inclui ambas as outras áreas.

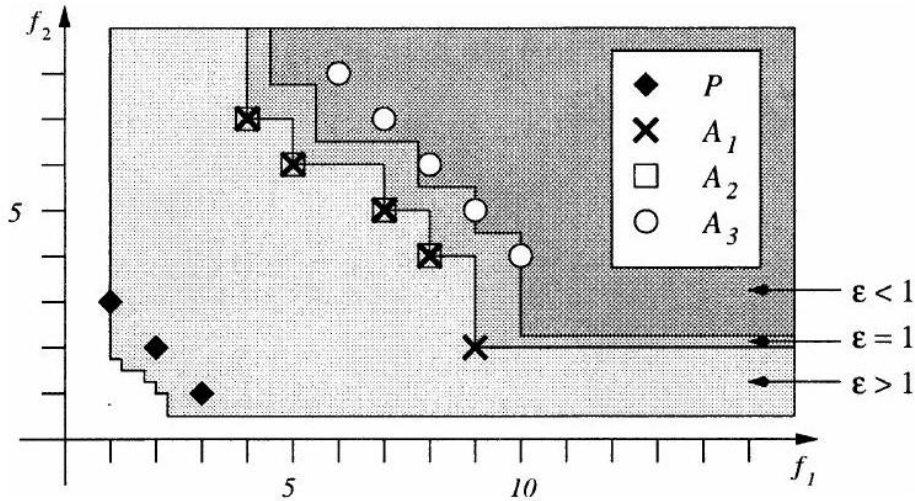


Figura 2.11: ϵ -Dominância. Fonte: Zitzler et al. (2003).

Analisando a figura, percebe-se que as soluções A_1 e A_2 possuem os mesmos valores, com a diferença de que A_1 possui uma solução a mais, o que torna maior seu subespaço de área sombreada média. A Tabela 2.1 apresenta todos os valores calculados do IQ para este exemplo.

Tabela 2.1: $I_{\epsilon+}(A, B)$ para todas as combinações dos conjuntos A_1, A_2, A_3 e P .

B	A			
	A_1	A_2	A_3	P
A_1	1	2	2	1/2
A_2	1	1	3/2	3/7
A_3	9/10	1	1	1/3
P	4	4	6	1

Examinando os resultados mostrados na Tabela 2.1, percebe-se que ao avaliar quão bem o conjunto A_1 se aproxima do conjunto A_2 , isto é, $I_{\epsilon+}(A_1, A_2) = 1$, não se diagnostica nenhuma diferença no IQ. Porém, abordando o contrário desta análise, alcança-se $I_{\epsilon+}(A_2, A_1) = 2$, e este resultado apresenta a diferença entre essas duas soluções, visto na Figura 2.11. Com isso, nota-se a possível falha na análise, dependendo da escolha de qual conjunto representa a referência. Então, considera-se neste trabalho que a melhor escolha como PF de referência para os futuros experimentos deve ser o conjunto B em $I_{\epsilon+}(A, B)$, conforme formulado na Definição 2.3.2. Vale ressaltar que a implementação deste exemplo foi realizada em Python na versão 11 para a validação do código desenvolvido.

Por fim, a seleção de IQs eficazes é essencial para a comparação entre diferentes MOEAs e suas abordagens, permitindo a avaliação dos conjuntos de soluções obtidos por eles. Neste estudo, optou-se por dois IQs específicos: o IGD^+ e o $I_{\epsilon+}$, devido à sua capacidade de fornecer comparações úteis em MaOPs. No entanto, é importante ressaltar que tanto o IGD^+ quanto o $I_{\epsilon+}$ apresentam limitações. O IGD^+ pode ser sensível a distribuições irregulares de soluções em problemas com muitos objetivos, enquanto o $I_{\epsilon+}$, por ser um indicador binário, exige a comparação de dois conjuntos distintos de soluções e pode não capturar completamente as diferenças sutis de desempenho entre soluções que se destacam em alguns objetivos, mas que apresentam resultados similares em termos gerais.

2.4 Desafios para Solucionar os Problemas com Muitos Objetivos

Ao lidar com um maior número de funções objetivo, surgem desafios ao utilizar um algoritmo multiobjetivo baseado na dominância de Pareto para obter um conjunto representativo de soluções do conjunto Pareto-ótimo. Neste contexto, existem três principais dificuldades que merecem atenção e são detalhadas nesta subseção (Coello Coello et al., 2020) (Nguyen

et al., 2020): Baixa escalabilidade; Alto custo computacional; Dificuldade de visualização das soluções.

2.4.1 Baixa Escalabilidade

A primeira dificuldade na resolução de MaOPs está relacionada à baixa escalabilidade do algoritmo, devido à deterioração da habilidade de busca. Isso ocorre porque, em problemas com um grande número de objetivos, quase toda a população se torna não dominada. Esse fenômeno torna a seleção baseada na dominância de Pareto ineficaz, resultando em uma degradação séria da convergência dos MOEAs (Saxena et al., 2012).

De acordo com Bentley et al. (1978), o número de vetores não dominados de dimensão k em um conjunto de tamanho n é da ordem de $\mathcal{O}(\log^{k-1} n)$. Isso implica que, em problemas com um grande número de objetivos, a seleção de soluções é feita de forma quase aleatória ou guiada pelo estimador de densidade. Efetivamente, Mostaghim e Schneck (2008) mostrou experimentalmente que um otimizador de busca randômica poderia alcançar melhores resultados que o NSGA-II em um problema com dez objetivos. Esse problema é a razão pela qual a maioria dos MOEAs baseados em IQs é capaz de lidar com MaOPs simplesmente utilizando um estimador de densidade mais poderoso para orientar a pesquisa.

Um exemplo ilustrativo desse problema é encontrado em Farina e Amato (2002). Na Figura 2.12, são apresentadas duas formulações de um MOP em $\mathbb{R}^2$ e $\mathbb{R}^3$. O quadrado e o cubo azuis representam as regiões que possuem soluções dominadas pela solução x , enquanto o quadrado e o cubo amarelos são as regiões que possuem soluções que dominam x . Já os quadrados e cubos em branco (linhas tracejadas) representam as regiões que possuem soluções não dominadas.

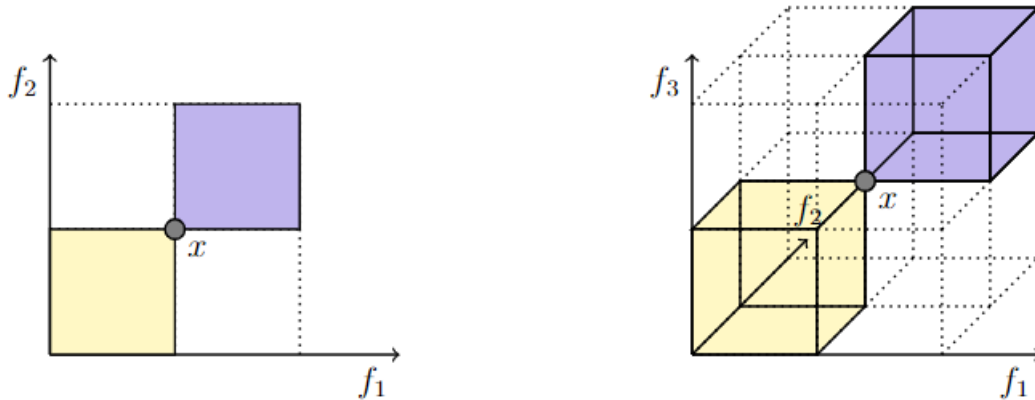


Figura 2.12: Aumento das regiões não dominadas. Fonte: Mendes (2022).

Ao comparar as regiões não dominadas, na formulação com dois objetivos, aproxima-

mente metade das regiões (quadrados tracejados) apresentam soluções não dominadas. No entanto, ao adicionar uma função objetivo, a proporção de regiões com soluções não dominadas aumenta para $6/8$, ou seja, $3/4$, como pode ser observado pelos cubos formados por linhas tracejadas na formulação com três objetivos.

2.4.2 Alto Custo Computacional

A segunda dificuldade está relacionada ao alto custo computacional, devido à chamada “maldição da dimensionalidade”, em que o número de pontos necessários para representar com precisão uma PF aumenta exponencialmente com o número de objetivos.

De acordo com Sen e Yang (2012), o número de pontos necessários para representar uma PF k -dimensional com uma resolução r é da ordem de $\mathcal{O}(kr^{k-1})$. Isso apresenta um desafio e tanto para as estruturas de dados que gerenciam eficientemente esse grande número de pontos na população e no arquivo externo, quanto para os estimadores de densidade que visam obter uma distribuição uniforme das soluções ao longo da PF. Esse desafio é relevante mesmo para os IQs (Ishibuchi et al., 2014).

Na prática, a maioria dos MOEAs tende a usar tamanhos populacionais relativamente pequenos, geralmente com menos de 300 indivíduos, mesmo ao lidar com MaOPs que possuem mais de seis objetivos. No entanto, esses tamanhos populacionais são claramente inadequados para a amostragem de PFs de alta dimensão (Coello Coello et al., 2020).

2.4.3 Dificuldade de Visualização das Soluções

A terceira dificuldade enfrentada é a visualização das soluções devido ao aumento do número de objetivos, tornando-se extremamente desafiador para os DMs escolherem uma solução final em MaOPs (Saxena et al., 2012).

Segundo Miller (1956), do ponto de vista psicológico, as pessoas não lidam bem com um grande volume de dados e têm dificuldade em lidar com mais de 7 ± 2 fatores de uma só vez, o que corresponde a um intervalo de 5 a 9 fatores. Em outras palavras, quando o número de opções aumenta, os indivíduos têm dificuldade em expressar suas preferências e tomar uma decisão. No contexto dos MaOPs, à medida que o número de soluções na PF aumenta, é difícil para uma pessoa comparar e escolher uma solução do conjunto, bem como determinar quais soluções são melhores para cada função objetivo.

É evidente que, com mais de três objetivos ($\mathbb{R}^3$), não é possível traçar a PF da maneira convencional. Isso representa um desafio significativo, já que a visualização desempenha

um papel fundamental em um processo adequado de tomada de decisão. Nos últimos anos, várias técnicas de visualização têm sido propostas para abordar MaOPs Coello Coello et al. (2020).

Uma dessas técnicas é o Gráfico de Coordenadas Paralelas (do inglês, Parallel Coordinate Plot – PCP), proposto por Wegman (1990), onde o eixo y representa as dimensões de um ponto no espaço e o eixo x representa os valores de cada coordenada. Nesse gráfico, cada solução é representada por uma linha que conecta os respectivos valores de cada coordenada.

Outro método de visualização é apresentado na publicação de He e Yen (2015), que mapeia as soluções de um espaço de alta dimensão em um gráfico polar de duas dimensões. Essa visualização preserva a relação de dominância, a localização e o formato da PF, além de manter a distribuição dos indivíduos.

Além disso, em Koochaksaraei et al. (2017), os autores mencionam que gráficos de bolhas em $\mathbb{R}^3$ podem representar soluções em $\mathbb{R}^5$ ao utilizar atributos adicionais, como tamanho e cor das bolhas. No entanto, a interpretação e compreensão dessas visualizações podem ser desafiadoras para os DMs. Como alternativa, eles propõem uma ferramenta chamada DC, que oferece outra forma de visualização de soluções em problemas com dimensões elevadas. Nessa abordagem, os pontos são representados em um formato circular, em que cada coordenada é representada por um arco.

Recentemente, em He et al. (2020), um método foi desenvolvido para identificar várias soluções de interesse em um grande conjunto de soluções na PF. Além disso, a fim de reduzir as dificuldades no processo de tomada de decisão em MaOPs, uma nova visualização em $\mathbb{R}^2$ foi desenvolvida, fornecendo informações sobre a forma e a localização da PF, bem como o grau de convergência e a distribuição das soluções.

2.5 Redução de Dimensionalidade

Para solucionar os desafios mencionados na Seção 2.4, surge a abordagem da redução de dimensionalidade em MaOPs como uma alternativa viável. Seu propósito consiste em eliminar funções objetivo redundantes, a fim de mitigar a complexidade inerente ao problema original. Essa estratégia é alcançada por meio da análise da relação entre as funções objetivo ou por meio da aplicação de técnicas de seleção ou extração de atributos (Brockhoff e Zitzler, 2006a; Ishibuchi et al., 2011).

No contexto de AM e reconhecimento de padrões, um atributo refere-se a uma propriedade mensurável ou característica individual de um fenômeno. Por outro lado, uma classe é

uma descrição que informa a um modelo de AM o significado de um determinado conjunto de atributos, permitindo que ele aprenda com padrões (Bishop e Nasrabadi, 2006). Por exemplo, na Figura 2.13a, o atributo f_1 é relevante, pois pode discriminar entre duas classes. No entanto, na Figura 2.13b, o atributo f_2 é redundante em relação a f_1 , pois está altamente correlacionado com ele. Já na Figura 2.13c, o atributo f_3 é considerado irrelevante, uma vez que não é capaz de separar as classes. Portanto, remover os atributos f_2 e f_3 não afetará negativamente o desempenho do processo de aprendizado.

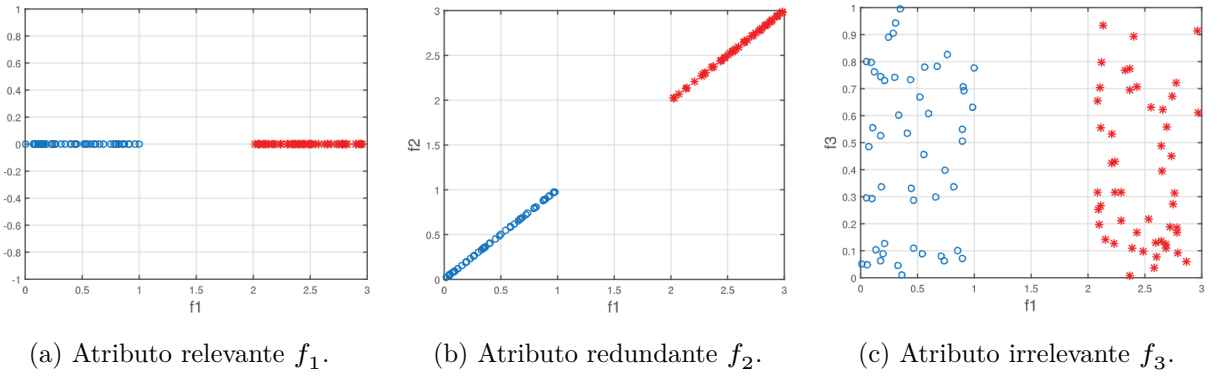


Figura 2.13: Um exemplo ilustrativo de atributos relevantes, redundantes e irrelevantes. Adaptado de: Li et al. (2017).

Notavelmente, os MOEAs geralmente tratam o MaOP simplesmente como um conjunto de dados. Sendo assim, ao aprender a estrutura de preferência dos vetores objetivos das soluções não dominadas obtidas por um MOEA, é possível fornecer um suporte mais efetivo à busca (Duro et al., 2014). Nesse sentido, as funções objetivo podem ser consideradas como atributos de um modelo de AM não supervisionado, isto é, sem informações de classe do mesmo.

A redução de dimensionalidade pode ser categorizada principalmente em duas abordagens principais: seleção de atributos e extração de atributos (Li et al., 2017):

1. Seleção de atributos: Nessa abordagem, seleciona-se um subconjunto relevante de atributos do conjunto de dados original. A ideia é identificar os atributos que têm maior poder preditivo ou explicativo em relação ao problema que está sendo tratado. Isso pode ser feito usando técnicas estatísticas, como a análise de correlação, ou algoritmos de seleção de atributos, como o algoritmo de filtro (que avalia os atributos independentemente do modelo de AM não supervisionado) ou o algoritmo de *wrapper* (que usa um modelo de AM para avaliar a importância dos atributos) ou o algoritmo híbrido (que usa uma combinação dos outros dois).

2. Extração de atributos: Nessa abordagem, criam-se novos atributos (chamadas de componentes) que são combinações lineares dos atributos originais. Essas componentes são selecionadas de forma que preservem a maior parte da variância dos dados originais.

Ao adotar técnicas de seleção de atributos em MaOPs, é importante considerar algumas questões. Primeiramente, a seleção inadequada de atributos leva à perda de informações relevantes se funções objetivo importantes forem removidas. Isso pode ocorrer se a seleção for mal executada ou se houver correlações complexas entre os objetivos. Além disso, se o conjunto de dados contiver ruídos ou valores discrepantes (*outliers*), a seleção de atributos pode ser sensível a esses elementos, resultando na escolha de objetivos incorretos ou irrelevantes.

Em relação à extração de atributos em MaOPs, é crucial avaliar a variância das novas funções objetivo. Caso não possuam um nível adequado de variância, essas combinações podem oferecer menos informações úteis. Além disso, ao extrair atributos complexos, pode ser mais desafiador interpretar o significado dessas novas funções objetivo em relação ao problema original. A relação entre os objetivos extraídos e as características do mundo real pode ser menos clara do que com as funções originais. Comparando com a seleção de atributos, as funções objetivo extraídas, por serem construídas a partir de combinações lineares ou não lineares das funções originais, tendem a ser mais robustas a ruídos e *outliers*.

Em resumo, ao adotar a seleção de atributos em MaOPs, é necessário cuidado para minimizar a perda de informações relevantes e a sensibilidade a ruídos e *outliers*. Já com a extração de atributos, é preciso considerar a variância das novas funções objetivo e a possível dificuldade na interpretação dos resultados. Neste trabalho são adotadas uma abordagem de seleção e uma de extração de atributos, sendo o LS e a PCA, respectivamente. As duas subseções a seguir detalham os procedimentos de cada uma.

2.5.1 Laplacian Score

Introduzido por He et al. (2005), o LS é um dos métodos de seleção de atributos não supervisionado de filtro univariado mais referenciados e relevantes da literatura (Solorio-Fernández et al., 2020). A ideia central por trás do LS é que os atributos discriminantes tendem a revelar estruturas de agrupamento nos dados. O método calcula o escore de cada atributo com base em sua contribuição para a preservação da estrutura de agrupamento (*cluster*) quando os dados são projetados em um espaço de menor dimensão.

Vale ressaltar que tais métodos univariados podem efetivamente identificar e remover atributos irrelevantes, mas são incapazes de remover os redundantes, pois não levam em

conta possíveis dependências entre os atributos (Solorio-Fernández et al., 2020).

De maneira geral, o processo de cálculo do LS envolve a construção de um grafo de similaridade, onde cada ponto de dados é representado por um nó e as arestas indicam a similaridade entre os pontos. Em seguida, é calculada a matriz Laplaciana do grafo, que captura as relações entre os pontos. A partir dessa matriz, é possível determinar o escore de cada atributo. Um atributo com um escore baixo indica que ele é relevante e informativo, enquanto um escore alto indica que o atributo possui pouco poder discriminante. O Algoritmo 2 a seguir trata-se do pseudocódigo que descreve este procedimento detalhadamente:

Algoritmo 2: Laplacian Score para seleção de atributos *offline*

Entrada: Matriz de funções objetivo normalizada $\mathbf{f}_{n \times m}$ com n amostras e m

atributos, número de vizinhos k , constante t

Saída: Escores $L = (L_1, L_2, \dots, L_m)$ para cada atributo

1 **início**

2 **para cada amostra** $\mathbf{f}_i$ **em** $\mathbf{f}_{n \times m}$ **faça**

3 Encontre os k vizinhos mais próximos de $\mathbf{f}_i$ usando a distância euclidiana

4 Conecte $\mathbf{f}_i$ com seus k vizinhos no grafo de vizinhança

/* Calcule a matriz de similaridade S */

5 **para cada par de nós** (i, j) **no grafo faça**

6 **se** i **e** j **estão conectados então**

7 $S_{ij} = \exp\left(-\frac{\|\mathbf{f}_i - \mathbf{f}_j\|^2}{2t}\right)$

8 **senão**

9 $S_{ij} = 0$

/* Calcule a matriz diagonal D e o grafo laplaciano L */

10 Calcule a matriz diagonal $D = \text{diag}(S\mathbf{1})$, onde $\mathbf{1} = [1, \dots, 1]^T$

11 Calcule a matriz laplaciana $L = D - S$

12 **para cada atributo** r **faça**

13 Calcule o vetor ajustado: $\tilde{\mathbf{f}}_r = \mathbf{f}_r - \frac{\mathbf{f}_r^T D \mathbf{1}}{\mathbf{1}^T D \mathbf{1}} \mathbf{1}$

14 Calcule o escore do r -ésimo atributo: $L_r = \frac{\tilde{\mathbf{f}}_r^T L \tilde{\mathbf{f}}_r}{\tilde{\mathbf{f}}_r^T D \tilde{\mathbf{f}}_r}$

Para ilustrar a transformação da estrutura em grafo para a matriz laplaciana no Al-

goritmo 2, considere um conjunto de seis amostras conectadas por relações de vizinhança ilustrado na Figura 2.14:

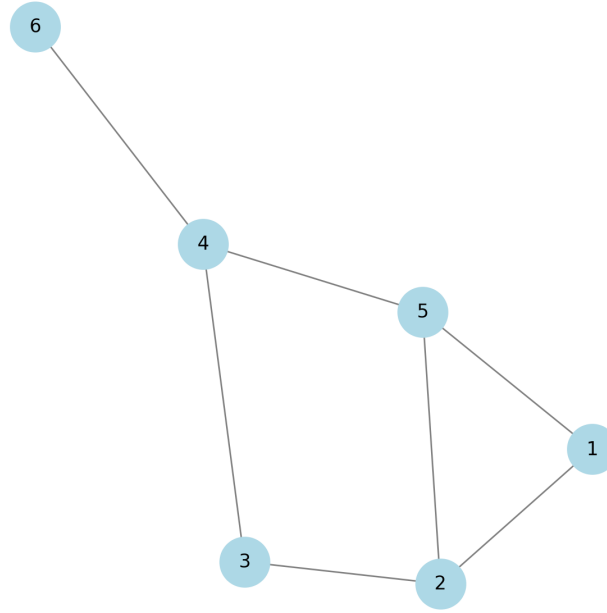


Figura 2.14: Grafo correspondente ao exemplo.

A estrutura das conexões é representada por uma matriz de adjacência, que indica as conexões diretas entre os nós do grafo antes da aplicação da função de similaridade¹. A matriz de adjacência A correspondente ao exemplo é:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Cada linha e cada coluna da matriz A representam um nó no grafo. As entradas A_{ij} indicam se há uma conexão direta entre os nós i e j . Um valor de 1 indica que há uma conexão direta (ou aresta) entre os nós i e j , enquanto 0 indica a ausência de conexão. Por exemplo, $A_{12} = 1$ significa que o nó 1 está conectado ao nó 2.

Com base nessa matriz de adjacência, constrói-se a matriz diagonal D , conforme definido

¹A matriz de similaridade S calculada no Algoritmo 2 está intimamente relacionada à matriz de adjacência A apresentada acima. Enquanto A é binária e apenas indica a presença ou ausência de uma conexão entre os nós (1 para conexão, 0 para ausência de conexão), S reflete a intensidade da relação entre as amostras. Assim, a matriz de similaridade pode ser vista como uma versão “ponderada” da matriz de adjacência, refletindo não apenas a conexão entre os nós, mas também a força dessa conexão com base nas distâncias entre as amostras.

no Algoritmo 2. Essa matriz contém a soma dos pesos das conexões de cada nó, resultando em:

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Assim como na matriz de adjacência, as linhas e colunas da matriz D correspondem aos nós do grafo. As entradas D_{ii} são a soma das conexões de cada nó i , ou seja, a soma de todos os valores da linha i da matriz de adjacência A . Por exemplo, $D_{11} = 2$ porque o nó 1 está conectado aos nós 2 e 5 (dois vizinhos), então a soma de suas conexões é $1 + 1 = 2$.

Por fim, a matriz laplaciana é obtida como $L = D - A$, resultando em:

$$L = \begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Novamente, as linhas e colunas da matriz Laplaciana L correspondem aos nós do grafo. A matriz Laplaciana é calculada subtraindo-se a matriz de adjacência A da matriz diagonal D . Para elementos da diagonal ($i = j$), as entradas L_{ii} são as somas das conexões de cada nó i , conforme já descrito pela matriz D . Para elementos fora da diagonal ($i \neq j$), as entradas L_{ij} são o oposto das entradas da matriz de adjacência, ou seja, $L_{ij} = -A_{ij}$. Isso indica a presença de uma conexão entre os nós i e j .

A fim de ajustar os valores dos dados em uma escala comum, de forma que todas as funções objetivo estejam no mesmo intervalo, utilizou-se a normalização máximo-mínimo. Nesse método, os valores dos dados são ajustados para um intervalo entre 0 e 1. Sua equação é descrita a seguir:

$$f_{norm} = \frac{f - f_{min}}{f_{max} - f_{min}} \quad (2.17)$$

sendo f_{norm} o valor normalizado, f o valor original, f_{min} e f_{max} o valor mínimo e máximo da coluna de amostras da matriz, respectivamente. Ressalta-se que esta normalização foi

aplicada apenas ao LS, visto que a PCA foi calculada utilizando a matriz de correlação dos respectivos problemas de teste, conforme descrito na Subseção 2.5.2 a seguir.

Os k vizinhos mais próximos são determinados utilizando o algoritmo *k-Nearest Neighbours* (k-NN) (Fix e Hodges, 1989), e a matriz de similaridade S do grafo é responsável por capturar a estrutura local do espaço de dados. A escolha do parâmetro k é um aspecto crucial e pode variar conforme as características do conjunto de dados. Em geral, valores maiores de k ajudam a suavizar o impacto do ruído, mas podem levar a limites menos definidos entre as classes. Para conjuntos de dados de tamanho moderado, uma escolha comum é $k = 5$, conforme demonstrado no estudo de He et al. (2005) e adotado neste trabalho. Para uma análise mais aprofundada sobre a definição do tamanho adequado de k em relação ao conjunto de dados, recomenda-se a leitura de Beyer et al. (1999), que aborda de forma detalhada essa relação.

Outro detalhe importante diz respeito à definição da constante adequada t . Esse ajuste é um processo de experimentação e otimização, onde diferentes valores são testados para determinar qual produz os melhores resultados para a tarefa específica em questão.

Uma vez que a equação de S_{ij} é formulada a partir da Função de Base Radial (do inglês, *Radial Basis Function* – RBF), é correto admitir que $t = \sigma_{ij}^2$, ou seja, a covariância entre $\mathbf{x}_i$ e $\mathbf{x}_j$. No entanto, a utilização de $\sigma_{ij}^2 = 1$ em RBF é uma escolha comum e frequentemente conveniente em algumas implementações e aplicações de AM. Essa atribuição pode levar a uma distribuição padrão em muitos contextos, onde valores de distância menores que 1 terão uma contribuição relativamente alta, enquanto valores de distância maiores que 1 terão uma contribuição cada vez menor (Bishop e Nasrabadi, 2006). Essa característica pode ser adequada para o método LS, pois a RBF tenderá a ser mais centralizada, decaindo rapidamente à medida que a distância aumenta. Portanto, neste trabalho, é definido $t = 1$.

No entanto, vale salientar que essa escolha de t não é uma regra fixa, e podem haver casos em que outros valores são mais adequados para resolver um determinado problema. Em alguns cenários, pode ser necessário ajuste manual ou realizar uma busca de hiperparâmetros para otimizar o desempenho do modelo de AM.

Ao calcular os escores de cada função objetivo do MaOP, torna-se possível gerar uma classificação (do inglês, *ranking*) de relevância e informação entre elas. Isso permite a aplicação da redução de dimensionalidade para um número desejado de objetivos, isto é, selecionar um subconjunto que pode variar desde a eliminação de objetivos menos relevantes até a seleção desses objetivos, como, por exemplo, somando-os.

Por fim, para comparar os resultados do MaOP original com os do MaOP reduzido,

foi necessário retornar o MaOP reduzido à sua dimensão original, ou seja, aumentar a sua dimensão. Isso foi feito utilizando a última geração de indivíduos do espaço de busca no MaOP reduzido e aplicando essa geração à formulação do problema original. Dessa forma, ambos os MaOPs passaram a estar na mesma dimensão no espaço de objetivos, permitindo uma comparação direta entre os resultados obtidos.

2.5.2 Análise de Componentes Principais

Introduzido por Pearson (1901) e fundamentado no artigo de Hotelling (1933), o objetivo principal da PCA é explicar a estrutura de variância e covariância de um conjunto de dados, composto de m -atributos, através da construção de combinações lineares dos atributos originais, sendo as constantes destas expressões os autovetores da matriz original de dados. Estas combinações lineares são chamadas de componentes principais e são não correlacionadas entre si.

Multiplicar a matriz de autovetores ordenada pela matriz original de dados resulta na projeção dos dados originais nos novos eixos principais definidos pelos autovetores, esse é um passo fundamental na PCA. A multiplicação das matrizes realiza a transformação linear dos dados originais para um novo espaço, onde os eixos são os autovetores.

Essencialmente, essa multiplicação realiza a transformação dos dados para um novo sistema de coordenadas, onde os eixos são alinhados com as direções de maior variância nos dados originais. Essas direções representam as principais variações ou padrões nos dados, permitindo a identificação das características mais significativas e a redução da dimensionalidade ao destacar os aspectos mais informativos dos dados.

Se temos m -atributos originais é possível obter-se m componentes principais. No entanto, a informação contida nos m -atributos originais é substituída pela informação contida em p ($p \leq m$) componentes principais não correlacionadas. Desta forma, o sistema de variabilidade do conjunto de dados composto dos m -atributos originais é aproximado pelo sistema de variabilidade do conjunto de dados que contém as p componentes principais. A qualidade da aproximação depende do número de componentes mantidas no sistema e pode ser medida através da avaliação da proporção de variância total explicada por essas componentes (Mingoti, 2007). No entanto, é importante notar que, ao reduzir a dimensionalidade, há a perda de informação, pois componentes com menor variação são descartadas, resultando em uma aproximação dos dados originais.

Uma transformação muito usual é a padronização do conjunto de dados pelas respectivas

médias e desvios padrões, gerando-se novos atributos centrados em zero e com variâncias iguais a 1. Neste caso, as componentes principais são determinadas a partir da matriz de covariâncias dos atributos originais padronizados, o que é equivalente a extrair-se as componentes principais utilizando-se a matriz de correlação dos atributos originais.

Uma vez determinadas as componentes principais, os seus valores numéricos podem ser calculados para cada elemento amostral. Deste modo, os valores de cada componente podem ser analisados, usando-se técnicas estatísticas usuais como análise de variância e análise de regressão, dentre outras. No Algoritmo 3 a seguir é descrito o pseudocódigo de extração da matriz de componentes principais via matriz de correlação:

Algoritmo 3: PCA via matriz de correlação para extração de atributos *offline*

Entrada: Matriz de funções objetivo $\mathbf{f}_{n \times m}$

Saída: $Y_{n \times p}$

1 **início**

2 Calcule a matriz de correlação amostral de $\mathbf{f}_{n \times m}$: $P_{m \times m}$

3 Calcule o vetor de autovalores de $P_{m \times m}$: $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$

4 Calcule a matriz de autovetores à direita normalizados de $P_{m \times m}$: $A_{m \times m}$

5 Ordene o vetor de autovalores em ordem decrescente: $\lambda_{1'} \geq \lambda_{2'} \geq \dots \geq \lambda_{m'}$

6 Ordene $A_{m \times m}$ de acordo com seus respectivos autovalores

7 Calcule o vetor de porcentagens de variância total explicadas pelas componentes:

$$\lambda_{\%} = \left(\frac{\lambda_{1'}}{|\lambda|}, \frac{\lambda_{2'}}{|\lambda|}, \dots, \frac{\lambda_{m'}}{|\lambda|} \right)$$

8 **repita**

9 Normalize a matriz de funções objetivo $\mathbf{f}_{n \times m}$ usando z-score

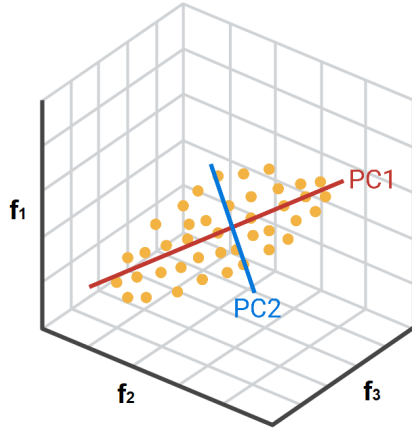
10 Calcule a matriz de componentes principais pela equação matricial:

$$Y_{n \times m} = \mathbf{f}_{n \times m} \times A_{m \times m}$$

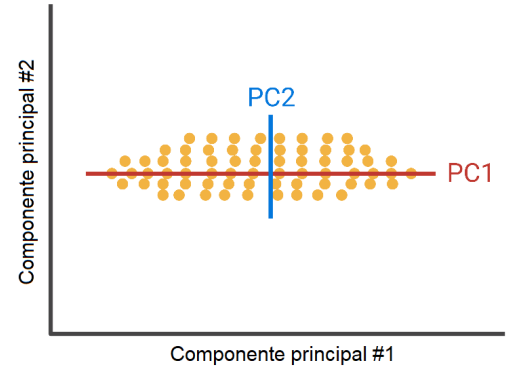
11 Construa uma nova matriz $Y_{n \times p}$, sendo $p = m, \dots, 2$, dependendo do MaOP

12 **até** critérios de parada do MOEA forem satisfeitos

A Figura 2.15 exemplifica a aplicação da PCA em um MOP com três funções objetivo, com a finalidade de reduzir a dimensionalidade do problema, transformando as três funções em duas componentes principais (PC1 e PC2):

Dados originais (alta dimensão)

Redução de
dimensionalidade por
PCA

**Incorporação em
menor dimensão**

- Maximiza a variância ao longo de PC1
- Minimiza os resíduos ao longo de PC2

Figura 2.15: Exemplo PCA.

O processo busca maximizar a variância ao longo de PC1, representando a principal direção de dispersão dos dados, enquanto minimiza os resíduos ao longo de PC2, garantindo que a segunda componente seja ortogonal à primeira e capture as variações restantes. Esse procedimento facilita a análise e visualização do espaço de soluções, mantendo as informações essenciais do MOP.

Neste trabalho, o Algoritmo 3 é empregado para reduzir a dimensionalidade dos MaOPs, preservando as direções mais relevantes (definidas pelos autovetores) e suas respectivas importâncias (associadas aos autovalores). Como os autovetores são ordenados com base na variância que explicam, os primeiros correspondem às direções de maior variabilidade nos dados originais. Detalhes sobre a escolha do tamanho das reduções dimensionais para os problemas de teste são apresentados no Capítulo 4.

Para o cálculo dos autovetores e autovalores, utilizou-se a função `eig` da biblioteca `numpy.linalg` para Python, que é uma ferramenta amplamente utilizada na análise de matrizes e decomposições. Esta função realiza a decomposição espectral de uma matriz quadrada e sua saída é composta especificamente pelos autovetores à direita e pelos autovalores correspondentes. Em notação matemática, se $\mathbf{P}$ é a matriz de correlação e $\mathbf{\Lambda}$ a matriz diagonal de autovalores, então:

$$\mathbf{P} \times \mathbf{A} = \mathbf{A} \times \mathbf{\Lambda} \quad (2.18)$$

A função `eig` calcula esses valores para a matriz $\mathbf{P}$ e retorna a matriz de autovetores $\mathbf{A}$ (cujas colunas são os autovetores). Destaca-se que a função `eig` fornece os autovetores

à direita, o que é crucial para o correto entendimento e aplicação dos resultados obtidos na PCA.

Para manter a concordância com a matriz de correlação utilizada de forma *offline*, é necessário ajustar os valores das funções objetivo de forma que tenham média zero e desvio padrão unitário durante o curso do MOEA. Para isso, utilizou-se a normalização z-score, esse método é fundamental quando os dados precisam ser transformados antes de serem multiplicados pelos autovetores para gerar a matriz de componentes principais. A equação para a normalização z-score é descrita a seguir:

$$f_{norm} = \frac{f - \mu}{\sigma} \quad (2.19)$$

onde f_{norm} é o valor normalizado, f é o valor original, μ é a média da coluna de amostras da matriz, e σ é o desvio padrão da coluna de amostras da matriz.

É importante destacar que a matriz $Y_{n \times p}$ deve, idealmente, possuir dimensões menores do que a matriz $Y_{n \times m}$. No entanto, existem situações em que é necessário incluir todas as componentes para preservar a informação crucial do MaOP, o que pode ser medido pela proporção da variância total explicada pelas componentes principais retidas. Isso implica que o número original de objetivos deve ser mantido ($p = m$), comprometendo a redução da dimensionalidade. Adicionalmente, essas combinações lineares constituem novas funções objetivo, caracterizando a retenção integral dos objetivos originais do MaOP, sem qualquer eliminação.

De forma semelhante ao procedimento realizado no LS, para comparar os resultados do MaOP original com os do MaOP reduzido, é necessário projetar os resultados do espaço reduzido de volta ao espaço de objetivos. Isso é feito aplicando uma transformação linear em $Y_{n \times p}$, ou seja, utilizando a PCA inversa na última geração da população no espaço de objetivos. A função `numpy.linalg.pinv` é utilizada para calcular a pseudo-inversa dos autovetores. A transformação inversa é realizada pela fórmula:

$$\mathbf{f}_{n \times m} = Y_{n \times p} \times A_{m \times m}^+ \quad (2.20)$$

A precisão da transformação inversa é crucial para garantir a validade dos resultados e a interpretação correta, pois qualquer erro pode comprometer a análise. A pseudo-inversa é adotada porque, ao contrário da inversa comum, ela pode ser aplicada a matrizes não quadradas e a matrizes singulares, fornecendo uma solução que minimiza o erro de reconstrução quando a inversa exata não está disponível. No contexto da PCA, a pseudo-inversa permite

a reconstrução aproximada dos dados originais, mesmo quando algumas informações foram perdidas na redução de dimensionalidade. Embora a eficácia dessa aproximação não tenha sido diretamente medida em experimentos neste estudo, ela é amplamente reconhecida na literatura como uma solução válida para a reconstrução dos dados originais de forma aproximada.

2.6 Formulação do Teste de Hipóteses

A fim de analisar estatisticamente a igualdade das médias do IGD^+ e do $I_{\epsilon+}$ para os MaOPs original e reduzido, foi realizado o Teste de Postos Sinalizados de Wilcoxon, porque a distribuição dos resultados é contínua, não paramétrica e as amostras independentes (Walpole et al., 1993). Desta maneira, o teste de hipóteses com nível de significância de $\alpha = 0.05$ é o seguinte:

$$\begin{cases} H_0 : \overline{IQ}_{orig} = \overline{IQ}_{red} \\ H_1 : \overline{IQ}_{orig} \neq \overline{IQ}_{red} \end{cases} \quad (2.21)$$

sendo H_0 a hipótese nula, H_1 a hipótese alternativa, $\overline{IQ}_{orig}$ a média amostral do IQ para o MaOP original e $\overline{IQ}_{red}$ a média amostral do IQ para o MaOP reduzido. A implementação computacional deste teste foi uma adaptação do módulo `stats.wilcoxon`, proveniente do pacote `scipy` dos autores Virtanen et al. (2020).

Especificamente, se a diferença observada estiver contida dentro de $(1 - \alpha)$ do intervalo de confiança, não é possível rejeitar a hipótese nula, indicando que os resultados dos IQs não são estatisticamente significativos. Caso contrário, a hipótese nula é rejeitada, permitindo afirmar que os resultados são diferentes ao nível de significância α . Além disso, os p-valores são calculados para todas as comparações e expostos nas tabelas.

Trabalhos Relacionados

Neste capítulo são explorados os trabalhos relacionados à resolução de MaOPs utilizando MOEAs baseados na abordagem de redução de dimensionalidade, sendo brevemente descritos e discutidos, destacando-se as interconexões existentes entre eles. Inicialmente, é apresentado um panorama histórico relacionado ao tema, fornecendo um contexto bibliográfico relevante. Em seguida, são discutidos os estudos atuais que exploram a aplicação de técnicas de AM não supervisionado nesse contexto. No capítulo seguinte, são apresentadas as bases teóricas que fundamentam o desenvolvimento deste trabalho, oferecendo um maior aprofundamento sobre o assunto.

3.1 Revisão Bibliográfica

Tratando-se da linha de pesquisa sobre o desenvolvimento de MOEAs para enfrentar os desafios dos MaOPs, o artigo de revisão Li et al. (2015) oferece uma base sólida para iniciar a revisão bibliográfica deste trabalho. Os autores apresentam um levantamento dos principais MOEAs existentes, classificando-os em sete categorias distintas, de acordo com suas ideias-chaves. Dentre elas, destaca-se a abordagem baseada em redução de dimensionalidade que, diferentemente das outras, visa lidar com objetivos redundantes.

Essa abordagem pode ser categorizada em duas classes: métodos *offline* e métodos *online*. Em linhas gerais, os métodos *offline* reduzem a dimensionalidade do problema após a obtenção de um conjunto de soluções Pareto-ótimas. Por outro lado, os métodos *online* realizam uma redução gradual da dimensionalidade do problema a cada iteração do MOEA. Em concordância com a Seção 1.3, a proposta deste trabalho se encaixa na classe dos métodos *offline*, por isso não são apontadas as publicações referentes aos métodos *online*. Ao

leitor interessado neste tema, recomenda-se consultar Li et al. (2015).

Seguindo a denominação proposta por Singh et al. (2011), os métodos *offline* podem ser divididos conforme as principais técnicas usadas. As três subclasses são detalhadas a seguir nas Subseções 3.1.1, 3.1.2 e 3.1.3.

3.1.1 Baseados em Correlação

Estes métodos visam examinar a correlação, seja linear ou não linear, entre as funções objetivo do MaOP, por meio da remoção das dependências de segunda ordem e de ordem superior nas soluções não dominadas. A primeira publicação a empregar este conceito foi Deb et al. (2006), na qual os autores propuseram um método para reduzir o número de objetivos com base na PCA, utilizando a correlação linear. Neste método, a correlação entre os objetivos é medida através da matriz de covariância, que é usada para calcular a matriz de correlação linear. A principal suposição é que, se dois objetivos estão negativamente correlacionados, então esses objetivos são conflitantes.

Por meio de uma análise dos autovetores, ou seja, das componentes principais da matriz de correlação linear, é possível identificar os dois objetivos mais conflitantes ao escolher os elementos mais negativos e mais positivos do primeiro autovetor, que carrega mais informações sobre a variância. Para agregar mais objetivos ao conjunto de objetivos conflitantes, o restante dos autovetores é analisado de forma semelhante até que a contribuição cumulativa dos autovalores ultrapasse um limite de corte (do inglês, *threshold cut*).

Seguindo a mesma técnica da publicação anterior, com vistas a obter melhorias na remoção de dependências de ordem superior, os autores em Saxena et al. (2012) propuseram dois algoritmos: o L-PCA, baseado na PCA, e o NL-MVU-PCA, baseado no *Maximum Variance Unfolding* (MVU) (Weinberger e Saul, 2006), para a redução de objetivos linear e não linear, respectivamente.

O MVU é um método baseado em grafos que calcula a representação de baixa dimensão para “desdobrar” o conjunto de dados de alta dimensão. Com o resultado obtido, a PCA pode ser aplicada no novo conjunto de dados. Para o estudo de desempenho, o *framework* foi aplicado em problemas de teste da família DTLZ, com o número de objetivos variando até 50, além de ser utilizado em dois problemas do mundo real e no problema teste WFG3 dos autores Huband et al. (2005), com o número de objetivos escalando até 25. Também foi proposta uma medida de erro para avaliar a qualidade das soluções.

Os resultados apresentados pelos autores na publicação Saxena et al. (2012) evidenciam

que tanto o L-PCA quanto o NL-MVU-PCA, bem como os algoritmos gulosos e exatos baseados na preservação da estrutura de dominância (Subseção 3.1.2), demonstram alta precisão na identificação de conjuntos de objetivos conflitantes em uma variedade de MaOPs, sejam eles redundantes ou não. Essa precisão é alcançada mesmo em conjuntos de soluções caracterizados apenas por soluções Pareto-ótimas que definem a verdadeira PF, ou seja, sem a presença de ruído. Por fim, os autores concluem que, de uma perspectiva prática (Saxena et al., 2012, pág. 33):

“Há uma necessidade de algoritmos de redução de objetivos que possam lidar com conjuntos de soluções tipicamente ruidosos obtidos por MOEAs e determinar um conjunto conflitante de objetivos para uma variedade de problemas, com uma precisão razoavelmente alta e eficiência computacional.”

Assim como os autores citados acima, este trabalho também utiliza a PCA como uma técnica alternativa de redução de dimensionalidade em MaOPs. No entanto, a abordagem adotada difere, pois aqui os coeficientes de combinação linear – autovetores da matriz de correlação – são empregados para extrair novas funções. Em seguida, as componentes com a maior taxa de variância são selecionadas a fim de reduzir o número de objetivos.

3.1.2 Baseados em Estrutura de Dominância

Estes métodos consideram as relações de dominância de Pareto entre as soluções obtidas por toda uma execução de um MOEA. Quanto mais a estrutura de dominância é preservada na população, menos efeitos colaterais causará a perda de informação que acompanha a redução de dimensionalidade.

As primeiras publicações a adotarem esta ideia foram Brockhoff e Zitzler (2006a,b). Nelas, os autores definiram dois tipos de problemas de redução de objetivos e desenvolveram dois algoritmos correspondentes para resolvê-los. Aqui, o conflito é definido usando a mudança na relação de dominância induzida pelo conjunto de objetivos em um conjunto de soluções no espaço de objetivos. Ou seja, se a relação de dominância entre os vetores não muda quando um objetivo é descartado, então esse objetivo não está em conflito com os outros e, portanto, é considerado redundante. Dessa forma, o grau de mudança na relação de dominância induzida no conjunto de soluções pode ser considerado uma medida de conflito entre dois conjuntos de objetivos.

O primeiro problema, denominado Subconjunto Mínimo de Objetivos (do inglês, *Minimum Objective Subset* – δ -MOSS), consiste em encontrar o menor subconjunto de objetivos

que produz um erro δ dado, medido pelo grau de mudança na relação de dominância. Já o segundo problema, denominado k -EMOSS, trata-se de encontrar um subconjunto de objetivos de tamanho k , com o menor erro possível. Uma vez que ambos os problemas são NP-difíceis, os autores propuseram tanto um algoritmo exato quanto um algoritmo guloso para cada um deles.

Os resultados experimentais indicaram que o método proposto não apenas é útil para auxiliar na análise da saída de MOEAs pelos DMs, mas também é qualificado para ser utilizado juntamente de um MOEA da literatura. Além disso, a capacidade deste método foi demonstrada com os resultados em quatro problemas de teste diferentes (problema da mochila multiobjetivo, DTLZ2, DTLZ5 e DTLZ7) e com a comparação em nível ao método de Deb et al. (2006).

Visando progredir com esta linha de pesquisa, os autores propuseram em Brockhoff e Zitzler (2009) uma mudança no cálculo do erro em ambos os problemas δ -MOSS e k -EMOSS, utilizando desta vez o conceito de ϵ -dominância de Laumanns et al. (2002). Já em Brockhoff e Zitzler (2010), diferentemente das outras abordagens, eles apresentaram uma heurística que agrega iterativamente dois objetivos em um novo, utilizando uma combinação linear ponderada. Em citação direta (Brockhoff e Zitzler, 2010, pág. 96):

“Os resultados dos testes estatísticos apoiam a conclusão principal de que a agregação apresenta algumas vantagens em relação à omissão de objetivos, especialmente quando o conjunto de objetivos é reduzido para apenas duas funções.”

Apesar deste trabalho não adotar nenhum método baseado em estrutura de dominância, é importante ressaltar as publicações que foram realizadas nesta linha de pesquisa e como diferentes ideias de redução de dimensionalidade podem ser empregadas na otimização com muitos objetivos.

3.1.3 Baseados em Seleção de Atributos

Estes métodos utilizam técnicas da seleção de atributos não supervisionada (do inglês, *unsupervised feature selection*) para reduzir o número de objetivos. Sua finalidade é identificar as funções que têm maior poder explicativo em relação ao MaOP, o que pode ser alcançado por meio de técnicas de agrupamento (do inglês, *clustering*) ou outros algoritmos de seleção de atributos não supervisionados.

A primeira publicação a adotar esta abordagem foi feita por López Jaimes et al. (2008), onde os autores empregaram uma técnica de seleção de atributos inspirada no trabalho

de Mitra et al. (2002), que utiliza métricas de similaridade para avaliar a redundância e a relevância entre atributos, sem o uso direto de uma matriz de correlação. No entanto, em López Jaimes et al. (2008), foi utilizada uma matriz de correlação linear, derivada de uma matriz de covariâncias, para medir o grau de conflito entre pares de objetivos. De maneira semelhante, Deb et al. (2006) também empregaram uma matriz de correlação linear, calculada a partir de um conjunto de aproximações da PF gerada por um MOEA, para identificar e reduzir conflitos entre objetivos de forma eficiente.

No contexto do trabalho de López Jaimes et al. (2008), o coeficiente de correlação entre um par de objetivos é redefinido para variar em um intervalo entre 0 e 2. Nesse caso, um valor de 0 indica que os objetivos estão totalmente correlacionados positivamente, enquanto um valor de 2 indica que eles estão totalmente correlacionados negativamente. Essa redefinição visa evidenciar os conflitos entre objetivos, de modo que, quanto mais próximo de 2 for o coeficiente, maior será o grau de conflito entre o par de objetivos.

Os autores propuseram dois algoritmos baseados nessa abordagem: o primeiro busca encontrar o subconjunto mínimo de objetivos conflitantes com o menor erro possível, enquanto o segundo visa identificar um subconjunto de tamanho pré-determinado k , também minimizando o erro. De forma geral, esses algoritmos operam particionando o conjunto de objetivos em agrupamentos homogêneos em torno de cada objetivo, utilizando o conflito como uma métrica de distância. Quanto maior o conflito entre um par de objetivos, maior é a distância entre eles no espaço de objetivos. Os agrupamentos mais “compactos” descartam seus objetivos, mantendo apenas a informação do centro de cada agrupamento. Para gerar as soluções não dominadas após a redução, é utilizada uma implementação do *Non-dominated Sorting Genetic Algorithm II* (NSGA-II), conforme descrito em Deb et al. (2002a).

Nos experimentos computacionais, os dois algoritmos propostos foram testados de formas distintas. Para avaliar a efetividade do primeiro algoritmo, os autores o compararam com a abordagem apresentada por Deb et al. (2006), além do algoritmo guloso proposto por Brockhoff e Zitzler (2006a,b), desenvolvido para resolver o problema δ -MOSS com $\delta = 0$. Os problemas abordados foram variações do DTLZ5, e os três algoritmos demonstraram competência na identificação dos objetivos mais conflitantes, evidenciando a robustez da abordagem proposta por López Jaimes et al. (2008).

Para a avaliação do segundo algoritmo, os autores buscaram solucionar o problema k -EMOSS, novamente comparando seus resultados com relação ao algoritmo guloso de Brockhoff e Zitzler (2006a,b) e ao algoritmo de Deb et al. (2006). No entanto, dado que o último não aborda diretamente o problema k -EMOSS, os autores experimentaram dife-

rentes valores de limite de corte para obter conjuntos de objetivos reduzidos com tamanhos variados. A avaliação contemplou duas instâncias com 10 e 20 objetivos do conhecido problema da mochila 0/1, composto por 100 itens, bem como duas instâncias com 10 e 20 objetivos de uma variação do problema DTLZ2 (Brockhoff e Zitzler, 2007). Além disso, o erro δ e o IQ *Inverted Generational Distance* (IGD) foram utilizados para comparação dos três algoritmos.

Os resultados experimentais mostraram que os algoritmos propostos possuem uma vantagem sobre os dois algoritmos de referência em seu baixo tempo computacional. Por último, em citação direta aos autores (López Jaimes et al., 2008, pág. 680):

“Embora, em geral, as medidas de desempenho tenham resultados consistentes, em alguns casos específicos produziram resultados contraditórios. Uma possível explicação para esse comportamento é que um bom desempenho de um algoritmo de redução de objetivos em relação à tomada de decisão não implica em um bom desempenho em relação ao processo de busca em geral. No entanto, mais experimentos são necessários para provar ou refutar esta afirmação.”

No caso deste trabalho, é adotado o LS como algoritmo de seleção de atributos para redução de dimensionalidade. Nesta abordagem, os objetivos são tratados de forma individual, em vez de pares, pois o LS calcula uma medida de preservação da estrutura de agrupamento. Portanto, os objetivos com conjunto de dados mais “compactos” são considerados mais representativos para o MaOP. Por fim, ressalta-se que este trabalho não realiza comparações diretas com resultados de outros artigos da referência, bem como não utiliza as métricas de erro até aqui apresentadas.

3.1.4 Outras Publicações

Com o propósito de apresentar outros artigos para o leitor interessado sobre esta linha de pesquisa, a seguir são listadas e brevemente descritas quatro publicações que utilizam outras técnicas para redução de dimensionalidade em MaOPs:

1. No artigo de Zhou et al. (2009), os autores propuseram um algoritmo para identificar o objetivo mais redundante de um problema, visando obter um conjunto reduzido de funções que facilite a busca e a tomada de decisão. O método utilizado é baseado nos mínimos quadrados, o qual ajusta cada objetivo em uma linha, concentrando-se exclusivamente na sua inclinação. Em seguida, são selecionados dois objetivos redundantes

com base no valor mínimo das diferenças entre suas inclinações. Por fim, identifica-se como redundância aquele cuja média das inclinações é menor em relação às demais. Além disso, o algoritmo utiliza um parâmetro para avaliar a presença de objetivos redundantes.

2. No trabalho de Sinha et al. (2013), os autores apresentaram um *framework* que combina uma simplificação do problema guiado pelo NL-MVU-PCA (Saxena et al., 2012), com um procedimento de solução fundamentado no envolvimento interativo do DM, conhecido como PI-EMO-VF.
3. Em de Freitas et al. (2015), é apresentado um método baseado em árvore para visualizar a relação entre objetivos, utilizando medidas de conflito e harmonia para inferir sua redutibilidade. Além disso, os nós de agregação são coloridos nas visualizações com base em uma medida de região de conflito, indicando onde o conflito não paramétrico ocorre com maior intensidade.
4. Por último, na publicação de Wang e Yao (2016), os autores empregaram uma métrica não linear, denominada NCIE, a fim de medir a correlação entre os objetivos. Além disso, utilizaram um método de seleção de objetivos simples sem nenhum parâmetro pré-definido, o que resulta na robustez da abordagem proposta.

3.2 Estado da Arte

O artigo de Mendes et al. (2017) tem como foco o Problema de Roteamento de Veículos com Transporte Reativo à Demanda (do inglês, *Vehicle Routing Problem with Demand Responsive Transport - VRPDRT*), um tipo de serviço que possibilita o deslocamento de passageiros até seus destinos de forma compartilhada, buscando minimizar os custos da empresa e oferecer um serviço de qualidade, atendendo às necessidades dos passageiros.

Uma modelagem de MaOP é aplicada ao VRPDRT, com sete funções objetivo diferentes. Para resolver este problema através de MOEAs tradicionais, os autores propuseram o uso da análise de agrupamentos para redução de dimensionalidade, especificamente o método de ligação completa. Sendo assim, a matriz de correlação entre as sete funções é calculada, depois ela é transformada em uma matriz de distância subtraindo a matriz identidade, e uma formulação com dois objetivos é selecionada com base na análise do dendrograma. Para a busca das soluções não dominadas são utilizados o NSGA-II e o *Strength Pareto Evolutionary Algorithm 2* (SPEA2) de Zitzler et al. (2001).

Os resultados dos experimentos computacionais revelaram que os MOEAs alcançaram desempenhos estatisticamente distintos, com o NSGA-II obtendo um maior número de soluções não dominadas em comparação ao SPEA2. Adicionalmente, esses resultados foram comparados a uma abordagem presente na literatura que utiliza outra forma de redução da dimensionalidade deste problema, também em uma formulação de dois objetivos. Por fim, os autores concluíram que o procedimento de análise de agrupamentos mostrou-se uma metodologia competitiva nesse contexto e a forma como a redução é realizada tem um impacto significativo no comportamento do MOEA.

É importante ressaltar que a redução pode ser realizada para resultar em um número diferente de objetivos e que outros métodos da análise de agrupamentos podem – e devem – ser empregados, por exemplo, o método de ligação simples. Cada método pode gerar dendrogramas distintos e cada problema com suas particularidades deve ser analisado pelo DM. Embora este trabalho não utilize a análise de agrupamentos, é importante ressaltar que pesquisas que adotam abordagens de redução de dimensionalidade na resolução de MaOPs do mundo real demonstram que fomentar essa linha de pesquisa pode ser útil para enfrentar antigos e novos desafios.

Outra publicação recente, conduzida pelos autores Nguyen et al. (2020), teve como propósito investigar a interação entre MOEAs e algoritmos de redução de dimensionalidade. O estudo avaliou o impacto de um MOEA tradicional no desempenho de um algoritmo de redução de objetivos, além dos benefícios que um MOEA mais recente pode trazer quando combinado com esse tipo de técnica. Para o processo de redução de objetivos, foi implementado o L-PCA de Saxena et al. (2012). Foram utilizados quatro MOEAs: os tradicionais NSGA-II e SPEA2, e os mais recentes NSGA-III e SPEA2+SDE (Li et al., 2013).

Os resultados experimentais do estudo revelaram que a qualidade dos MOEAs desempenha um papel significativo no desempenho do L-PCA. A combinação do L-PCA com o NSGA-III e o SPEA2+SDE foi capaz de remover com sucesso objetivos redundantes, mesmo em situações com um grande número de objetivos originais. No entanto, ao combinar o L-PCA com o NSGA-II e o SPEA2, a remoção de objetivos redundantes foi menos efetiva, especialmente quando o número de objetivos originais era grande.

Para destacar os benefícios de um algoritmo de redução de dimensionalidade em conjunto com um MOEA recente, foi comparado o desempenho de diferentes combinações de MOEAs com o L-PCA em relação ao uso isolado do MOEA. O desempenho foi avaliado utilizando-se as métricas *Generational Distance* (GD) (Van Veldhuizen, 1999) e IGD. Os resultados mostraram que a abordagem com o algoritmo de redução de dimensionalidade teve um

desempenho significativamente superior em relação ao uso exclusivo do MOEA.

É importante ressaltar que o estudo concentrou-se exclusivamente no L-PCA para a redução de dimensionalidade, e há outros métodos que merecem ser avaliados em combinação com esses MOEAs. Além das duas publicações recentes mencionadas, existem outras duas que merecem atenção, e suas ideias são brevemente descritas a seguir:

1. Em Li e Zhang (2021), os autores apresentaram uma abordagem multitarefa para a seleção de atributos visando a redução de objetivos. Nesse método, cada objetivo é formulado como uma combinação linear positiva de um pequeno conjunto de objetivos essenciais, e a regularização esparsa é utilizada para identificar objetivos redundantes. O desempenho do algoritmo proposto, denominado MTFSOR, foi avaliado por meio de comparações com diversos métodos de ponta para a redução de objetivos em problemas de teste, como a família DTLZ. Os resultados obtidos demonstraram a eficácia e robustez do MTFSOR.
2. Por último, em Yao et al. (2021), foi apresentado um método para resolver MaOPs de larga escala (LSMaOPs). Esse método é baseado na redução de dimensionalidade e utiliza um Algoritmo Evolutivo Guiado pelo Conhecimento (KGEA) para a busca das soluções não dominadas.

A abordagem de redução de dimensionalidade aplicada agrupa e agrega as funções objetivo com base em sua correlação, resultando em uma efetiva redução da dimensão do LSMaOP original. Isso contribui para uma melhor representação de diferentes preferências, uma vez que as correlações entre os objetivos reduzidos são mantidas relativamente baixas. Posteriormente, o KGEA é empregado para resolver o LSMaOP transformado.

Experimentos Computacionais

Este capítulo apresenta os experimentos computacionais realizados para avaliar a eficácia das abordagens de redução de dimensionalidade propostas. As implementações dos problemas da família DTLZ¹ e do NSGA-III² foram baseadas no *framework* `pymoo`, criado por Blank e Deb (2020). Desenvolvido em Python com foco na pesquisa acadêmica, esse pacote visa contribuir para a área de otimização, oferecendo ferramentas para a resolução e análise de problemas mono e multiobjetivo. Com o intuito de integrar melhor os algoritmos propostos, todas as ferramentas computacionais desenvolvidas neste trabalho foram criadas e adaptadas em Python na versão 11. Adicionalmente, a implementação do LS foi baseada no pacote `scikit-feature`, conforme descrito por Li et al. (2017).

4.1 Parâmetros Atribuídos aos Algoritmos

Uma vez formulados os problemas de teste, o próximo passo foi gerar seus conjuntos de dados para aplicação dos Algoritmos 2 e 3. Para realizar a análise, foram geradas matrizes $\mathbf{f}_{2500 \times m}$ por meio da extração de 2500 amostras aleatórias simples de cada função objetivo nos sete problemas da família DTLZ, considerando os cenários com $m = 5, 10, 15, 20$ e 30 objetivos. Vale ressaltar que estes 35 bancos de dados são os mesmos para ambos os métodos em *offline*. Sendo assim, devido à dimensão destas matrizes, tornou-se impraticável inserir esses resultados no corpo do texto, pois isso comprometeria a legibilidade.

Visto que o NSGA-III inicia com a descrição de um número de conjuntos predefinidos de direções de referência em um simplexo unitário, o `pymoo` disponibiliza diferentes métodos

¹Todas as alterações realizadas foram testadas para garantir conformidade com a produção original do *framework* `pymoo`.

²Detalhes de implementação deste algoritmo podem ser encontrados em Blank et al. (2019).

para realizar este procedimento. Neste trabalho, foi escolhido o Riesz s-Energy.

Publicado por Blank et al. (2020), os autores propõem o Riesz s-Energy como um novo método para criar um conjunto de pontos bem espaçados para qualquer número arbitrário de pontos e dimensões, visando também atender à necessidade da pesquisa em MOEAs a respeito de uma abordagem mais genérica para amostras de direções de referência. Desta maneira, o número de pontos escolhido neste trabalho foi o tamanho da população de indivíduos, bem como a dimensão, sendo a mesma do espaço de objetivos.

Os parâmetros do NSGA-III foram adotados em conformidade com os padrões estabelecidos na literatura original. A seleção por torneio, o cruzamento SBX com um índice de distribuição de 30 e probabilidade de cruzamento de 1, além da mutação polinomial com um índice de distribuição de 20 e uma probabilidade de mutação especificada como $1/n$, sendo n o número de variáveis de decisão.

Para cada um dos cenários com $m = 5, 10, 15, 20$ e 30 , o tamanho da população escolhido foi de 300 indivíduos e o limite superior para o número de avaliações de função objetivo foi de 30000, portanto, o número de gerações para cada execução do NSGA-III é de 100. Foram adotadas 21 execuções, sendo assim, 21 PFs e um resultado de IGD^+ para cada (as tabelas apresentam a média destes resultados). Apesar de cada execução utilizar uma semente diferente entre elas, a população inicial aleatória é a mesma para todas as execuções em cada cenário.

Adicionalmente, conforme descrito na Seção 2.3, para os cálculos dos IQs é necessário o conjunto de soluções de referência. Neste trabalho, foi utilizada a combinação das 21 PFs obtidas em cada cenário. Este procedimento foi feito por meio da aplicação da relação de dominância de Pareto entre estas soluções, resultando assim em uma PF combinada (ou “final”) sem limitação do número de indivíduos, para comparar com o resultado de cada execução individualmente e calcular o respectivo IGD^+ e I_e (Knowles et al., 2006).

Além disso, os dois métodos escolhidos compartilham a mesma estrutura na apresentação dos resultados para análise. Em cada um dos cinco casos selecionados para cada DTLZ, os resultados são organizados em tabelas de escores (L_r) para o LS e de porcentagens de variância total explicada ($\lambda_{\%}$) para a PCA. São apresentadas também as tabelas que contêm os resultados dos IQs e seus respectivos *boxplots* das 21 execuções. Todos os cenários foram submetidos à análise estatística e, posteriormente, discutidos.

Por fim, para determinar o número de objetivos após a redução, adotou-se uma abordagem de decrescimento progressivo até que se chegasse a um MOP com dois objetivos. Essa metodologia também possibilita a análise de eventuais padrões nas respostas dos IQs. Com

os detalhes do procedimento experimental apresentados, passa-se à exposição dos resultados e à discussão dos experimentos.

4.2 Laplacian Score: Análise dos Resultados

Com o conjunto de dados normalizado $\mathbf{f}_{2500 \times m}$, onde $m = 5, 10, 15, 20$ e 30 , o LS foi aplicado conforme o Algoritmo 2 descrito na Subseção 2.5.1. No Apêndice A, são apresentadas as tabelas com os escores das funções objetivo (L_r), ordenados de forma crescente, para cada caso tratado. Ressalta-se que escores mais baixos indicam maior relevância da função objetivo para o MaOP, permitindo a redução de dimensionalidade. A seguir, são discutidos os resultados para cada problema DTLZ.

Conforme evidenciado no Apêndice A, os resultados para o DTLZ1 (Seção A.1) mostram um padrão em que as últimas funções objetivo apresentam escores mais baixos, sugerindo maior relevância. Esse comportamento é explicado pela multiplicação de variáveis de decisão na formulação do DTLZ1, onde a quantidade de multiplicadores diminui à medida que o índice da função objetivo aumenta, resultando em soluções menos esparsas. No entanto, à medida que a dimensão do problema aumenta, a inversão da ordem dos escores torna-se menos rigorosa, provavelmente devido à maior dispersão das amostras. A redução de dimensionalidade mostrou-se eficaz para melhorar a convergência do NSGA-III nos cenários $m = 5, 10, 30$.

Os resultados para o DTLZ2 (Seção A.2) seguem um padrão semelhante ao do DTLZ1, com as últimas funções objetivo apresentando maior relevância. Esse comportamento é justificado pela mesma estrutura de multiplicação de variáveis de decisão. No entanto, à medida que a dimensionalidade do problema foi reduzida, o desempenho do MOEA piorou gradualmente, conforme evidenciado pelos valores de IQs. A dispersão dos resultados, avaliada por meio de *boxplots*, foi baixa para o IGD^+ , especialmente no cenário biobjetivo.

O DTLZ3 (Seção A.3), que compartilha as mesmas funções objetivo do DTLZ2, apresentou resultados de escores semelhantes. A seleção de atributos mostrou-se eficaz nos cenários $m = 5$, $m = 20$ e $m = 30$, com evidências estatísticas de melhoria no desempenho do MOEA. Nos demais cenários, a hipótese nula não foi rejeitada. O objetivo do DTLZ3, que é avaliar a capacidade de convergência para a PF global, explica os ganhos de desempenho observados após a redução de dimensionalidade.

A modificação no parâmetro α do DTLZ4 (Seção A.4) resulta em um conjunto denso de soluções, o que afeta a análise do LS. Os escores calculados foram baixos, refletindo a

concentração das amostras e a incerteza sobre a relevância das funções objetivo. Além disso, a ordem dos escores não seguiu o padrão observado nos DTLZ anteriores, provavelmente devido à aleatoriedade das variáveis de decisão. Experimentos com diferentes sementes podem produzir resultados variados, sem um padrão claro de *ranking*.

O DTLZ5 (Seção A.5), que testa a capacidade de convergência para uma curva específica, apresentou resultados semelhantes aos do DTLZ1 e DTLZ3. A redução de dimensionalidade mostrou ganhos de desempenho na maioria dos cenários, com escores formando um *ranking* consistente. Essa consistência é justificada pela estrutura de multiplicação de variáveis de decisão, que diminui à medida que o índice da função objetivo aumenta. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), o LS demonstrou eficácia em reduções moderadas e agressivas ($m = 6, 4, 2$), com melhorias consistentes nos IQs e baixa dispersão dos resultados.

O DTLZ6 (Seção A.6), uma variação do DTLZ5, foi projetado para dificultar a convergência do MOEA. Apesar disso, a redução de dimensionalidade mostrou ganhos de desempenho, com escores seguindo um padrão semelhante aos outros DTLZ. A ordem dos *rankings* foi consistente, reforçando a eficácia da abordagem de seleção de atributos. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), o LS manteve sua eficácia, com ganhos de desempenho em reduções agressivas ($m = 3, 2$) e dispersão consistentemente baixa. A redução para $m = 2$ destacou-se novamente, com resultados estatisticamente significativos e dispersão mínima.

O DTLZ7 (Seção A.7) apresentou um padrão de aleatoriedade na ordenação dos escores, exceto para a última função objetivo, que foi calculada de forma distinta. A piora gradual dos IQs com o aumento da dimensionalidade sugere que a abordagem de seleção de atributos pode não ser tão eficaz para este problema. Além disso, a concentração das soluções em uma única região Pareto-ótima pode introduzir vieses nos resultados. Para uma avaliação mais completa, sugere-se a incorporação de métricas adicionais que avaliem a distribuição uniforme das soluções, bem como a aplicação de IQs complementares mencionados na literatura.

4.3 PCA: Análise dos Resultados

Com o conjunto de dados não normalizado, a PCA foi aplicada conforme o Algoritmo 3 descrito na Subseção 2.5.2. Para evitar comprometer a legibilidade, as matrizes de correlação amostral e de autovetores não foram incluídas no corpo do texto. As tabelas no Apêndice B apresentam as porcentagens de variância total explicada ($\lambda_{\%}$) das componentes extraídas ($Y_{n \times m}$), ordenadas de forma decrescente, para cada caso tratado. Conforme a PCA, quanto

maior $\lambda_{\%}$, mais variância a componente representa para o MaOP, permitindo a redução de dimensionalidade ao eliminar componentes com menor variância. A seguir, são discutidos os resultados para cada problema DTLZ.

Conforme evidenciado no Apêndice B, os resultados para o DTLZ1 (Seção B.1) mostram que não há um valor único de variância total explicada adequado para todas as dimensões. Observa-se também uma inconsistência entre as médias dos IQs: uma determinada variância pode ser suficiente para melhorar uma métrica, mas não outra. A redução de dimensionalidade contribui para melhorar a eficiência do NSGA-III, especialmente em problemas de alta dimensionalidade, mitigando a dificuldade de convergência para o hiperplano Pareto-ótimo proposto pelo DTLZ1. Em cenários de menor dimensionalidade ($m = 5, 10$), a PCA mostrou-se eficaz, com ganhos significativos nos IQs para reduções moderadas ($m = 4, 6$). No entanto, em cenários de maior dimensionalidade ($m = 15, 20, 30$), a eficácia da PCA diminuiu, com resultados estatisticamente insignificantes para algumas reduções.

Os resultados para o DTLZ2 (Seção B.2) exibem um comportamento semelhante ao do DTLZ1, com variabilidade total das componentes não sendo única para todos os cenários. A PCA parece atenuar o desafio de lidar com um grande número de objetivos, preservando informações relevantes do modelo por meio de combinações lineares das funções. Isso resulta em melhorias no desempenho do MOEA, conforme sugerem os resultados positivos. Em cenários de menor dimensionalidade ($m = 5, 10$), a PCA mostrou-se eficaz, com ganhos significativos nos IQs para reduções moderadas ($m = 4, 6$). No entanto, em cenários de maior dimensionalidade ($m = 15, 20, 30$), a eficácia da PCA diminuiu, com resultados estatisticamente insignificantes para algumas reduções.

O DTLZ3 (Seção B.3), que compartilha as mesmas funções objetivo do DTLZ2, apresentou resultados consistentes com os problemas anteriores. A PCA mostrou-se eficaz em melhorar a convergência para a PF global, especialmente nos cenários de menor dimensionalidade. A inconsistência entre as médias dos IQs persiste, reforçando a necessidade de uma análise cuidadosa ao escolher o valor de variância total explicada. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), a PCA mostrou ganhos de desempenho em algumas reduções específicas ($m = 6, 8, 10$), mas com resultados estatisticamente insignificantes em outros casos.

No DTLZ4 (Seção B.4), as variâncias não se concentram fortemente nas primeiras componentes principais, devido à natureza da sua PF, onde as soluções são agrupadas em regiões específicas. No entanto, as primeiras componentes ainda capturam as direções mais relevantes, melhorando a convergência e a diversidade das soluções. A PCA simplifica o problema

de otimização, facilitando a busca de soluções pelo NSGA-III. Em cenários de menor dimensionalidade ($m = 5, 10$), a PCA mostrou-se eficaz, com ganhos significativos nos IQs para reduções moderadas ($m = 4, 6$). No entanto, em cenários de maior dimensionalidade ($m = 15, 20, 30$), a eficácia da PCA diminuiu, com resultados estatisticamente insignificantes para algumas reduções.

O DTLZ5 (Seção B.5) apresentou maior concentração de variância na primeira componente ($\approx 70\%$), diferentemente dos demais problemas da família DTLZ. Isso resulta em uma variância total alta, mesmo em reduções com apenas duas componentes. A PCA aprimorou o desempenho do NSGA-III na convergência para uma curva específica, conforme o objetivo do DTLZ5. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), a PCA mostrou ganhos de desempenho em algumas reduções específicas ($m = 6, 8, 10$), mas com resultados estatisticamente insignificantes em outros casos.

O DTLZ6 (Seção B.6), uma variação do DTLZ5, apresentou um comportamento mais próximo ao dos problemas DTLZ1, 2 e 3, com concentração de variância nas primeiras componentes. Apesar do aumento na dificuldade de convergência proposto pelo DTLZ6, a PCA mostrou-se eficaz em melhorar o desempenho do NSGA-III, embora inconsistências nas médias dos IQs tenham sido observadas em alguns cenários. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), a PCA mostrou ganhos de desempenho em algumas reduções específicas ($m = 6, 8, 10$), mas com resultados estatisticamente insignificantes em outros casos.

O DTLZ7 (Seção B.7) apresentou uma grande disparidade entre os valores dos IQs das reduções em comparação ao problema original. Embora isso sugira ganhos significativos, é importante investigar se as soluções se concentraram em apenas uma região Pareto-ótima, o que poderia enviesar os resultados. Novas análises, como visualizações das subpopulações não dominadas ou o uso de IQs complementares, são recomendadas para validar a efetividade da abordagem. Em cenários de maior dimensionalidade ($m = 15, 20, 30$), a PCA mostrou ganhos de desempenho em algumas reduções específicas ($m = 6, 8, 10$), mas com resultados estatisticamente insignificantes em outros casos.

4.4 Comparação dos Resultados de LS e PCA

A Tabela 4.1 e a Tabela 4.2 resumem os casos em que os métodos LS e PCA apresentaram melhorias estatisticamente significativas nos IQs, representadas por $\checkmark$, e aqueles em que não houve melhoria, representados por $\times$. A análise comparativa destaca o desempenho relativo

de cada método em diferentes cenários de dimensionalidade ($m = 5, 10, 15, 20, 30$) para os problemas DTLZ1 a DTLZ7.

Tabela 4.1: Resultados comparativos de LS e PCA - Parte 1 (DTLZ1 a DTLZ4)

DTLZ	m	LS	PCA
DTLZ1	5	✓	✓
	10	✓	✓
	15	×	✓
	20	×	✓
	30	✓	✓
DTLZ2	5	×	✓
	10	×	✓
	15	×	✓
	20	×	✓
	30	×	✓
DTLZ3	5	✓	✓
	10	×	✓
	15	×	✓
	20	✓	✓
	30	✓	✓
DTLZ4	5	×	✓
	10	×	✓
	15	×	✓
	20	×	✓
	30	×	✓

Tabela 4.2: Resultados comparativos de LS e PCA - Parte 2 (DTLZ5 a DTLZ7)

DTLZ	m	LS	PCA
DTLZ5	5	×	✓
	10	✓	✓
	15	✓	✓
	20	✓	✓
	30	✓	✓
DTLZ6	5	✓	✓
	10	✓	✓
	15	✓	✓
	20	✓	✓
	30	✓	✓
DTLZ7	5	×	✓
	10	✓	✓
	15	✓	✓
	20	✓	✓
	30	✓	✓

De forma geral, a PCA apresenta um desempenho superior ao LS na maioria dos cenários testados. Para problemas com um número menor de funções objetivo (5 e 10), o LS oferece resultados comparáveis à PCA em alguns casos, como no DTLZ1 e DTLZ3, mas a vantagem da PCA fica evidente à medida que o número de funções aumenta. Especificamente, em DTLZ1 e DTLZ3, a PCA supera o LS em todos os cenários com 15 ou mais funções objetivo. Em outros problemas, como DTLZ2 e DTLZ4, o LS não obteve resultados positivos, enquanto a PCA mostrou-se consistente, alcançando melhorias significativas em todos os casos com mais de 5 funções.

Em resumo, a PCA se mostrou mais eficaz, particularmente em cenários de alta dimensionalidade, o que reflete sua capacidade superior de lidar com MaOPs, enquanto o LS teve

desempenho mais limitado, sendo eficiente apenas em alguns casos com menor número de objetivos.

Conclusões

Este estudo propôs a integração de técnicas de redução de dimensionalidade a um MOEA com o intuito de reduzir a complexidade de MaOPs de *benchmark* da literatura, diminuindo o número de funções conflitantes para duas ou três. O objetivo foi alcançar resultados comparáveis ou superiores aos obtidos pelos MOEAs de última geração aplicados ao MaOP original. Para isso, foi realizada uma extensa revisão da literatura, cobrindo as abordagens mais relevantes para redução de dimensionalidade e as publicações mais recentes que direcionam essa área de pesquisa.

As bases teóricas deste trabalho foram apresentadas com a formulação geral dos MaOPs, definições-chave no campo da otimização multiobjetivo, e as formulações específicas dos problemas de teste da família DTLZ. Também foi abordada a origem dos MOEAs e a base de seus funcionamentos, com especial destaque para o NSGA-III, incluindo uma análise de sua estrutura. Além disso, foram introduzidos os IQs IGD^+ e $I_{\epsilon+}$, que serviram como métricas principais de desempenho, junto com uma discussão dos desafios específicos na solução de MaOPs.

Adicionalmente, discutiram-se os princípios da redução de dimensionalidade no contexto da otimização multiobjetivo, com ênfase nos dois métodos de AM não supervisionado aplicados: a seleção de atributos pelo LS e a extração de atributos pela PCA. Foram descritos os algoritmos correspondentes de forma detalhada e clara, estabelecendo as bases metodológicas para a condução dos experimentos. Por fim, delineou-se o planejamento dos experimentos computacionais e os resultados foram apresentados e discutidos em profundidade para cada um dos cinco cenários dos sete DTLZ.

A seguir, estão destacados alguns pontos importantes observados durante a produção desta pesquisa:

-
- Conforme concluído por Brockhoff e Zitzler (2010) e citado na Subseção 3.1.2, a redução de dimensionalidade por meio da seleção de atributos resultou na perda de informações ao se aplicar a eliminação de funções. Essa perda foi evidenciada pelos resultados dos IQs, que se mostraram inferiores em comparação com a abordagem baseada em agregação. Embora esta análise não tenha sido o foco principal deste trabalho, os resultados apresentados sugerem uma possível confirmação das conclusões previamente estabelecidas na literatura.

Em abordagens onde as funções objetivo são descartadas, como em Deb et al. (2006), alguma informação ainda é preservada por meio dos chamados objetivos compactos (de baixo conflito). Nesse caso, os agrupamentos formados descartam objetivos considerados redundantes ou menos conflitantes, enquanto preservam a informação dos “centros” desses agrupamentos. Isso implica que, apesar de algumas funções objetivo serem eliminadas, a técnica retém uma representação das informações eliminadas por meio desses centros.

Em resumo, não ocorre uma eliminação total das informações associadas aos objetivos descartados, mas sim uma simplificação do problema. Essa simplificação preserva as informações essenciais, garantindo que o comportamento global do conjunto de objetivos seja mantido, mesmo após a redução dimensional.

- A utilização *offline* dos autovetores gerados para realizar a transformação inversa da PCA deve ser cuidadosamente considerada, uma vez que é crucial avaliar a coerência dos dados. Se a distribuição dos dados mudar significativamente durante as iterações do MOEA, a transformação inversa utilizando esses autovetores pode introduzir erros. Por outro lado, se a distribuição dos dados permanecer relativamente estável ao longo do tempo, o uso desses autovetores continua sendo válido. No entanto, em cenários de alta variação, é recomendável reavaliar periodicamente os autovetores gerados ou empregar técnicas adaptativas de PCA, que ajustam os autovetores dinamicamente conforme as mudanças na distribuição dos dados.
- Após a redução dimensional pela PCA, em alguns casos foi percebida uma inconsistência entre as médias dos IQs, que pode ser atribuída às características intrínsecas dessas métricas. O IGD^+ mede a proximidade da solução à verdadeira PF, enquanto o $I_{\epsilon+}$ avalia a dominância relativa entre conjuntos de soluções. Ou seja, em alguns MaOPs, a proximidade das soluções melhora (menor IGD^+), mas a dominância relativa (medida pelo $I_{\epsilon+}$) pode não melhorar proporcionalmente, refletindo a diferença inerente

entre esses indicadores. Além disso, a presença de indivíduos negativos na PCA pode contribuir para essa discrepância, onde a redução pode beneficiar uma métrica mais do que a outra.

- Ainda nos resultados da PCA, notou-se uma diferença nas dispersões entre os resultados dos MaOPs reduzidos e os originais, que pode ser atribuída à aplicação da transformação inversa pela Equação 2.20, algo que não ocorre nos espaços de objetivo originais. Ambos os problemas mencionados podem ser mitigados ao adotar uma abordagem *online* em vez de *offline*, o que aumentaria a robustez da metodologia em futuros MaOPs, além da família DTLZ.
- Apesar das dúvidas sobre a aplicabilidade da PCA *offline*, o LS pode ser útil para a redução em modelos com dimensões iguais a 2 ou 3, pois cria oportunidades adicionais de visualização para o DM, utilizando ferramentas mais eficientes e aproveitando a simplicidade das dimensões reduzidas. Além disso, este trabalho não apresentou os gráficos DC; essa ausência é justificada pela extensão dos casos tratados. No entanto, esse problema poderia ter sido mitigado se um *threshold* tivesse sido aplicado inicialmente para ambos os métodos, a fim de decidir qual o número final de objetivos, o que reduziria o número excessivo de resultados.
- Por último, uma outra constatação ocorreu durante a implementação em Python do LS para a redução dos problemas de teste DTLZ. Como a ordem das funções é alterada conforme os escores calculados, foi necessário testar a troca de *index* para assegurar a repetibilidade dos experimentos, independentemente da ordem das funções. No entanto, foi observado que, ao inverter completamente a ordem das funções neste vetor para alimentar o algoritmo do problema de teste, a PF final não era a mesma que a da ordem original. Isso não deveria ser verdade, uma vez que as funções são as mesmas, mudando apenas a ordem das colunas na matriz de objetivos. Além disso, conforme descrito em (Blank e Deb, 2020), a função `minimize` cria uma cópia profunda do objeto do algoritmo antes da execução, garantindo que duas execuções independentes com o mesmo algoritmo e a mesma semente aleatória produzam os mesmos resultados, sem efeitos colaterais.

5.1 Trabalhos Futuros

Apesar deste trabalho se tratar do texto final de dissertação, ainda existem uma série de experimentos inexplorados que devem ser conduzidos para o avanço do conhecimento nesta área. Os trabalhos futuros abrangem:

- A utilização/inclusão da ferramenta de visualização DC, desenvolvida pelos autores Koochaksaraei et al. (2017), como descrito na Seção 2.4. Essa abordagem possibilitará a representação visual dos espaços de objetivo de forma clara e interpretável, permitindo que os DMs compreendam melhor as interações entre as funções e as *trade-offs* envolvidas.
- A aplicação dos mesmos métodos utilizados neste trabalho em uma abordagem *online*. Como mencionado em Mendes (2022), a abordagem *online* de redução de dimensionalidade baseada em seleção pode (i) gerar a mesma fronteira não dominada que o problema original ou (ii) preservar as relações de dominância do problema original. Essa aplicação pode contribuir para uma maior flexibilidade e adaptabilidade nas soluções encontradas em tempo real.
- A aplicação de métricas de qualidade específicas para modelos de AM não supervisionado em diferentes cenários de otimização. Isso incluirá a avaliação da eficácia das abordagens de redução de dimensionalidade em relação a métricas padrão da literatura, como, por exemplo, Silhouette Score (Rousseeuw, 1987) e Davies-Bouldin Index (Davies e Bouldin, 1979), possibilitando comparações mais robustas entre métodos e suas respectivas implementações.
- A utilização de métodos de otimização, como o Optuna (biblioteca Python), para o ajuste de hiperparâmetros do LS utilizados em AM, conforme mencionado na Subseção 2.5.1. A automatização do ajuste de hiperparâmetros pode levar a melhorias significativas no desempenho do modelo e aumentar a eficiência do processo do MOEA.
- O desenvolvimento de um algoritmo que calcule o erro associado a diferentes abordagens de redução de dimensionalidade e que encontre o melhor número de objetivos para um determinado MaOP com base nesse cálculo. Essa iniciativa pode ajudar a minimizar o número de cenários a serem analisados e discutidos pelo DM, tornando o processo mais eficiente e direcionado.

- Medir o erro associado à inversa (ou pseudo-inversa) utilizada no PCA durante o processo de reconstrução dos dados. Essa análise permitirá avaliar a precisão da transformação inversa e determinar em que medida a redução de dimensionalidade preserva as características originais do problema.
- A pesquisa e implementação de novos métodos *offline* de AM não supervisionado, conforme apresentados no artigo de revisão de Solorio-Fernández et al. (2020). Esses novos métodos podem trazer *insights* valiosos e expandir a aplicabilidade de técnicas de redução de dimensionalidade em diversos contextos de otimização multiobjetivo.
- Finalmente, a exploração de técnicas de *ensemble* que combinam múltiplos métodos de redução de dimensionalidade. Ao integrar diferentes abordagens, é possível aproveitar as vantagens de cada uma delas, potencializando a qualidade das soluções obtidas e oferecendo uma perspectiva mais abrangente sobre os MaOPs.

Referências Bibliográficas

- Adra, S. F. e Fleming, P. J. (2010). Diversity management in evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 15(2):183–195.
- Bellman, R. (1957). *Dynamic Programming*. Princeton University Press, Princeton, NJ, USA, 1 edition.
- Bentley, J. L., Kung, H.-T., Schkolnick, M. e Thompson, C. D. (1978). On the average number of maxima in a set of vectors and applications. *Journal of the ACM (JACM)*, 25(4):536–543.
- Beyer, K., Goldstein, J., Ramakrishnan, R. e Shaft, U. (1999). When is “nearest neighbor” meaningful? Em *Database Theory—ICDT’99: 7th International Conference Jerusalem, Israel, January 10–12, 1999 Proceedings* 7, páginas 217–235. Springer.
- Bishop, C. M. e Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.
- Blank, J. e Deb, K. (2020). pymoo: Multi-objective optimization in python. *IEEE Access*, 8:89497–89509.
- Blank, J., Deb, K., Dhebar, Y., Bandaru, S. e Seada, H. (2020). Generating well-spaced points on a unit simplex for evolutionary many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 25(1):48–60.
- Blank, J., Deb, K. e Roy, P. C. (2019). Investigating the normalization procedure of nsga-iii. Em *International Conference on Evolutionary Multi-Criterion Optimization*, páginas 229–240. Springer.

- Brockhoff, D. e Zitzler, E. (2006a). Are all objectives necessary? on dimensionality reduction in evolutionary multiobjective optimization. Em *Parallel Problem Solving from Nature-PPSN IX: 9th International Conference, Reykjavik, Iceland, September 9-13, 2006, Proceedings*, páginas 533–542. Springer.
- Brockhoff, D. e Zitzler, E. (2006b). Dimensionality reduction in multiobjective optimization with (partial) dominance structure preservation: Generalized minimum objective subset problems. *TIK Report*, 247.
- Brockhoff, D. e Zitzler, E. (2007). Offline and online objective reduction in evolutionary multiobjective optimization based on objective conflicts. tik report 269. *Computer Engineering and Networks Laboratory (TIK), ETH Zurich*.
- Brockhoff, D. e Zitzler, E. (2009). Objective reduction in evolutionary multiobjective optimization: Theory and applications. *Evolutionary computation*, 17(2):135–166.
- Brockhoff, D. e Zitzler, E. (2010). Automated aggregation and omission of objectives for tackling many-objective problems. Em *New developments in multiple objective and goal programming*, páginas 81–102. Springer.
- Carvalho, J. P. G., Carvalho, E. C. R., Vargas, D. E. C., Hallak, P. H., Lima, B. S. L. P. e Lemonge, A. C. C. (2021). Multi-objective optimum design of truss structures using differential evolution algorithms. *Computers & Structures*, 252:106544.
- Coello, C. A. C. (2007). *Evolutionary algorithms for solving multi-objective problems*. Springer.
- Coello Coello, C. A., González Brambila, S., Figueroa Gamboa, J., Castillo Tapia, M. G. e Hernández Gómez, R. (2020). Evolutionary multiobjective optimization: open research areas and some challenges lying ahead. *Complex & Intelligent Systems*, 6:221–236.
- Coello Coello, C. A. e Reyes Sierra, M. (2004). A study of the parallelization of a coevolutionary multi-objective evolutionary algorithm. Em *MICAI 2004: Advances in Artificial Intelligence: Third Mexican International Conference on Artificial Intelligence, Mexico City, Mexico, April 26-30, 2004. Proceedings 3*, páginas 688–697. Springer.
- Davies, D. L. e Bouldin, D. W. (1979). A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, 2:224–227.

- de Freitas, A. R., Fleming, P. J. e Guimaraes, F. G. (2015). Aggregation trees for visualization and dimension reduction in many-objective optimization. *Information Sciences*, 298:288–314.
- Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, Inc., USA.
- Deb, K. e Jain, H. (2013). An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints. *IEEE transactions on evolutionary computation*, 18(4):577–601.
- Deb, K., Pratap, A., Agarwal, S. e Meyarivan, T. (2002a). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197.
- Deb, K., Saxena, D. et al. (2006). Searching for pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. Em *Proceedings of the world congress on computational intelligence (WCCI-2006)*, páginas 3352–3360.
- Deb, K., Thiele, L., Laumanns, M. e Zitzler, E. (2002b). Scalable multi-objective optimization test problems. Em *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, volume 1, páginas 825–830. IEEE.
- DeRonne, K. W. e Karypis, G. (2013). Pareto optimal pairwise sequence alignment. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10(2):481–493.
- Duro, J. A., Saxena, D. K., Deb, K. e Zhang, Q. (2014). Machine learning based decision support for many-objective optimization problems. *Neurocomputing*, 146:30–47.
- Falcón-Cardona, J. G., Zapotecas-Martínez, S. e García-Nájera, A. (2021). Pareto compliance from a practical point of view. Em *Proceedings of the Genetic and Evolutionary Computation Conference*, páginas 395–402.
- Farina, M. e Amato, P. (2002). On the optimal solution definition for many-criteria optimization problems. Em *2002 annual meeting of the North American fuzzy information processing society proceedings. NAFIPS-FLINT 2002 (Cat. No. 02TH8622)*, páginas 233–238. IEEE.

- Fix, E. e Hodges, J. L. (1989). Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review/Revue Internationale de Statistique*, 57(3):238–247.
- Fleming, P. J., Purshouse, R. C., Lygoe, R. J. et al. (2005). Many-objective optimization: An engineering design perspective. Em *EMO*, volume 5, páginas 14–32. Springer.
- He, X., Cai, D. e Niyogi, P. (2005). Laplacian score for feature selection. *Advances in neural information processing systems*, 18.
- He, Z. e Yen, G. G. (2015). Visualization and performance metric in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 20(3):386–402.
- He, Z., Yen, G. G. e Ding, J. (2020). Knee-based decision making and visualization in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 25(2):292–306.
- Herrero, J. G., Berlanga, A. e López, J. M. M. (2008). Effective evolutionary algorithms for many-specifications attainment: Application to air traffic control tracking filters. *IEEE Transactions on Evolutionary Computation*, 13(1):151–168.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417.
- Huband, S., Barone, L., While, L. e Hingston, P. (2005). A scalable multi-objective test problem toolkit. Em *Evolutionary Multi-Criterion Optimization: Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings 3*, páginas 280–295. Springer.
- Ishibuchi, H., Akedo, N. e Nojima, Y. (2014). Behavior of multiobjective evolutionary algorithms on many-objective knapsack problems. *IEEE Transactions on Evolutionary Computation*, 19(2):264–283.
- Ishibuchi, H., Akedo, N., Ohyanagi, H. e Nojima, Y. (2011). Behavior of emo algorithms on many-objective optimization problems with correlated objectives. Em *2011 IEEE Congress of Evolutionary Computation (CEC)*, páginas 1465–1472. IEEE.
- Ishibuchi, H., Masuda, H. e Nojima, Y. (2016). Sensitivity of performance evaluation results by inverted generational distance to reference points. Em *2016 IEEE Congress on evolutionary computation (CEC)*, páginas 1107–1114. IEEE.

- Ishibuchi, H., Masuda, H., Tanigaki, Y. e Nojima, Y. (2015). Modified distance calculation in generational distance and inverted generational distance. Em *Evolutionary Multi-Criterion Optimization: 8th International Conference, EMO 2015, Guimarães, Portugal, March 29–April 1, 2015. Proceedings, Part II* 8, páginas 110–125. Springer.
- Jain, H. e Deb, K. (2013). An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part ii: Handling constraints and extending to an adaptive approach. *IEEE Transactions on evolutionary computation*, 18(4):602–622.
- Jolliffe, I. T. (2002). *Principal component analysis for special types of data*. Springer.
- Kasprzyk, J. R., Reed, P. M., Characklis, G. W. e Kirsch, B. R. (2012). Many-objective de novo water supply portfolio planning under deep uncertainty. *Environmental Modelling & Software*, 34:87–104.
- Knowles, J. D., Thiele, L. e Zitzler, E. (2006). A tutorial on the performance assessment of stochastic multiobjective optimizers. *TIK-report*, 214.
- Kollat, J. B. e Reed, P. M. (2006). Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design. *Advances in Water Resources*, 29(6):792–807.
- Koochaksaraei, R. H., Meneghini, I. R., Coelho, V. N. e Guimaraes, F. G. (2017). A new visualization method in many-objective optimization with chord diagram and angular mapping. *Knowledge-Based Systems*, 138:134–154.
- Laumanns, M., Thiele, L., Deb, K. e Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary computation*, 10(3):263–282.
- Li, B., Li, J., Tang, K. e Yao, X. (2015). Many-objective evolutionary algorithms: A survey. *ACM Computing Surveys (CSUR)*, 48(1):1–35.
- Li, G. e Zhang, Q. (2021). Multitask feature selection for objective reduction. Em *International Conference on Evolutionary Multi-Criterion Optimization*, páginas 77–88. Springer.
- Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J. e Liu, H. (2017). Feature selection: A data perspective. *ACM computing surveys (CSUR)*, 50(6):1–45.

- Li, M., Yang, S. e Liu, X. (2013). Shift-based density estimation for pareto-based algorithms in many-objective optimization. *IEEE Transactions on Evolutionary Computation*, 18(3):348–365.
- López Jaimes, A., Coello Coello, C. A. e Chakraborty, D. (2008). Objective reduction using a feature selection technique. Em *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, páginas 673–680.
- Mendes, R., Wanner, E., Martins, F. e Sarubbi, J. (2017). Dimensionality reduction approach for many-objective vehicle routing problem with demand responsive transport. Em *Evolutionary Multi-Criterion Optimization: 9th International Conference, EMO 2017, Münster, Germany, March 19-22, 2017, Proceedings 9*, páginas 438–452. Springer.
- Mendes, R. S. (2022). *Redução de dimensionalidade online para problema de roteamento de veículos com transporte reativo a demanda*. Tese de doutorado, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, MG.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2):81.
- Mingoti, S. A. (2007). *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Editora UFMG.
- Mitra, P., Murthy, C. e Pal, S. K. (2002). Unsupervised feature selection using feature similarity. *IEEE transactions on pattern analysis and machine intelligence*, 24(3):301–312.
- Mostaghim, S. e Schmeck, H. (2008). Distance based ranking in many-objective particle swarm optimization. Em *Parallel Problem Solving from Nature–PPSN X: 10th International Conference, Dortmund, Germany, September 13-17, 2008. Proceedings 10*, páginas 753–762. Springer.
- Mühlenbein, H., Schomisch, M. e Born, J. (1991). The parallel genetic algorithm as function optimizer. *Parallel computing*, 17(6-7):619–632.
- Narukawa, K. e Rodemann, T. (2012). Examining the performance of evolutionary many-objective optimization algorithms on a real-world application. Em *2012 Sixth International Conference on Genetic and Evolutionary Computing*, páginas 316–319. IEEE.

- Nguyen, X. H., Bui, L. T. e Tran, C. T. (2020). Improving many objective optimisation algorithms using objective dimensionality reduction. *Evolutionary Intelligence*, 13:365–380.
- Pearson, K. (1901). Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 2(11):559–572.
- Riquelme, N., Von Lüken, C. e Baran, B. (2015). Performance metrics in multi-objective optimization. Em *2015 Latin American computing conference (CLEI)*, páginas 1–11. IEEE.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.
- Saxena, D. K., Duro, J. A., Tiwari, A., Deb, K. e Zhang, Q. (2012). Objective reduction in many-objective optimization: Linear and nonlinear algorithms. *IEEE Transactions on Evolutionary Computation*, 17(1):77–99.
- Sen, P. e Yang, J.-B. (2012). *Multiple criteria decision support in engineering design*. Springer Science & Business Media.
- Singh, H. K., Isaacs, A. e Ray, T. (2011). A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Transactions on Evolutionary Computation*, 15(4):539–556.
- Sinha, A., Saxena, D. K., Deb, K. e Tiwari, A. (2013). Using objective reduction and interactive procedure to handle many-objective optimization problems. *Applied Soft Computing*, 13(1):415–427.
- Solorio-Fernández, S., Carrasco-Ochoa, J. A. e Martínez-Trinidad, J. F. (2020). A review of unsupervised feature selection methods. *Artificial Intelligence Review*, 53(2):907–948.
- Song, H., Triguero, I. e Özcan, E. (2019). A review on the self and dual interactions between machine learning and optimisation. *Progress in Artificial Intelligence*, 8(2):143–165.
- Thiele, L., Miettinen, K., Korhonen, P. J. e Molina, J. (2009). A preference-based evolutionary algorithm for multi-objective optimization. *Evolutionary computation*, 17(3):411–436.
- Van Veldhuizen, D. A. (1999). *Multiobjective evolutionary algorithms: classifications, analyses, and new innovations*. Air Force Institute of Technology.

- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P. e SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- Walpole, R. E., Myers, R. H., Myers, S. L. e Ye, K. (1993). *Probability and statistics for engineers and scientists*, volume 5. Macmillan New York.
- Wang, H. e Yao, X. (2016). Objective reduction based on nonlinear correlation information entropy. *Soft Computing*, 20:2393–2407.
- Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411):664–675.
- Weinberger, K. Q. e Saul, L. K. (2006). Unsupervised learning of image manifolds by semidefinite programming. *International journal of computer vision*, 70:77–90.
- Yao, X., Zhao, Q., Gong, D. e Zhu, S. (2021). Solution of large-scale many-objective optimization problems based on dimension reduction and solving knowledge guided evolutionary algorithm. *IEEE Transactions on Evolutionary Computation*.
- Zhang, Q. e Li, H. (2007). Moea/d: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on evolutionary computation*, 11(6):712–731.
- Zhou, C., Zheng, J., Li, K. e Lv, H. (2009). Objective reduction based on the least square method for large-dimensional multi-objective optimization problem. Em *2009 Fifth International Conference on Natural Computation*, volume 4, páginas 350–354. IEEE.
- Zitzler, E. e Künzli, S. (2004). Indicator-based selection in multiobjective search. Em *International conference on parallel problem solving from nature*, páginas 832–842. Springer.
- Zitzler, E., Laumanns, M. e Thiele, L. (2001). Spea2: Improving the strength pareto evolutionary algorithm. *TIK-report*, 103.
- Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M. e Da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132.

Zou, X., Chen, Y., Liu, M. e Kang, L. (2008). A new evolutionary algorithm for solving many-objective optimization problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(5):1402–1412.

Resultados Detalhados da Seleção de Atributos: LS

Este apêndice complementa a Seção 4.2 do Capítulo 4, apresentando os resultados detalhados dos experimentos computacionais realizados com o método de Laplacian Score (LS) para redução de objetivos em MaOPs. As tabelas e gráficos aqui contidos correspondem às análises das funções de *benchmark* DTLZ1 a DTLZ7, fornecendo uma visão completa dos resultados obtidos para cada problema de teste.

A.1 DTLZ1

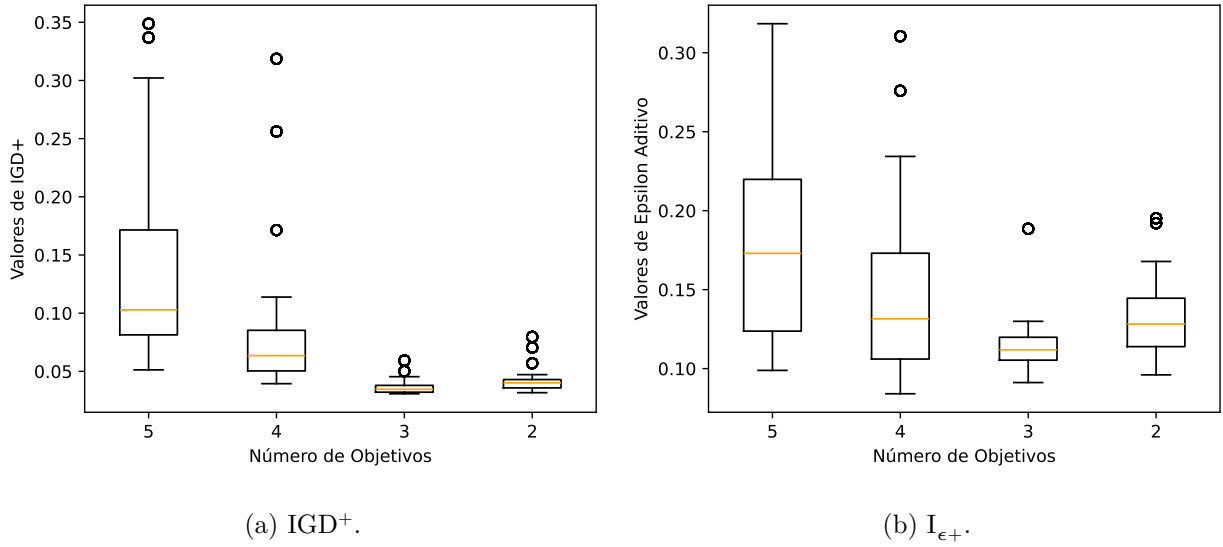
O problema DTLZ1, definido pela Equação 2.3, é utilizado para avaliar a capacidade de um MOEA em convergir para um hiperplano. Nesta seção, são apresentados os resultados da aplicação do LS, incluindo escores das funções objetivo, IQs, e respectivos *boxplots*. Os resultados mostram um padrão em que as últimas funções objetivo tendem a apresentar escores mais baixos, sugerindo maior relevância.

Tabela A.1: Escores: DTLZ1 com $m = 5$.

f_5	f_4	f_3	f_2	f_1
0.0137	0.0211	0.0383	0.0441	0.0657

Tabela A.2: Resultados da seleção de objetivos: DTLZ1 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.1431	0.0896	0.0371	0.0432
p-valor	1	0.0018	< 0.001	< 0.001
I _{ϵ+}	0.1809	0.1483	0.1157	0.1336
p-valor	1	0.0761	< 0.001	0.0259

Figura A.1: *Boxplots* dos IQs: DTLZ1 por seleção com $m = 5$.

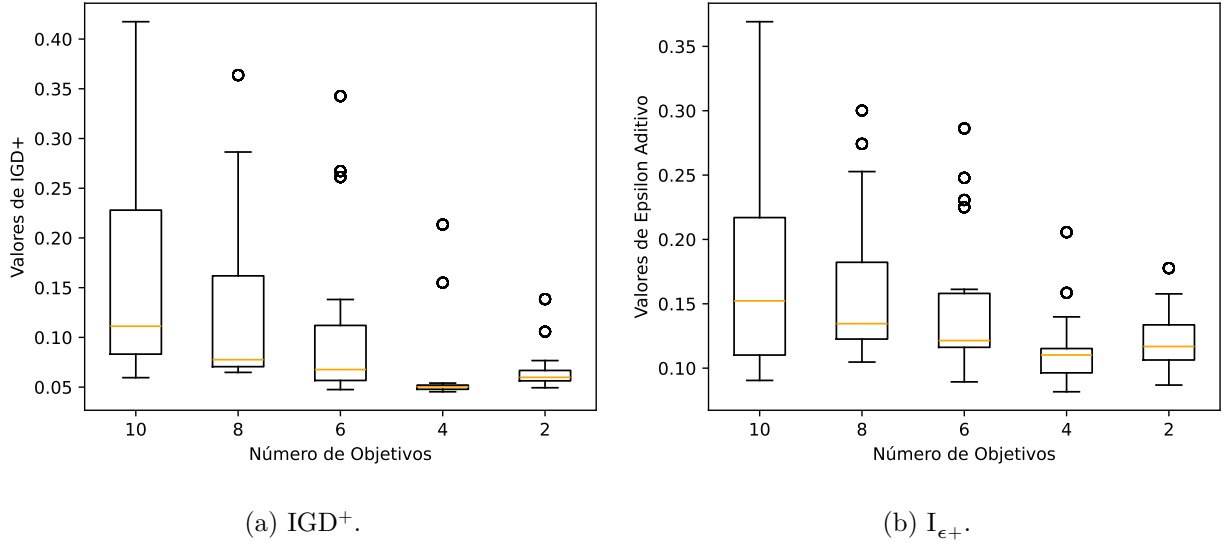
Analisando os escores na Tabela A.1, nota-se que eles aumentaram à medida que o índice da função decresceu. Conforme a Tabela A.2, todos os resultados apresentaram uma melhora significativa nos IQs para a seleção de funções, especialmente na abordagem triobjetivo. Os *boxplots* da Figura A.1 ilustram esta tendência. Além disso, o p-valor < 0.05 fornece evidências estatísticas para a rejeição da hipótese nula, indicando um ganho de desempenho em relação aos valores de IGD⁺ e I _{ϵ +} observados no MaOP original, após a aplicação da seleção de funções.

Tabela A.3: Escores: DTLZ1 com $m = 10$.

f_{10}	f_9	f_8	f_7	f_6	f_4	f_5	f_3	f_1	f_2
0.0252	0.0338	0.0472	0.0691	0.0946	0.1247	0.2054	0.3142	0.344	0.6131

Tabela A.4: Resultados da seleção de objetivos: DTLZ1 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	0.1597	0.1330	0.1043	0.0625	0.0659
p-valor	1	0.1951	0.0079	< 0.001	< 0.001
$I_{\epsilon+}$	0.1706	0.1625	0.1478	0.1128	0.1219
p-valor	1	0.9498	0.4136	0.0024	0.0152

Figura A.2: *Boxplots* dos IQs: DTLZ1 por seleção com $m = 10$.

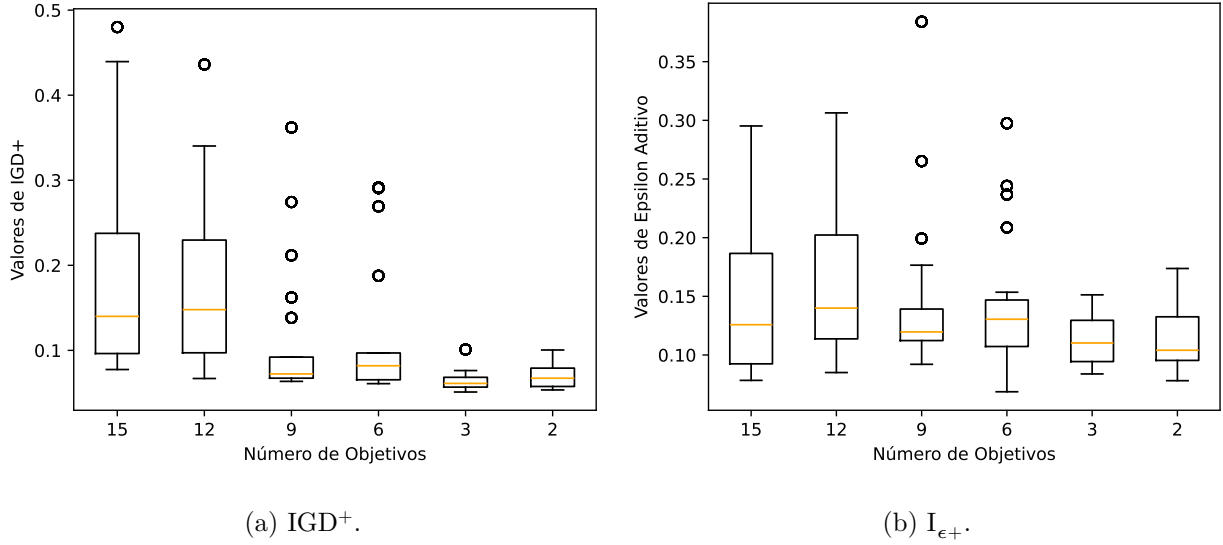
A Tabela A.3 apresenta os escores para este MaOP com $m = 10$, sendo possível notar que, embora haja uma tendência geral de índices das funções em ordem decrescente, esse critério não é seguido à risca. Similarmente ao cenário anterior, a seleção de funções apresentou efeitos positivos, de acordo com a Tabela A.4. Destacam-se os resultados para $m = 4$ e $m = 2$, onde o p-valor foi menor que 5% para ambas as médias dos IQs, porém o primeiro sugere um ganho de desempenho do MOEA ligeiramente melhor. A respeito da Figura A.2, os *boxplots* representam a queda considerável nos valores de IQ para os resultados já destacados, além de uma alta ocorrência de *outliers* para $m = 6$.

Tabela A.5: Escores: DTLZ1 com $m = 15$.

f_{15}	f_{14}	f_{13}	f_{12}	f_9	f_{11}	f_{10}	f_7	f_3	f_6
0.0253	0.0321	0.0516	0.1026	0.1849	0.2128	0.2174	0.2372	0.2395	0.2456
f_8	f_5	f_2	f_4	f_1					
0.2697	0.3763	0.4074	0.421	0.4474					

Tabela A.6: Resultados da seleção de objetivos: DTLZ1 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.1989	0.1769	0.1103	0.1122	0.0638	0.0714
p-valor	1	0.5886	< 0.001	0.0015	< 0.001	< 0.001
$I_{\epsilon+}$	0.1482	0.1673	0.1436	0.1428	0.1119	0.1155
p-valor	1	0.2965	0.8899	0.8899	0.2039	0.3327

Figura A.3: *Boxplots* dos IQs: DTLZ1 por seleção com $m = 15$.

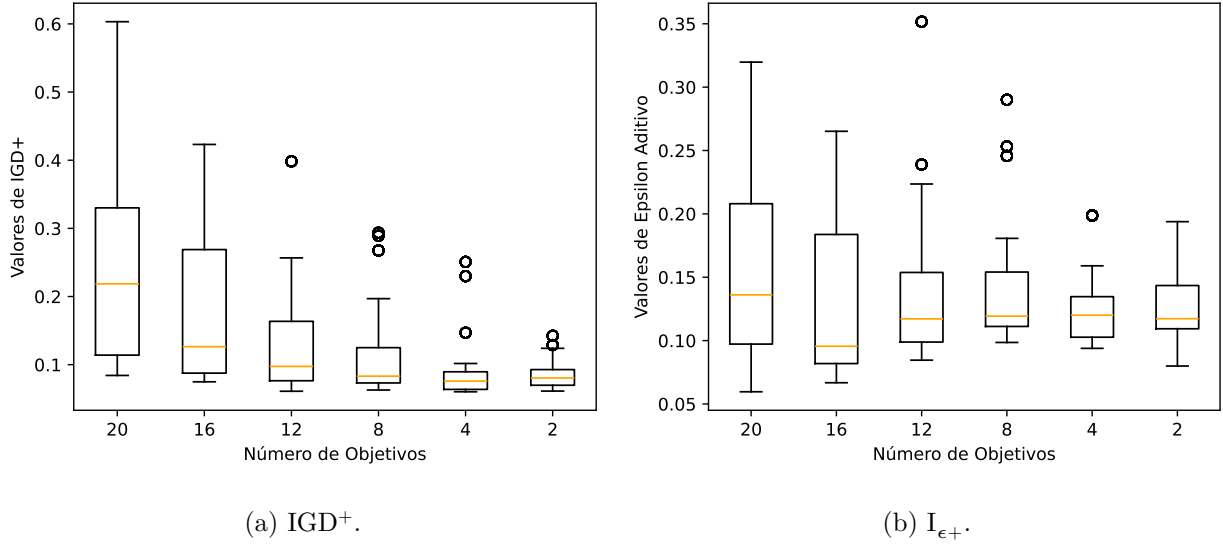
No cenário com $m = 15$, conforme a Tabela A.5, foi obtido um *ranking* similar ao dos cenários anteriores. Quanto à seleção, houve uma redução nas grandezas dos IQs em relação ao original, em especial para $m = 3$ e $m = 2$, conforme a Tabela A.6. A Figura A.3 mostra que estas duas abordagens obtiveram resultados mais consistentes que as demais, acarretando no valor mais baixo para suas médias. Porém, em ambos os casos, os testes de hipóteses não fornecem evidências estatísticas para rejeitar a hipótese nula do $I_{\epsilon+}$, visto que o p-valor foi > 0.05 .

Tabela A.7: Escores: DTLZ1 com $m = 20$.

f_{20}	f_{19}	f_{17}	f_{18}	f_{15}	f_{16}	f_{14}	f_{11}	f_{10}	f_{12}
0.0227	0.0286	0.0612	0.0717	0.125	0.1706	0.2579	0.2849	0.2982	0.3334
f_{13}	f_3	f_8	f_5	f_2	f_9	f_1	f_7	f_6	f_4
0.3425	0.376	0.4386	0.4405	0.4452	0.4514	0.5052	0.5228	0.5591	0.6319

Tabela A.8: Resultados da seleção de objetivos: DTLZ1 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	0.2527	0.1824	0.1350	0.1198	0.0935	0.0857
p-valor	1	0.0644	0.0024	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.1608	0.1357	0.1399	0.1430	0.1248	0.1275
p-valor	1	0.2421	0.5212	0.6597	0.2421	0.2523

Figura A.4: *Boxplots* dos IQs: DTLZ1 por seleção com $m = 20$.

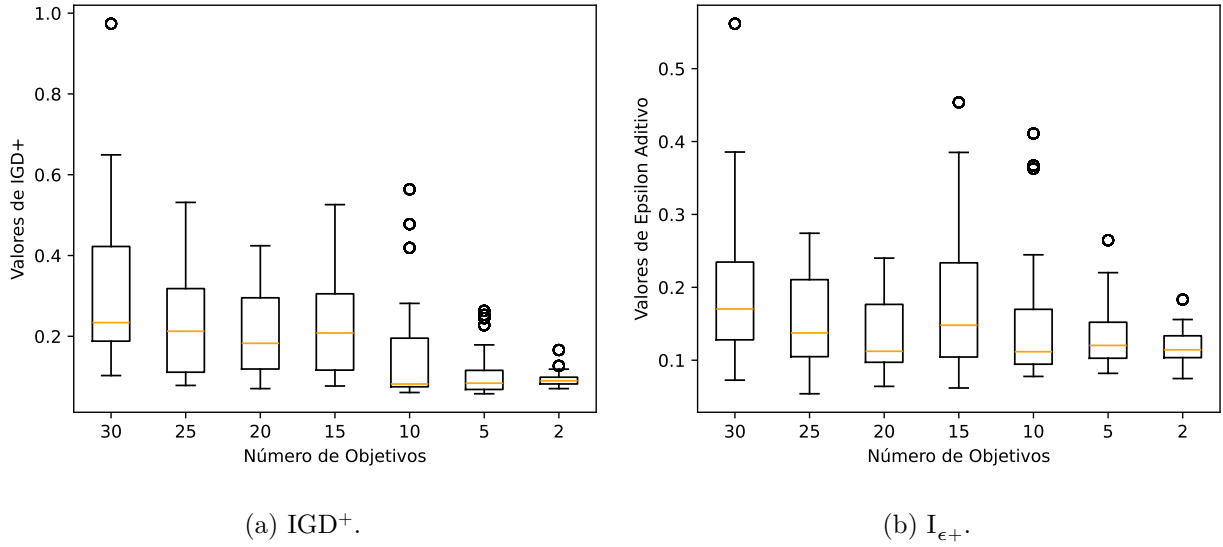
Para o cenário com $m = 20$, na Tabela A.7 verifica-se o mesmo padrão nos escores dos MaOPs anteriores. Já para os IQs da seleção na Tabela A.8, nota-se uma redução progressiva nos valores dos IQs e os *boxplots* da Figura A.4 ilustram este fenômeno, bem como as dispersões semelhantes para o ϵ -indicador. Similarmente ao cenário anterior, o $I_{\epsilon+}$ não alcançou um p-valor < 0.05 e sugere a falta de evidências estatísticas, indicando a ausência de diferença significativa entre as amostras do IQ.

Tabela A.9: Escores: DTLZ1 com $m = 30$.

f_{30}	f_{29}	f_{28}	f_{27}	f_9	f_{25}	f_{23}	f_{26}	f_1	f_4
0.0287	0.0319	0.061	0.0817	0.1548	0.1763	0.1841	0.2097	0.2214	0.2398
f_{22}	f_{11}	f_{24}	f_8	f_{14}	f_{19}	f_{21}	f_{20}	f_5	f_{16}
0.2401	0.2573	0.2964	0.3161	0.3265	0.3831	0.3844	0.3945	0.412	0.4158
f_{17}	f_6	f_7	f_3	f_{15}	f_{10}	f_{13}	f_{18}	f_{12}	f_2
0.4158	0.4257	0.4351	0.4418	0.4506	0.4604	0.4612	0.4885	0.5106	0.592

Tabela A.10: Resultados da seleção de objetivos: DTLZ1 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	0.3249	0.2329	0.2044	0.2403	0.1640	0.1161	0.0955
p-valor	1	0.1411	0.0356	0.1783	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.1971	0.1512	0.1346	0.1846	0.1562	0.1348	0.1198
p-valor	1	0.1480	0.0277	0.5376	0.0543	0.0132	0.0010

Figura A.5: *Boxplots* dos IQs: DTLZ1 por seleção com $m = 30$.

No último MaOP considerado, $m = 30$, é notável o mesmo comportamento dos escores perante os outros cenários, conforme a Tabela A.9. Em contrapartida, como pode ser visto na Tabela A.10, a seleção demonstrou uma redução nos valores dos IQs, onde os problemas $m = 5$ e $m = 2$ apresentaram o p-valor < 0.05 , em especial o resultado deste último, indicando um ganho de desempenho do MOEA neste caso. Os *boxplots* representados na Figura A.5 demonstram esta vantagem, uma vez que as dispersões destes são menores que as dos demais.

A.2 DTLZ2

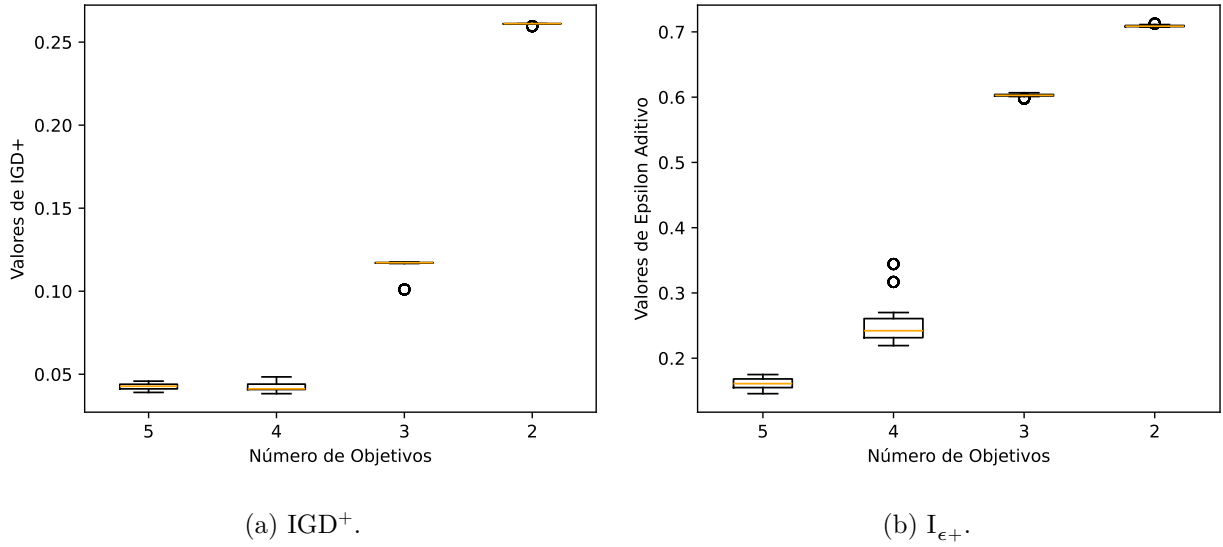
O problema DTLZ2, conforme a Equação 2.5, é projetado para testar a capacidade de um MOEA em lidar com um grande número de objetivos. Os resultados seguem um padrão semelhante ao observado no DTLZ1, com as últimas funções objetivo apresentando maior relevância.

Tabela A.11: Escores: DTLZ2 com $m = 5$.

f_5	f_4	f_3	f_1	f_2
0.0264	0.0265	0.031	0.0334	0.043

Tabela A.12: Resultados da seleção de objetivos: DTLZ2 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.0425	0.0424	0.1163	0.2610
p-valor	1	0.5212	< 0.001	< 0.001
I _{ϵ+}	0.1612	0.2512	0.6027	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001

Figura A.6: *Boxplots* dos IQs: DTLZ2 por seleção com $m = 5$.

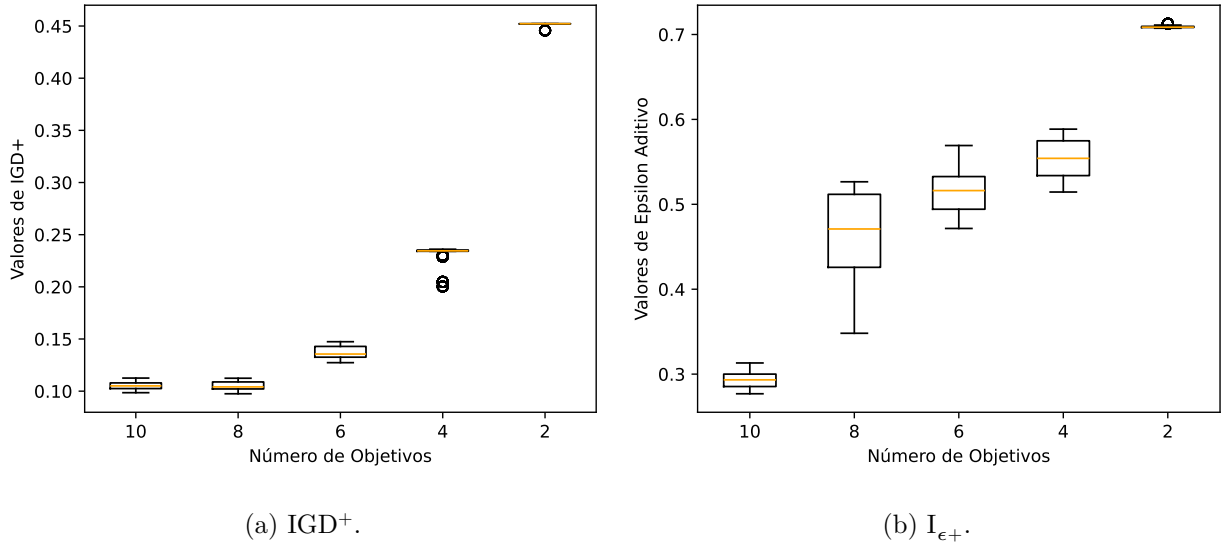
Primeiramente, analisando a Tabela A.11, nota-se que os escores alcançaram um padrão de *ranking* similar ao problema teste anterior, onde a ordem das funções objetivo é invertida. Visitando a Tabela A.12, os resultados evidenciam que houve uma piora gradual em ambos os IQs, exceto para o caso $m = 4$ (que agrega f_1 e f_2), onde o IGD⁺ demonstra evidências estatísticas de ser igual ao problema original. Estas observações também estão representadas na Figura A.6, visto que as medianas crescem gradualmente, exceto para o IGD⁺ em $m = 4$. Notam-se também baixas dispersões para todos os casos.

Tabela A.13: Escores: DTLZ2 com $m = 10$.

f_{10}	f_9	f_8	f_7	f_6	f_5	f_4	f_2	f_3	f_1
0.0361	0.0391	0.046	0.0595	0.0911	0.1276	0.2023	0.3052	0.3064	0.3066

Tabela A.14: Resultados da seleção de objetivos: DTLZ2 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	0.1051	0.1050	0.1374	0.2312	0.4518
p-valor	1	0.9899	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.2940	0.4643	0.5145	0.5521	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.7: *Boxplots* dos IQs: DTLZ2 por seleção com $m = 10$.

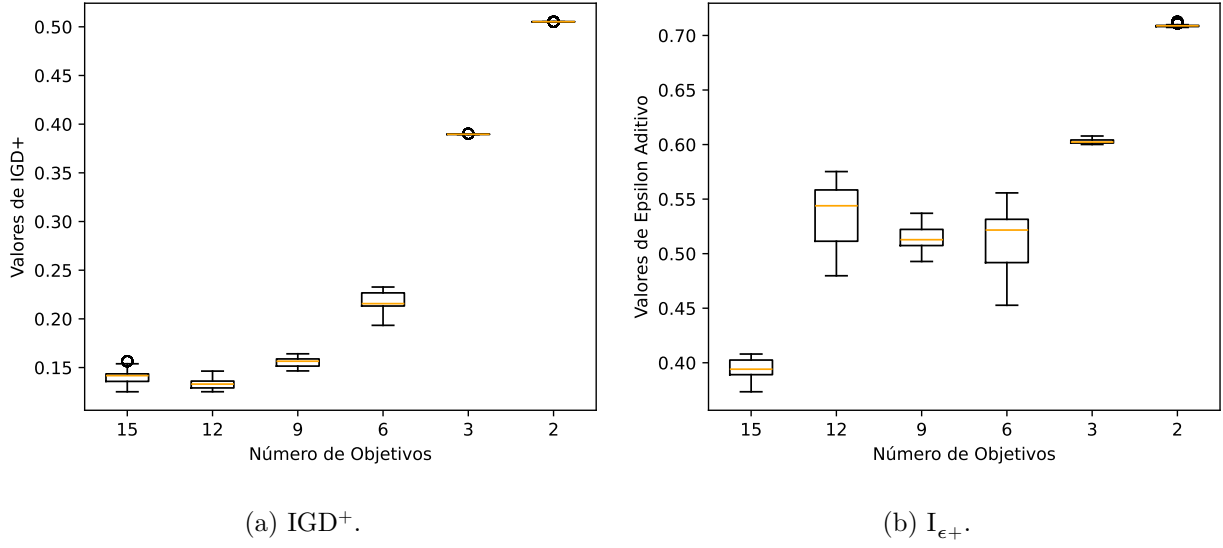
Observando os escores apresentados na Tabela A.13, percebe-se o mesmo comportamento deste *ranking* com o dos casos anteriores. Ao analisar a Tabela A.14, nota-se que no cenário $m = 8$ o valor de IGD⁺ foi estatisticamente igual ao problema original, de acordo com o p-valor. No entanto, quando os cenários começaram a incluir as funções com escores menores, observa-se uma piora significativa nos IQs. Isto também pode ser analisado nos *boxplots* na Figura A.7, que mais uma vez apresentam a baixa dispersão das médias das execuções do NSGA-III, principalmente no cálculo do IGD⁺.

Tabela A.15: Escores: DTLZ2 com $m = 15$.

f_{14}	f_{15}	f_{13}	f_{12}	f_{11}	f_{10}	f_9	f_6	f_5	f_8
0.0374	0.0421	0.0497	0.0714	0.1034	0.1492	0.1585	0.2848	0.3292	0.3466
f_7	f_4	f_2	f_1	f_3					
0.3476	0.3823	0.409	0.4429	0.5005					

Tabela A.16: Resultados da seleção de objetivos: DTLZ2 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.1410	0.1338	0.1554	0.2176	0.3896	0.5052
p-valor	1	0.0014	< 0.001	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3945	0.5345	0.5142	0.5157	0.6027	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.8: *Boxplots* dos IQs: DTLZ2 por seleção com $m = 15$.

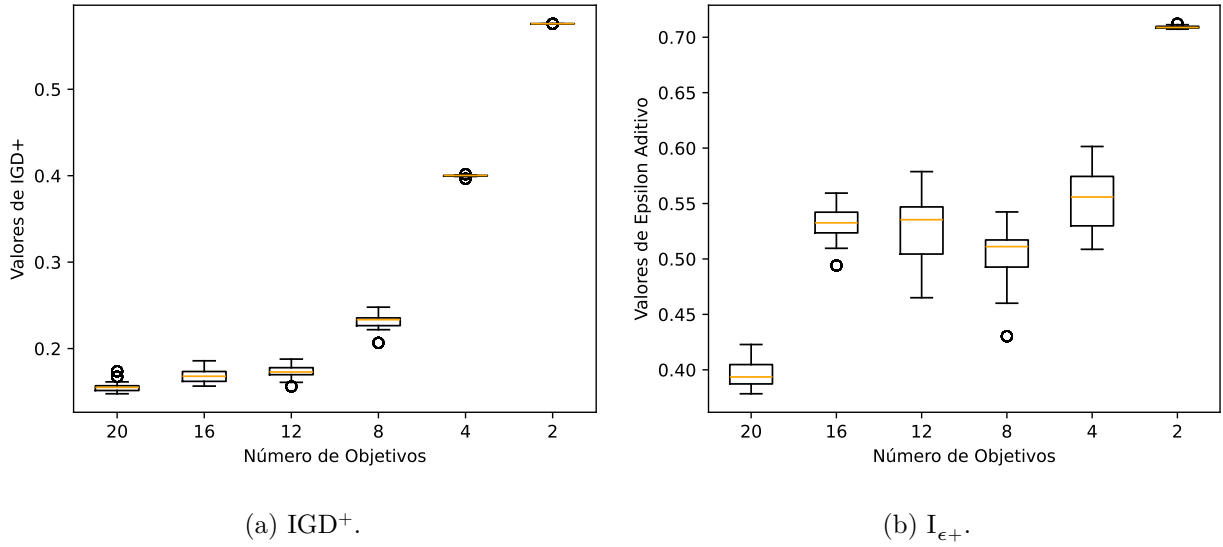
Os resultados do cenário $m = 15$ apresentados nas Tabelas A.15 e A.16 comportaram-se de maneira similar aos MaOPs anteriores. Portanto, à medida que a redução se agrava, observa-se uma piora evidente nos IQs. Destaque para o cenário $m = 12$ que mostrou seu valor de IGD⁺ consideravelmente próximo do inicial, assim como as reduções imediatas dos originais nos cenários $m = 5$ e $m = 10$, isto é, $m = 4$ e $m = 8$, respectivamente. Estes padrões também podem ser notados na Figura A.8, da mesma maneira que as dispersões menos expressivas nos *boxplots* do IGD⁺.

Tabela A.17: Escores: DTLZ2 com $m = 20$.

f_{19}	f_{20}	f_{18}	f_{17}	f_{16}	f_{15}	f_{13}	f_{14}	f_{12}	f_{11}
0.0375	0.0403	0.0589	0.0865	0.0947	0.1664	0.2039	0.2524	0.3021	0.3423
f_{10}	f_7	f_8	f_9	f_6	f_1	f_3	f_4	f_5	f_2
0.3798	0.3901	0.3991	0.4004	0.4026	0.4066	0.4194	0.442	0.4467	0.4807

Tabela A.18: Resultados da seleção de objetivos: DTLZ2 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	0.1553	0.1687	0.1732	0.2312	0.3999	0.5758
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3962	0.5313	0.5259	0.5040	0.5518	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.9: *Boxplots* dos IQs: DTLZ2 por seleção com $m = 20$.

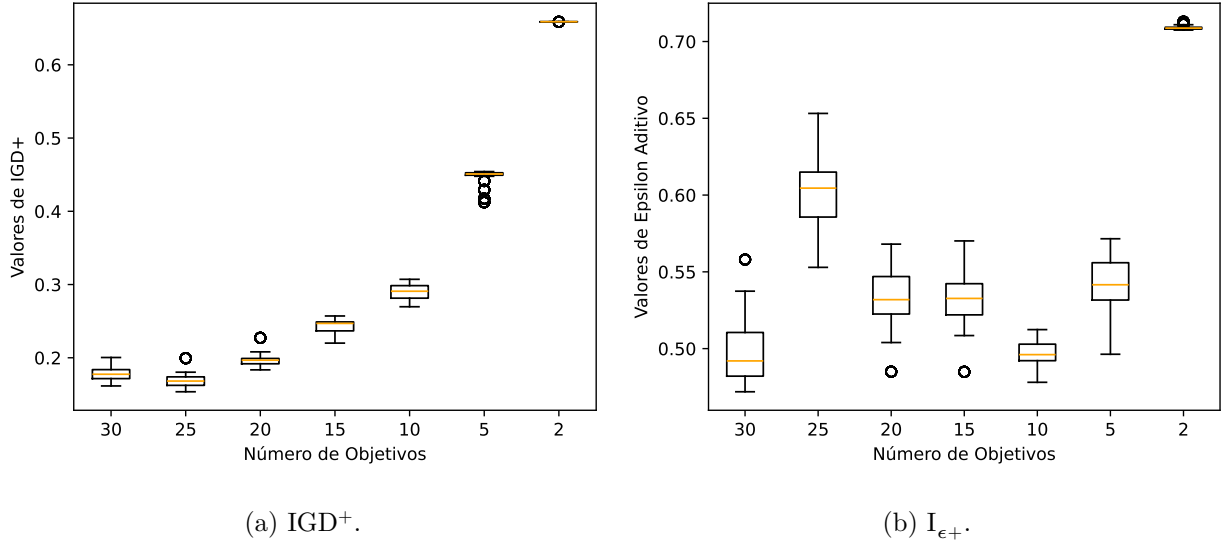
Da mesma forma, os resultados do cenário $m = 20$, apresentados nas Tabelas A.17 e A.18 mantiveram-se com comportamentos similares aos MaOPs anteriores, bem como os *boxplots* da Figura A.9. Sendo assim, à medida que a redução tornou-se mais expressiva, observa-se uma piora evidente em ambos os IQs. Além disso, a primeira redução ($m = 16$) também mostrou seu valor de IGD⁺ significativamente próximo ao problema inicial.

Tabela A.19: Escores: DTLZ2 com $m = 30$.

f_{30}	f_{29}	f_{28}	f_{27}	f_{26}	f_{25}	f_{24}	f_{23}	f_{21}	f_{22}
0.0364	0.0391	0.0513	0.0782	0.0919	0.1791	0.2035	0.2823	0.3083	0.329
f_{17}	f_9	f_1	f_{13}	f_{16}	f_{10}	f_7	f_{20}	f_{12}	f_{18}
0.3669	0.3694	0.3882	0.3884	0.3949	0.3983	0.4095	0.4215	0.4281	0.4308
f_{14}	f_3	f_8	f_6	f_2	f_{11}	f_{15}	f_{19}	f_5	f_4
0.4319	0.4333	0.4398	0.4472	0.4475	0.4693	0.4756	0.4792	0.4881	0.4927

Tabela A.20: Resultados da seleção de objetivos: DTLZ2 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	0.1783	0.1693	0.1971	0.2430	0.2899	0.4464	0.6588
p-valor	1	0.0114	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.4976	0.6017	0.5321	0.5333	0.4961	0.5425	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001	0.6780	< 0.001	< 0.001

Figura A.10: *Boxplots* dos IQs: DTLZ2 por seleção com $m = 30$.

Os resultados do cenário $m = 30$ nas Tabelas A.19 e A.20 demonstraram um padrão consistente com todos os MaOPs anteriores. Observa-se à medida que a redução se intensifica, segue a mesma deterioração significativa nos IQs. Novamente, percebe-se o padrão onde a redução mais próxima do original, neste caso $m = 25$, alcançou o valor do IGD⁺ estatisticamente próximo ao IQ do problema inicial. Estas análises também estão representadas na Figura A.10.

A.3 DTLZ3

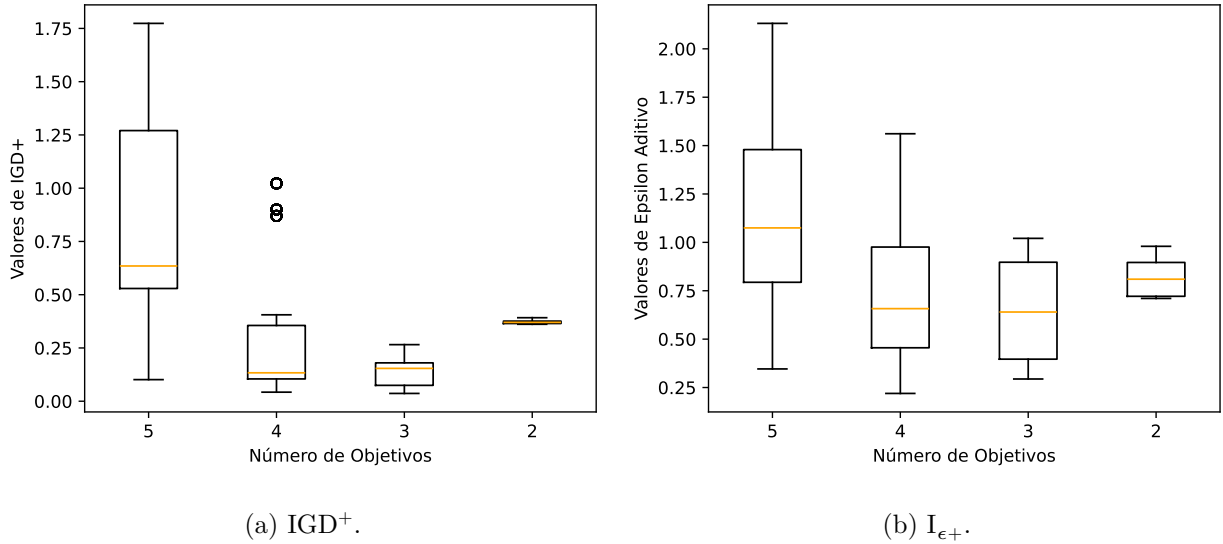
O problema DTLZ3, definido pela Equação 2.6, compartilha as mesmas funções objetivo do DTLZ2, mas com um funcional modificado para testar a convergência para a PF global. Os resultados mostram um padrão consistente com os observados nos problemas DTLZ1 e DTLZ2.

Tabela A.21: Escores: DTLZ3 com $m = 5$.

f_5	f_4	f_3	f_1	f_2
0.0265	0.0307	0.0396	0.0531	0.062

Tabela A.22: Resultados da seleção de objetivos: DTLZ3 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.8057	0.2767	0.1454	0.3720
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	1.1725	0.7301	0.6533	0.8137
p-valor	1	0.0054	< 0.001	0.0141

Figura A.11: *Boxplots* dos IQs: DTLZ3 por seleção com $m = 5$.

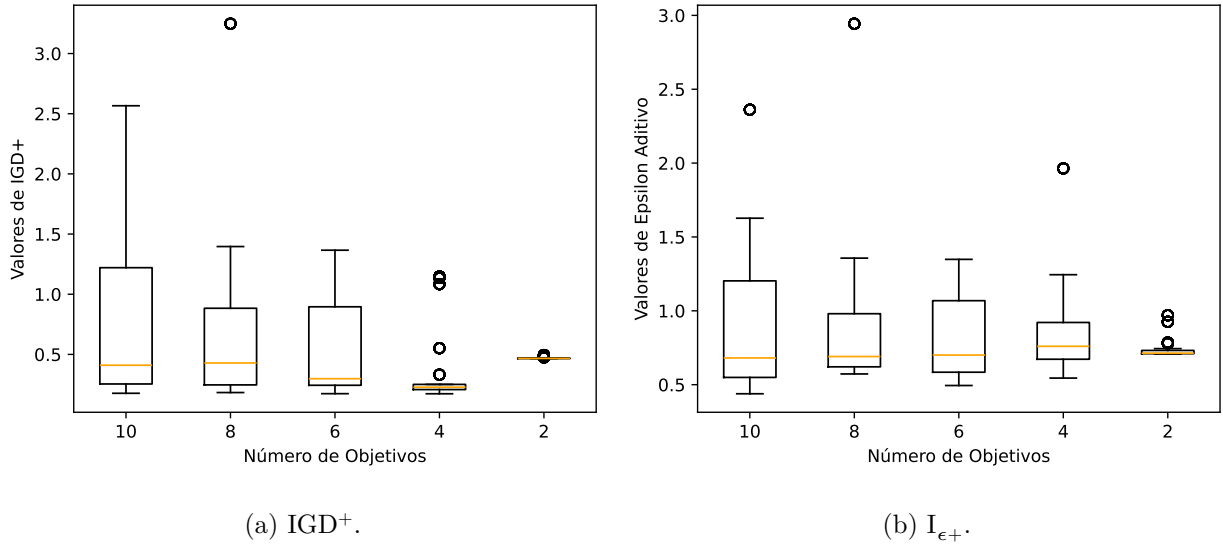
Para o cenário $m = 5$, observando a Tabela A.21, nota-se que os escores alcançaram um padrão de *ranking* idêntico ao problema teste anterior. Como pode ser visto na Tabela A.22, a abordagem com $m = 3$ destaca-se como uma ótima candidata para ganho de desempenho do MOEA, visto que demonstrou reduções expressivas em ambos os IQs e os p-valores < 5% sugerem evidências estatísticas para rejeição da hipótese nula. Os *boxplots* da Figura A.11 ilustram as melhorias observadas nas medianas dos IQs, além de evidenciar a redução significativa da dispersão, especialmente para o cenário $m = 2$.

Tabela A.23: Escores: DTLZ3 com $m = 10$.

f_{10}	f_9	f_8	f_6	f_7	f_5	f_3	f_2	f_1	f_4
0.043	0.0435	0.0605	0.0967	0.1016	0.1417	0.226	0.2646	0.2807	0.2927

Tabela A.24: Resultados da seleção de objetivos: DTLZ3 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	0.7500	0.6693	0.5038	0.3669	0.4699
p-valor	1	0.7916	0.2224	0.0031	0.1866
$I_{\epsilon+}$	0.8916	0.9040	0.7923	0.8526	0.7446
p-valor	1	0.4580	0.8701	0.3327	0.1703

Figura A.12: *Boxplots* dos IQs: DTLZ3 por seleção com $m = 10$.

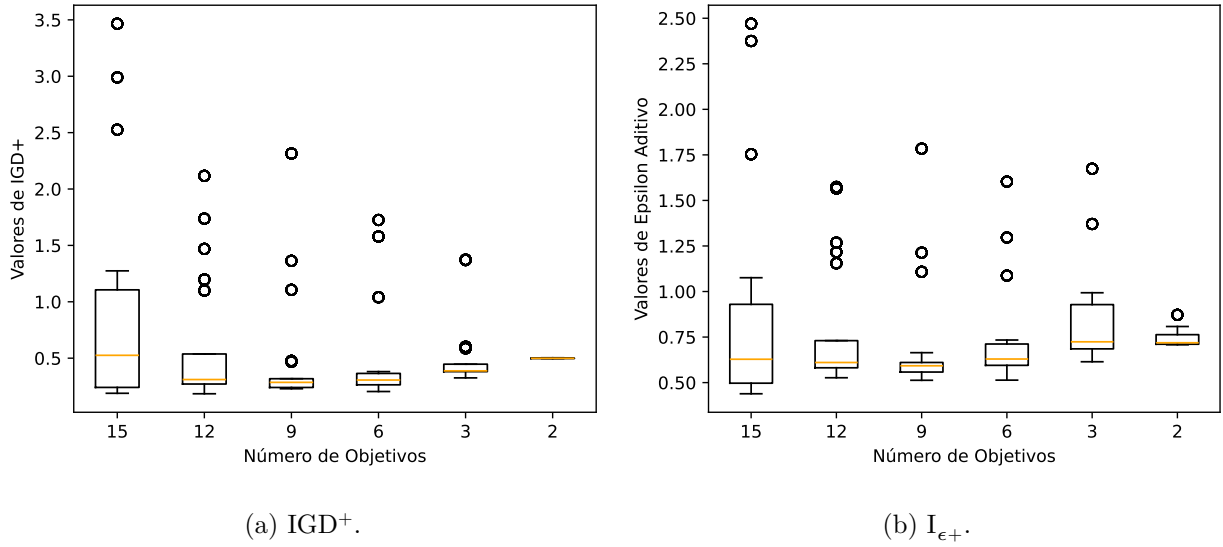
No MaOP com $m = 10$, segundo a Tabela A.23, outra vez os escores calculados mostram que os objetivos mais representativos praticamente invertem a ordem dos índices das funções. Conforme pode ser visto na Tabela A.24, a abordagem com $m = 4$ demonstrou evidências estatísticas para rejeitar a hipótese nula do resultado de IGD⁺, mas o mesmo não é válido para o $I_{\epsilon+}$. Analisando a Figura A.12, de fato os *boxplots* estão similares entre si, sendo $m = 4$ com mediana mais baixa, porém com mais *outliers*. Outra observação é a dispersão muito baixa em $m = 2$, mas desta vez com a presença de *outliers*.

Tabela A.25: Escores: DTLZ3 com $m = 15$.

f_{15}	f_{14}	f_{13}	f_{12}	f_{10}	f_{11}	f_9	f_7	f_8	f_5
0.0444	0.0469	0.0855	0.0882	0.1385	0.1472	0.2044	0.2339	0.2363	0.254
f_4	f_6	f_1	f_2	f_3					
0.4034	0.4089	0.458	0.498	0.6063					

Tabela A.26: Resultados da seleção de objetivos: DTLZ3 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.8825	0.6126	0.4765	0.4601	0.5085	0.4988
p-valor	1	0.5376	0.2965	0.3203	0.8899	0.7916
I _{ϵ+}	0.8699	0.7862	0.6912	0.7250	0.8361	0.7386
p-valor	1	0.5886	0.7531	0.6597	0.0720	0.1158

Figura A.13: *Boxplots* dos IQs: DTLZ3 por seleção com $m = 15$.

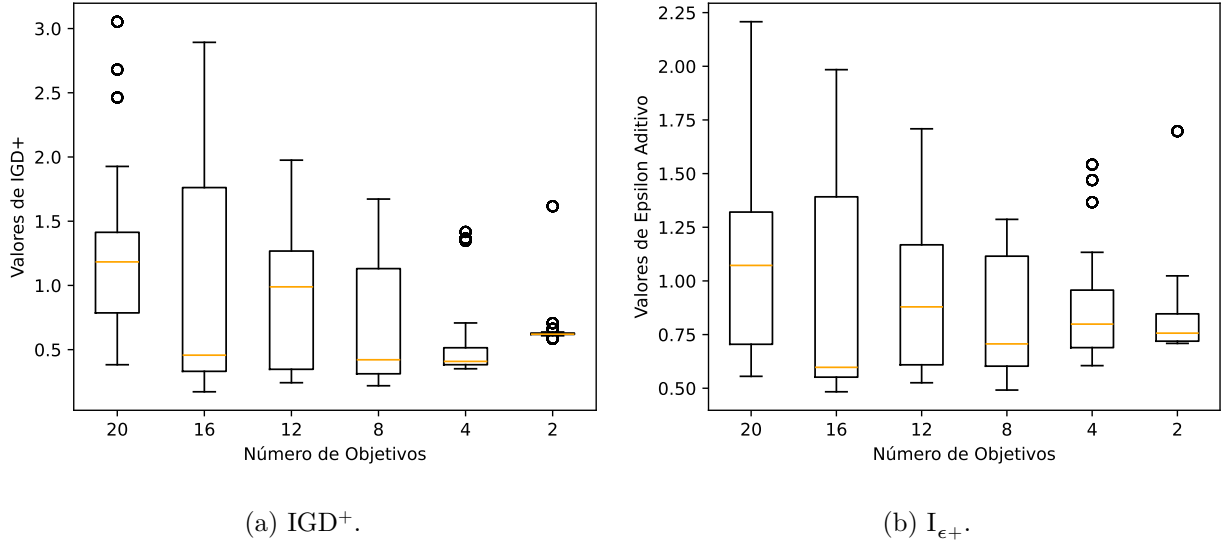
No cenário $m = 15$, os escores da Tabela A.25 mantiveram o padrão observado nos MaOPs anteriores. A Tabela A.26 evidencia melhorias nos resultados dos IQs, entre elas destaca-se a redução para $m = 6$, porém nenhum dos casos alcançou um p-valor $< 5\%$. Pelos *boxplots* da Figura A.13 verifica-se novamente as medianas similares e, apesar da baixa dispersão em todas as reduções, há a frequência de mais *outliers* que no cenário anterior. Exceto para dimensão $m = 2$, que é praticamente inexistente a dispersão nos IQs.

Tabela A.27: Escores: DTLZ3 com $m = 20$.

f_{19}	f_{20}	f_{18}	f_{17}	f_{16}	f_{13}	f_{15}	f_{14}	f_{12}	f_{10}
0.049	0.0509	0.0644	0.0779	0.1505	0.1806	0.211	0.2117	0.2711	0.3425
f_{11}	f_9	f_7	f_8	f_4	f_2	f_5	f_3	f_1	f_6
0.3874	0.4325	0.444	0.4848	0.4983	0.5039	0.5047	0.5138	0.5221	0.5507

Tabela A.28: Resultados da seleção de objetivos: DTLZ3 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	1.2776	0.9999	0.9312	0.6751	0.6019	0.6713
p-valor	1	0.0681	0.0895	0.0024	< 0.001	0.0015
I _{ϵ+}	1.1158	0.9523	0.9616	0.8132	0.8796	0.8404
p-valor	1	0.0720	0.2965	0.0403	0.2523	0.2039

Figura A.14: *Boxplots* dos IQs: DTLZ3 por seleção com $m = 20$.

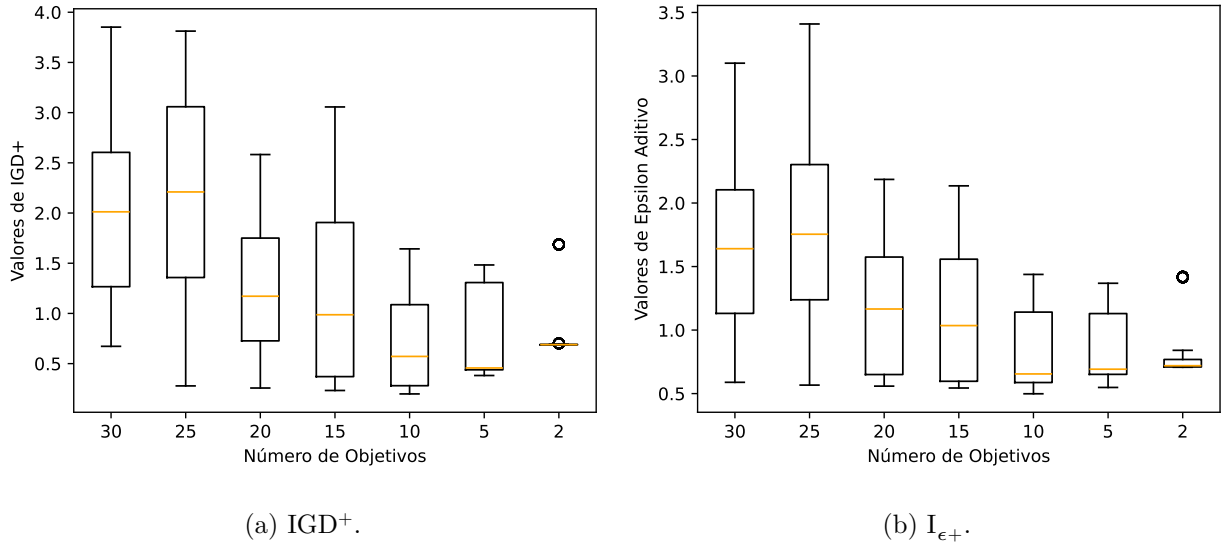
Analisando a Tabela A.27 do DTLZ3 com $m = 20$, os escores comportaram-se da mesma maneira percebida nos cenários anteriores. A respeito dos resultados da Tabela A.28, diferentemente dos dois MaOPs anteriores, a seleção de atributos desempenhou ganhos nos IQs, onde os p-valores da redução para $m = 8$ sugerem evidências estatísticas para rejeição da hipótese nula. Visitando a Figura A.14, percebe-se a redução nas medianas de $m = 15$ e $m = 8$, porém o primeiro apresentou uma dispersão mais elevada que o segundo, influenciando no valor da média. Ademais, as reduções $m = 4$ e $m = 2$ possuem baixa dispersão, mas os *outliers* definiram os resultados negativos nas análises estatísticas.

Tabela A.29: Escores: DTLZ3 com $m = 30$.

f_{30}	f_{29}	f_{28}	f_{27}	f_{26}	f_{25}	f_{22}	f_{19}	f_{20}	f_{13}
0.0501	0.0514	0.0613	0.0858	0.1249	0.1833	0.2458	0.269	0.2702	0.2904
f_{24}	f_1	f_{17}	f_{12}	f_{16}	f_{23}	f_{18}	f_9	f_{21}	f_{14}
0.3245	0.3508	0.3764	0.3879	0.4174	0.4186	0.4262	0.4282	0.4411	0.4467
f_{15}	f_7	f_8	f_{11}	f_6	f_2	f_3	f_4	f_{10}	f_5
0.4574	0.4583	0.4682	0.4819	0.4961	0.4973	0.5238	0.535	0.5511	0.557

Tabela A.30: Resultados da seleção de objetivos: DTLZ3 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	2.0352	2.1418	1.2430	1.2023	0.7370	0.7331	0.7846
p-valor	1	0.6780	0.0043	0.0054	< 0.001	< 0.001	< 0.001
I _{ϵ+}	1.6348	1.8052	1.1820	1.1183	0.8499	0.8275	0.8015
p-valor	1	0.4890	0.0243	0.0079	< 0.001	< 0.001	< 0.001

Figura A.15: *Boxplots* dos IQs: DTLZ3 por seleção com $m = 30$.

Por último, no MaOP com $m = 30$, os escores da Tabela A.29 mantiveram o padrão percebido nos demais cenários do DTLZ3. Analisando os resultados da Tabela A.30, notam-se ganhos de desempenho nos IQs, corroborados com os p-valores abaixo de 5%, onde as reduções para $m = 10$, $m = 5$ e $m = 2$ destacaram-se entre as alternativas. As medianas dos *boxplots* da Figura A.15 reafirmam esta análise. Percebe-se também a redução gradativa na dispersão a partir de $m = 20$ e, como em todos os outros casos, chegando a quase nula em $m = 2$, apesar do solitário *outlier* também visto no problema de teste anterior.

A.4 DTLZ4

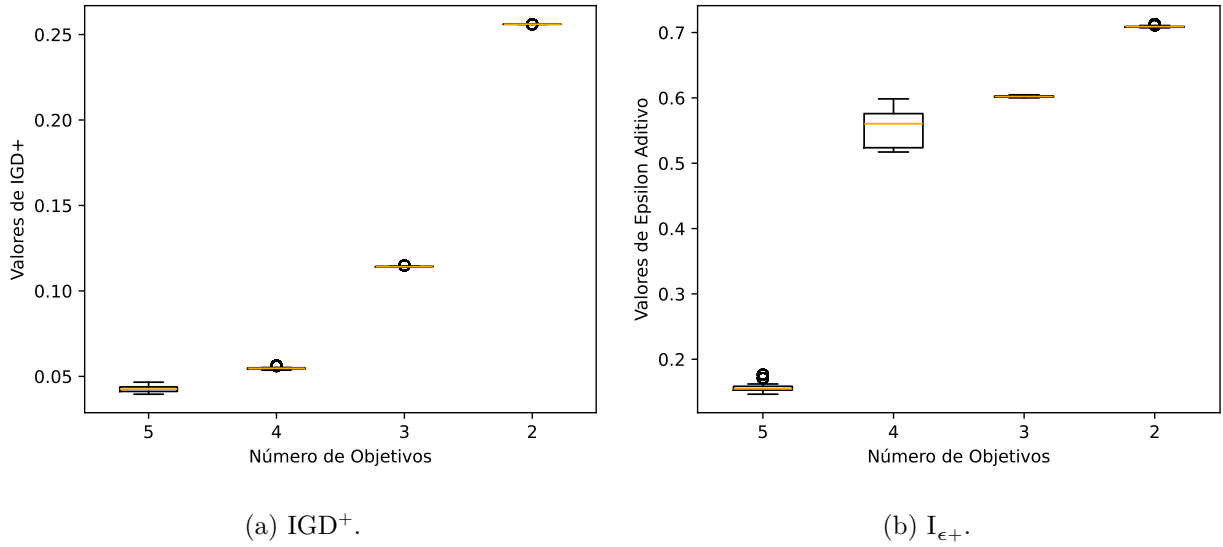
O problema DTLZ4, conforme a Equação 2.7, é uma variação do DTLZ2 com um parâmetro α modificado, resultando em um conjunto denso de soluções. Os resultados mostram uma distribuição de escores distinta, refletindo a influência da densidade das amostras.

Tabela A.31: Escores: DTLZ4 com $m = 5$.

f_2	f_5	f_4	f_1	f_3
0.0091	0.0132	0.0153	0.0181	0.0204

Tabela A.32: Resultados da seleção de objetivos: DTLZ4 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.0427	0.0547	0.1142	0.2560
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.1566	0.5532	0.6020	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001

Figura A.16: *Boxplots* dos IQs: DTLZ4 por seleção com $m = 5$.

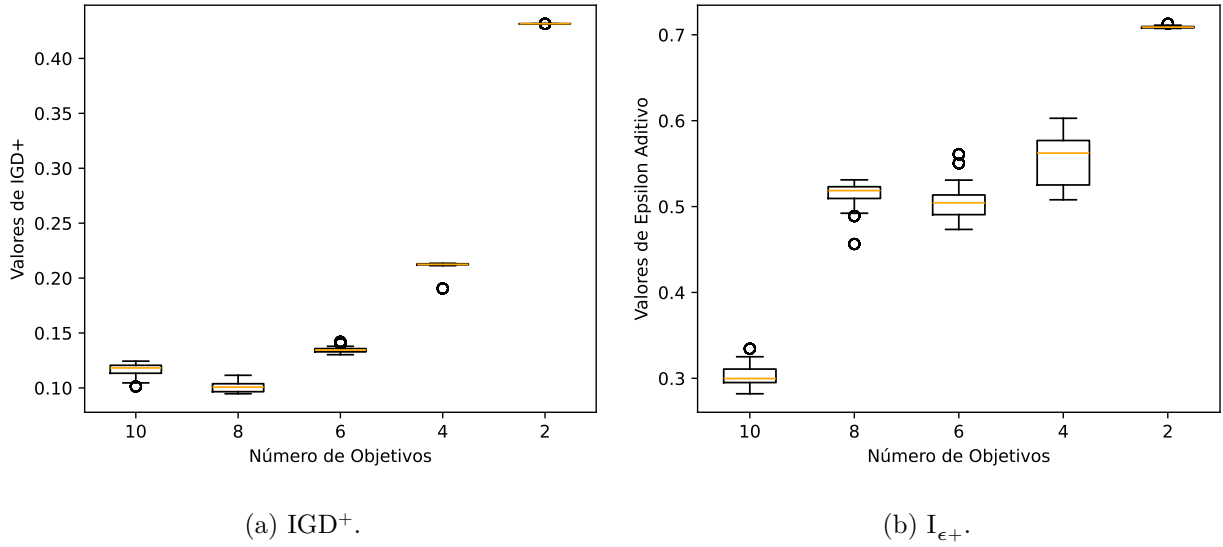
Analisando a Tabela A.31 para este MaOP com $m = 5$, nota-se que os escores apresentaram uma ordenação incomum quando comparado com os demais DTLZ discutidos até aqui. Como pode ser observado na Tabela A.32, nenhum dos cenários de redução apresentou resultados superiores ao MaOP original. Além disso, é perceptível o crescimento dos IQs nos *boxplots* da Figura A.16, assim como uma dispersão muito baixa em todos os casos, com exceção do $I_{\epsilon+}$ em $m = 4$.

Tabela A.33: Escores: DTLZ4 com $m = 10$.

f_8	f_6	f_9	f_4	f_{10}	f_5	f_7	f_1	f_2	f_3
0.0151	0.0201	0.0235	0.0257	0.0258	0.0305	0.0318	0.032	0.0333	0.045

Tabela A.34: Resultados da seleção de objetivos: DTLZ4 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	0.1162	0.1010	0.1347	0.2114	0.4315
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3025	0.5127	0.5076	0.5545	0.7090
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.17: *Boxplots* dos IQs: DTLZ4 por seleção com $m = 10$.

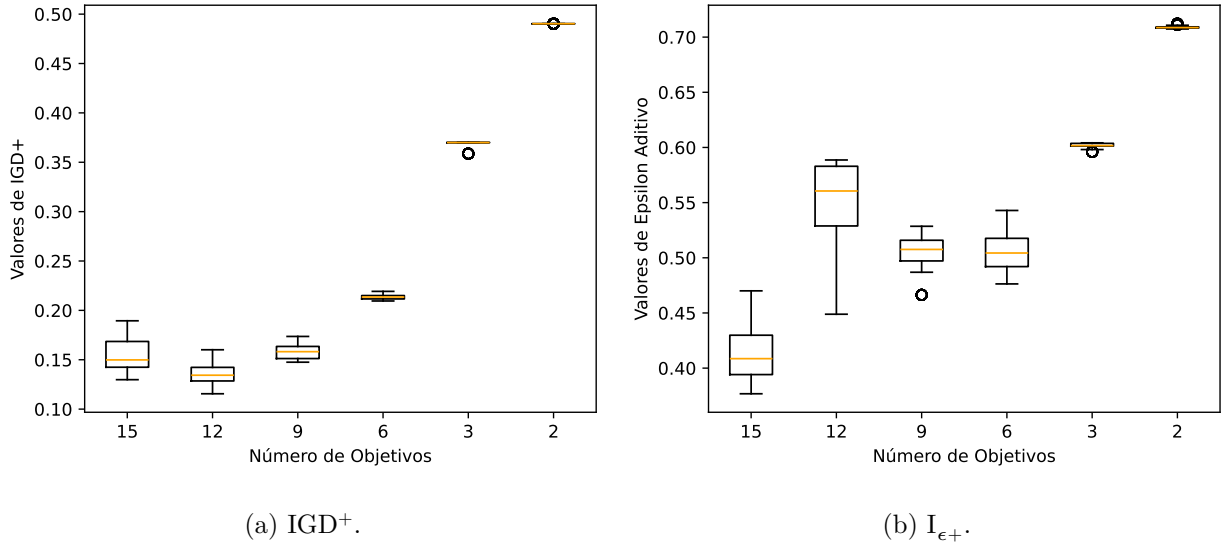
Para o cenário com $m = 10$, observando a Tabela A.33, observa-se que a escala dos valores dos escores está bem menor que os demais DTLZ até aqui, indicando a concentração de soluções atribuída ao DTLZ4. A Tabela A.34 mostra os resultados dos IQs e nota-se apenas um ligeiro ganho no caso $m = 8$ para o IGD⁺ com evidências estatísticas para rejeição da hipótese nula, por outro lado, houve uma piora no I _{ϵ +}, comprometendo esta redução. Analisando-se a Figura A.17, percebe-se que os *boxplots* dos IQs diferem entre si nas distribuições, exceto para $m = 2$, onde é praticamente nula em ambos. Também verifica-se a mediana de $m = 8$ do IGD⁺ abaixo da mediana original.

Tabela A.35: Escores: DTLZ4 com $m = 15$.

f_7	f_{12}	f_8	f_3	f_5	f_6	f_{14}	f_4	f_1	f_9
0.014	0.0145	0.0192	0.0198	0.024	0.0253	0.0261	0.0274	0.036	0.036
f_{11}	f_{10}	f_2	f_{15}	f_{13}					
0.0381	0.0433	0.0438	0.0443	0.0689					

Tabela A.36: Resultados da seleção de objetivos: DTLZ4 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.1545	0.1347	0.1576	0.2136	0.3694	0.4902
p-valor	1	< 0.001	0.3455	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.4152	0.5487	0.5061	0.5060	0.6018	0.7089
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.18: *Boxplots* dos IQs: DTLZ4 por seleção com $m = 15$.

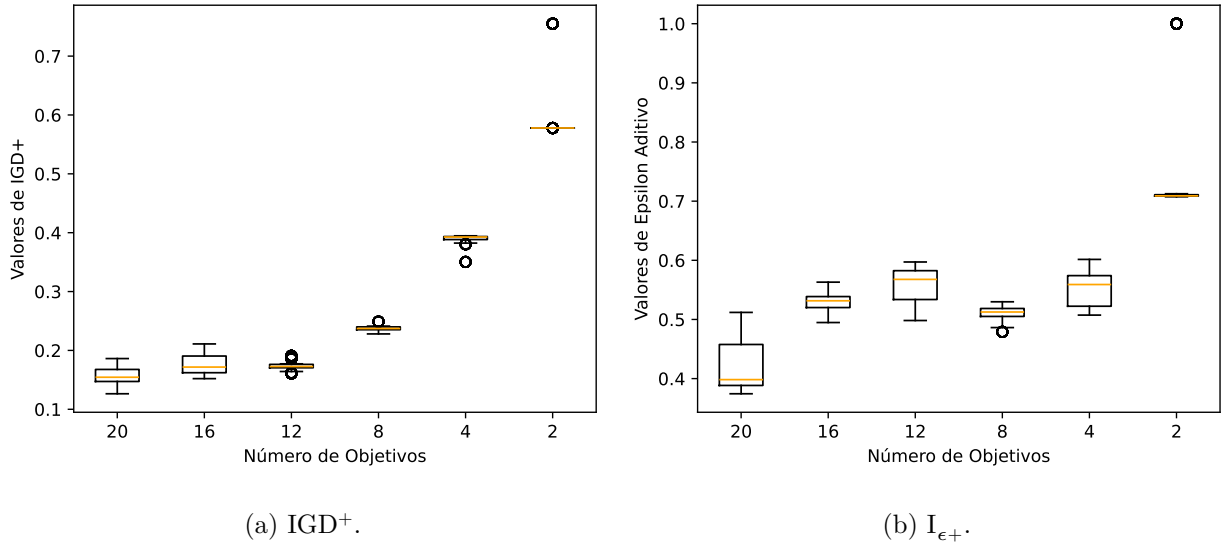
Para este MaOP com $m = 15$, a Tabela A.35 exibe *ranking* e analisando-o percebe-se o mesmo padrão que nos cenários anteriores, reforçando a dificuldade da abordagem LS em identificar padrões nos dados gerados do DTLZ4. Observando a Tabela A.36, nota-se que os IQs apresentaram uma piora gradativa com relação ao problema original. Apenas um tênue ganho no caso $m = 12$ para o IGD⁺ com evidências estatísticas para rejeição da hipótese nula, por outro lado, houve uma piora no $I_{\epsilon+}$ que compromete esta redução, semelhante ao resultado do cenário anterior com $m = 10$. A respeito da Figura A.18, há também uma semelhança nas dispersões dos *boxplots* em ambos os IQs, visto que $I_{\epsilon+}$ produziu um conjunto mais distribuído de resultados, salvo as reduções para $m = 3$ e $m = 2$.

Tabela A.37: Escores: DTLZ4 com $m = 20$.

f_3	f_7	f_4	f_6	f_{15}	f_8	f_{19}	f_{12}	f_9	f_{20}
0.0282	0.0297	0.0319	0.0335	0.0338	0.0347	0.0353	0.0371	0.0399	0.0404
f_1	f_{17}	f_{13}	f_{10}	f_2	f_{18}	f_{16}	f_5	f_{14}	f_{11}
0.0405	0.0408	0.0434	0.0445	0.0464	0.0465	0.0489	0.0527	0.0593	0.0882

Tabela A.38: Resultados da seleção de objetivos: DTLZ4 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	0.1558	0.1749	0.1732	0.2369	0.3890	0.6031
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.4207	0.5315	0.5582	0.5099	0.5548	0.7506
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.19: *Boxplots* dos IQs: DTLZ4 por seleção com $m = 20$.

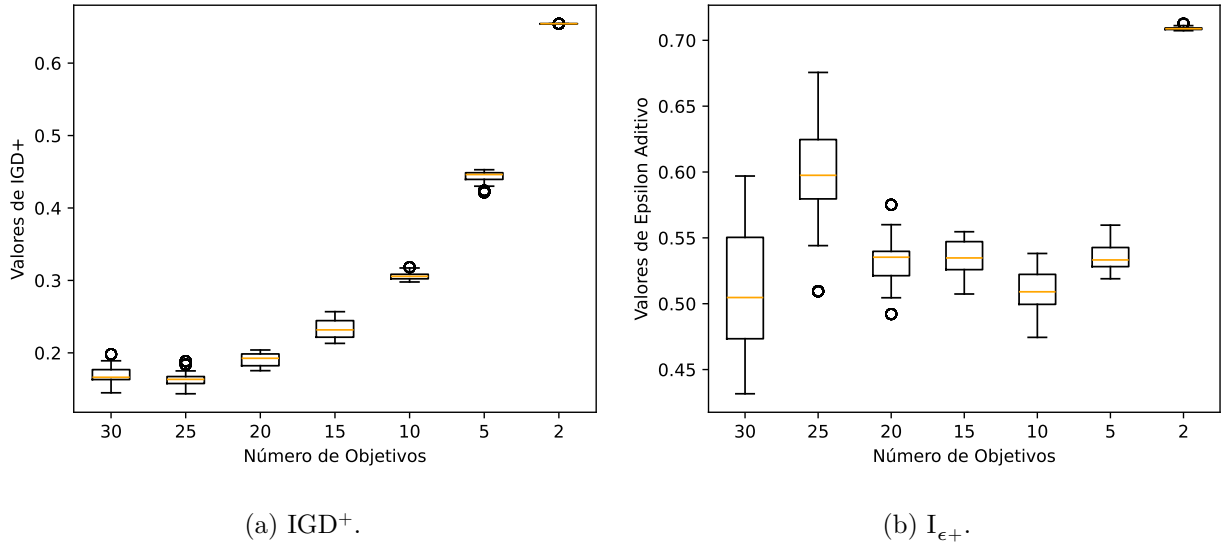
Avaliando a Tabela A.37, percebe-se um comportamento nos escores semelhante ao dos cenários anteriores deste MaOP. Além disso, a mesma análise pode ser feita para os dados exibidos na Tabela A.38, que mostram uma perda gradual de desempenho manifestada nos IQs e nos seus *boxplots* na Figura A.19. Repete-se também a equivalência nas dispersões, se comparado com os MaOP anteriores desta seção, porém desta vez ocorreu uma solução *outlier* em $m = 2$.

Tabela A.39: Escores: DTLZ4 com $m = 30$.

f_7	f_{27}	f_2	f_{20}	f_3	f_9	f_{16}	f_{22}	f_{25}	f_6
0.0295	0.0408	0.042	0.0422	0.0441	0.0447	0.0469	0.0494	0.0505	0.051
f_5	f_1	f_{12}	f_{19}	f_{10}	f_{30}	f_{28}	f_{23}	f_{11}	f_{24}
0.0521	0.0563	0.0571	0.0572	0.0594	0.0608	0.0618	0.0622	0.0649	0.0661
f_8	f_{14}	f_{13}	f_{18}	f_{17}	f_4	f_{15}	f_{21}	f_{26}	f_{29}
0.0671	0.0671	0.0713	0.0726	0.0776	0.0809	0.0821	0.0982	0.0994	0.1603

Tabela A.40: Resultados da seleção de objetivos: DTLZ4 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	0.1693	0.1634	0.1904	0.2320	0.3067	0.4422	0.6543
p-valor	1	0.1480	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.5122	0.5985	0.5318	0.5357	0.5085	0.5364	0.7090
p-valor	1	< 0.001	0.1951	0.1101	0.8701	0.0994	< 0.001

Figura A.20: *Boxplots* dos IQs: DTLZ4 por seleção com $m = 30$.

Por último, o MaOP com $m = 30$ apresentou o mesmo padrão no *ranking* dos escores até aqui, conforme mostrado na Tabela A.39. Novamente, a Tabela A.40 e a Figura A.20 revelam a mesma perda exponencial de desempenho para o IGD⁺ após a aplicação do método LS. Analisando as médias e os *boxplots* do $I_{\epsilon+}$, verificam-se resultados mais próximos entre si, apesar da distribuição pior que do outro IQ, com exceção de $m = 2$ como nos cenários anteriores.

A.5 DTLZ5

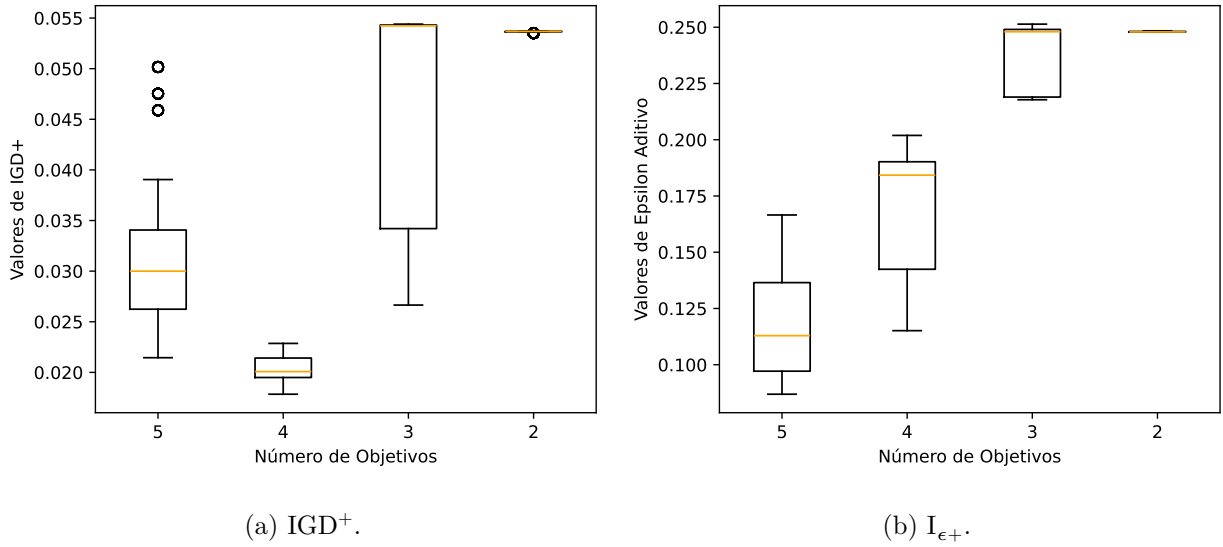
O problema DTLZ5, definido pela Equação 2.8, é utilizado para avaliar a capacidade de um MOEA em convergir para uma curva específica. Os resultados seguem um padrão semelhante aos observados nos DTLZ1, DTLZ2 e DTLZ3.

Tabela A.41: Escores: DTLZ5 com $m = 5$.

f_5	f_4	f_1	f_3	f_2
0.011	0.0142	0.021	0.0217	0.0258

Tabela A.42: Resultados da seleção de objetivos: DTLZ5 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.0317	0.0204	0.0467	0.0536
p-valor	1	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.1199	0.1688	0.2388	0.2479
p-valor	1	< 0.001	< 0.001	< 0.001

Figura A.21: *Boxplots* dos IQs: DTLZ5 por seleção com $m = 5$.

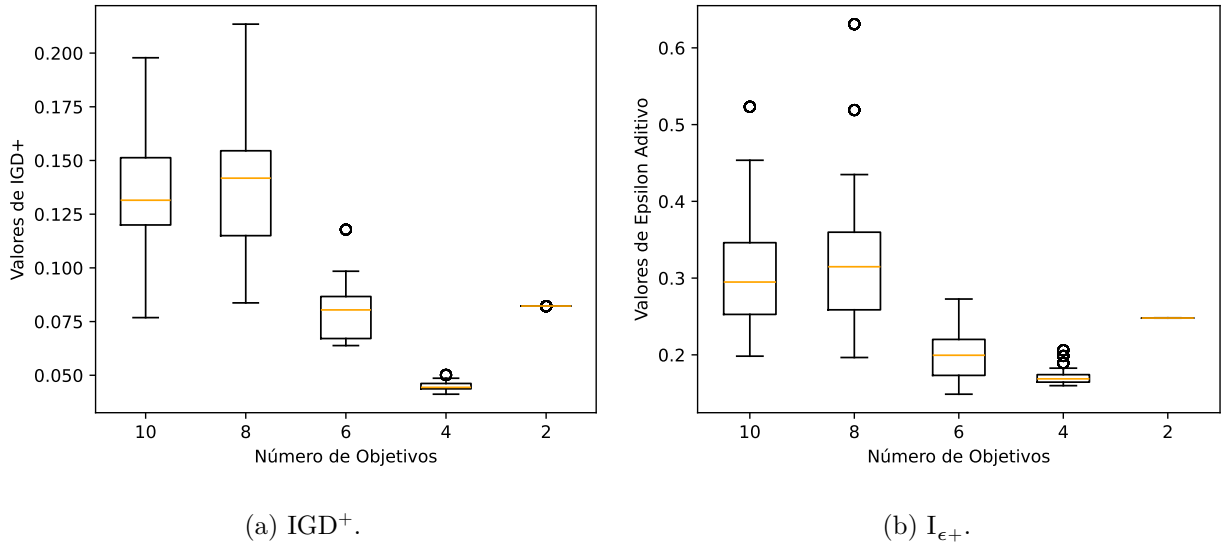
Primeiramente, a Tabela A.41 dispõe os escores calculados para o cenário com $m = 5$ e, diferentemente do DTLZ4, este problema indica que seguirá o padrão evidenciado pelos outros DTLZ tratados até aqui, uma vez que as funções com maiores índices (f_5 e f_4) resultaram nos menores valores de escores. Examinando a Tabela A.43 que contém os resultados dos IQs, nota-se que apenas o cenário com $m = 4$ alcançou um resultado positivo de IGD⁺ com p-valor < 5%, porém o mesmo não foi obtido pelo I _{ϵ +} e os demais IQs dos outros casos. Esta análise pode ser verificada na Figura A.21, com adição da mediana do $m = 3$ no terceiro quartil e da baixíssima distribuição dos resultados na última redução ($m = 2$).

Tabela A.43: Escores: DTLZ5 com $m = 10$.

f_{10}	f_9	f_8	f_7	f_2	f_1	f_3	f_4	f_5	f_6
0.0276	0.0305	0.0414	0.0516	0.0599	0.061	0.064	0.0719	0.073	0.0735

Tabela A.44: Resultados da seleção de objetivos: DTLZ5 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	0.1351	0.1377	0.0798	0.0449	0.0822
p-valor	1	0.9498	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3130	0.3299	0.2001	0.1723	0.2479
p-valor	1	0.6597	< 0.001	< 0.001	0.0036

Figura A.22: *Boxplots* dos IQs: DTLZ5 por seleção com $m = 10$.

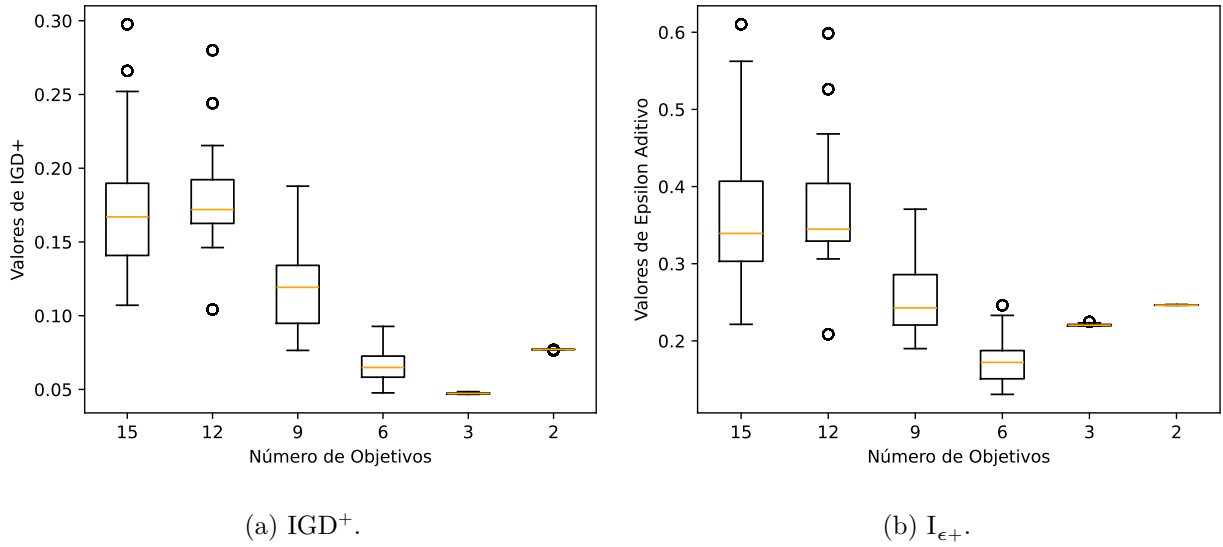
Em sequência, as Tabelas A.43 e A.44 dispõem os resultados do MaOP com $m = 10$, observa-se que os escores reforçam a afirmativa feita no cenário anterior, pois a ordem dos índices dos objetivos passou a ser decrescente novamente. Além disso, a redução para $m = 4$ apresentou resultados positivos em ambos os IQs, com evidências estatísticas para rejeição da hipótese nula. A respeito dos *boxplots*, nota-se as medianas de $m = 4$ as mais baixas e as de $m = 2$ acima de $m = 6$ na Figura A.22, assim como as dispersões diminuíram conforme a redução do problema.

Tabela A.45: Escores: DTLZ5 com $m = 15$.

f_{15}	f_{14}	f_{13}	f_{11}	f_{12}	f_{10}	f_3	f_5	f_7	f_8
0.035	0.0414	0.0532	0.0585	0.0594	0.0701	0.0808	0.0845	0.0922	0.0947
f_2	f_4	f_1	f_6	f_9					
0.1048	0.1068	0.1097	0.117	0.15					

Tabela A.46: Resultados da seleção de objetivos: DTLZ5 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.1751	0.1801	0.1209	0.0663	0.0473	0.0770
p-valor	1	0.3854	0.2546	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3623	0.3732	0.2546	0.1748	0.2207	0.2464
p-valor	1	0.4429	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.23: *Boxplots* dos IQs: DTLZ5 por seleção com $m = 15$.

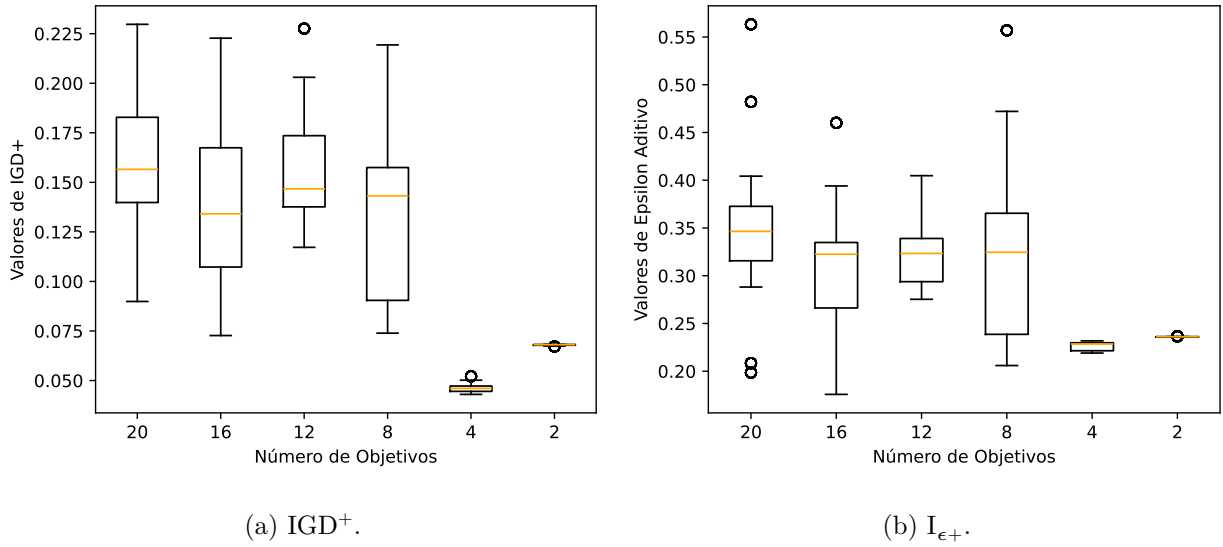
A respeito dos resultados do MaOP com $m = 15$, dispostos nas Tabelas A.45 e A.46, o *ranking* do LS apresentou o mesmo comportamento do cenário para $m = 10$, assim como ambos os IQs apontam com evidências estatísticas o ganho de desempenho do MOEA em todas as abordagens, exceto para $m = 12$. Com destaque para os casos $m = 6$ e $m = 3$, pois indicam os valores mais baixos entre todos. Pela Figura A.23 é possível notar que, como no cenário anterior, a penúltima abordagem ($m = 3$) com medianas menores que as da última ($m = 2$), onde também as duas alcançaram distribuições quase nulas.

Tabela A.47: Escores: DTLZ5 com $m = 20$.

f_{20}	f_{19}	f_{16}	f_{15}	f_{18}	f_{14}	f_{12}	f_{17}	f_1	f_5
0.0423	0.0504	0.0728	0.0732	0.0745	0.0769	0.0892	0.0913	0.0921	0.0984
f_{11}	f_{10}	f_9	f_6	f_8	f_3	f_{13}	f_7	f_2	f_4
0.1041	0.1048	0.1095	0.1098	0.1124	0.1131	0.1147	0.1151	0.1186	0.1188

Tabela A.48: Resultados da seleção de objetivos: DTLZ5 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	0.1573	0.1405	0.1557	0.1317	0.0462	0.0679
p-valor	1	0.1626	0.6237	0.0403	< 0.001	< 0.001
$I_{\epsilon+}$	0.3482	0.3119	0.3261	0.3220	0.2261	0.2360
p-valor	1	0.1046	0.1552	0.2421	< 0.001	< 0.001

Figura A.24: *Boxplots* dos IQs: DTLZ5 por seleção com $m = 20$.

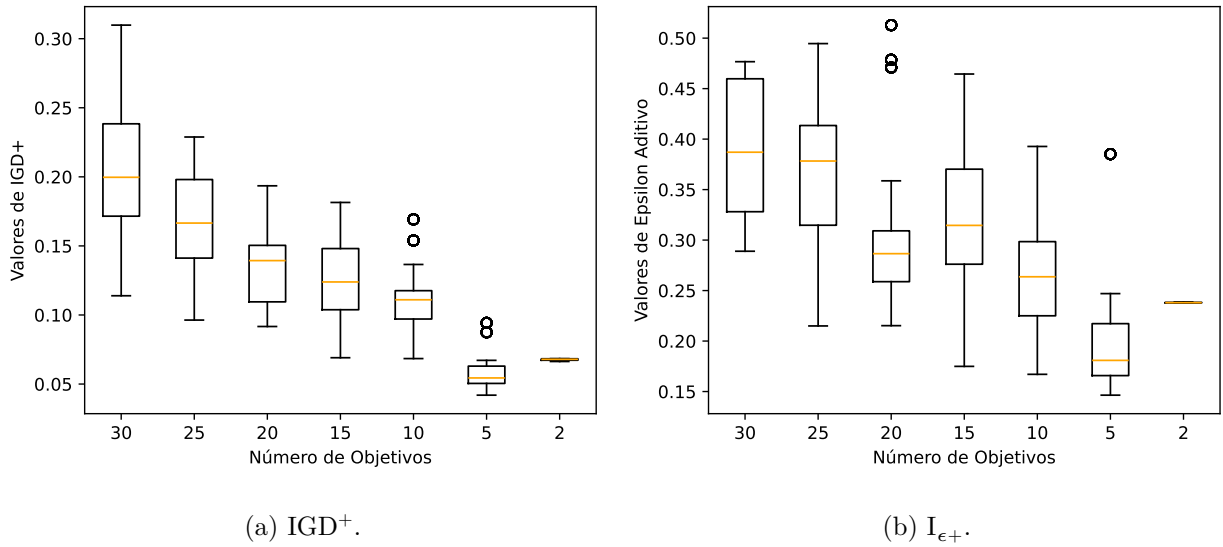
Analisando a Tabela A.47, percebe-se um desempenho nos resultados dos escores similar ao dos cenários anteriores do DTLZ5. Sobre a Tabela A.48, houveram melhorias no IGD⁺ e no $I_{\epsilon+}$ para todas as reduções de dimensionalidade, porém apenas as abordagens $m = 4$ e $m = 2$ apresentaram evidências estatísticas para rejeição da hipótese nula, com destaque para a primeira, que alcançou grandezas abaixo das demais. Verificando a Figura A.24, reforça-se esta análise e, novamente, a baixa dispersão nas duas últimas dimensões.

Tabela A.49: Escores: DTLZ5 com $m = 30$.

f_{30}	f_{29}	f_{28}	f_{25}	f_{27}	f_{26}	f_5	f_{13}	f_{11}	f_{18}
0.0459	0.046	0.0625	0.0821	0.0827	0.0916	0.0995	0.1027	0.1029	0.1041
f_{24}	f_{14}	f_6	f_{15}	f_9	f_7	f_{19}	f_8	f_3	f_{16}
0.1072	0.109	0.1094	0.1116	0.1123	0.1155	0.1191	0.12	0.1292	0.1306
f_{10}	f_{12}	f_4	f_{23}	f_{20}	f_{22}	f_2	f_{21}	f_{17}	f_1
0.1314	0.1321	0.1327	0.1333	0.1339	0.1345	0.1367	0.1367	0.1438	0.1442

Tabela A.50: Resultados da seleção de objetivos: DTLZ5 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	0.2016	0.1703	0.1346	0.1252	0.1103	0.0580	0.0676
p-valor	1	0.0295	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
I _{ϵ+}	0.3873	0.3631	0.3106	0.3180	0.2646	0.1973	0.2380
p-valor	1	0.3993	< 0.001	0.0028	< 0.001	< 0.001	< 0.001

Figura A.25: *Boxplots* dos IQs: DTLZ5 por seleção com $m = 30$.

Como último cenário, o MaOP com $m = 30$ praticamente manteve o comportamento observado nos cenários anteriores, conforme a Tabela A.49, e é plausível dizer que há um padrão nos resultados dos escores para este problema de teste. Avaliando a Tabela A.50 que dispõe os resultados dos IQs, conclui-se que houveram melhorias de desempenho em todas as abordagens, exceto para $m = 25$ que apresentou um p-valor $> 5\%$ no I _{ϵ +}. Destaca-se a redução para $m = 5$ que obteve os menores valores para ambos os IQs. Os *boxplots* da Figura A.25 mostram, assim como nos três cenários anteriores, que a mediana da penúltima

abordagem ($m = 5$) é menor que a da última e também que as duas apresentaram as menores dispersões, com ênfase para $m = 2$.

A.6 DTLZ6

O problema DTLZ6, conforme a Equação 2.9, é uma variação do DTLZ5 projetada para dificultar a convergência do MOEA. Os resultados mostram um padrão consistente com os observados nos problemas anteriores.

Tabela A.51: Escores: DTLZ6 com $m = 5$.

f_5	f_4	f_3	f_2	f_1
0.0115	0.0135	0.0205	0.0287	0.03

Tabela A.52: Resultados da seleção de objetivos: DTLZ6 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.6422	0.2871	0.1817	0.1375
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	1.0798	0.7104	0.5304	0.4110
p-valor	1	< 0.001	< 0.001	< 0.001

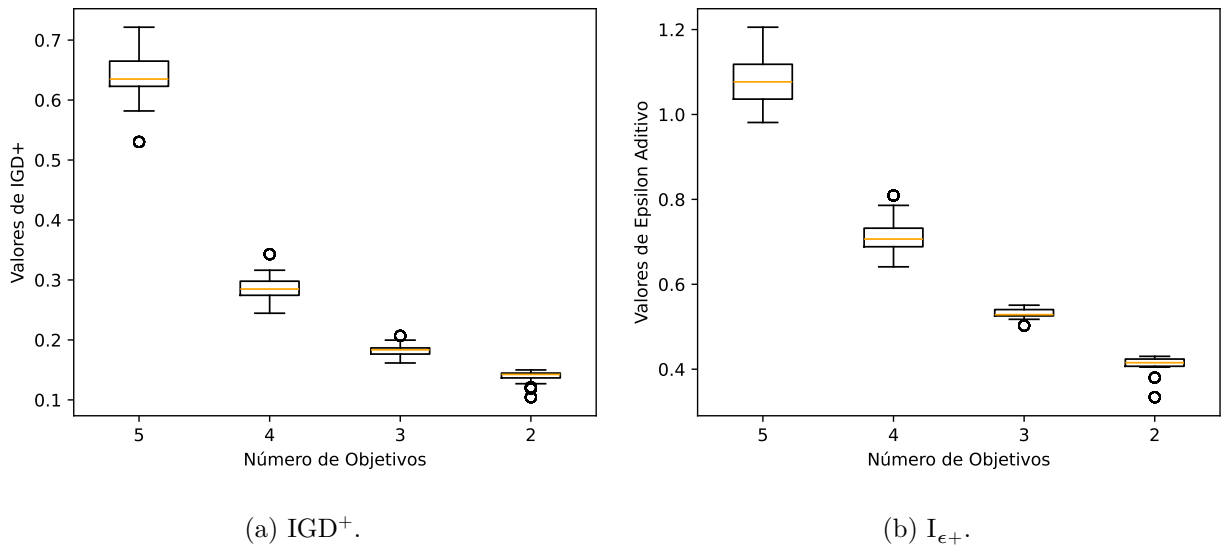


Figura A.26: *Boxplots* dos IQs: DTLZ6 por seleção com $m = 5$.

Começando pelo cenário com $m = 5$, a Tabela A.51 apresenta os escores obtidos pelo LS para cada objetivo. Nela, nota-se que a ordem dos índices das funções foi rigorosamente

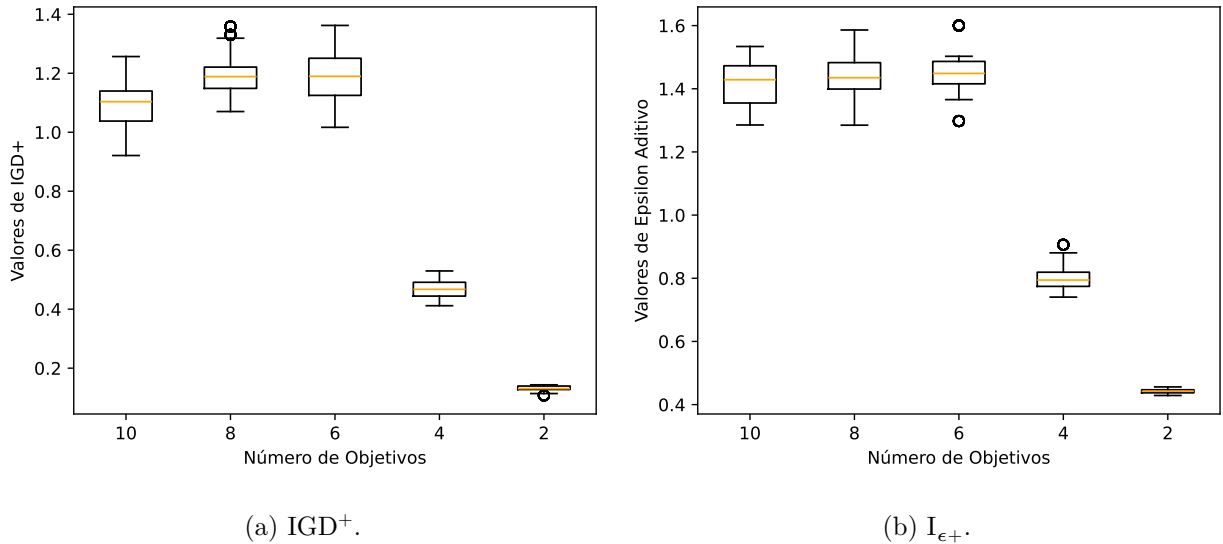
invertida. A Tabela A.52 e a Figura A.26 mostram os resultados dos IQs e seus *boxplots* para esta análise, respectivamente, permitindo concluir que as médias e as medianas das reduções foram gradativamente menores à medida que a redução se agravava, bem como as distribuições dos IQs nos diagramas. Por fim, o cenário com melhor ganho de desempenho foi o de dimensão $m = 2$, além disso, todas as abordagens obtiveram p-valor $< 5\%$, indicando estatisticamente a rejeição da hipótese nula.

Tabela A.53: Escores: DTLZ6 com $m = 10$.

f_{10}	f_9	f_8	f_7	f_6	f_5	f_4	f_1	f_2	f_3
0.0233	0.0265	0.0412	0.056	0.0918	0.1289	0.1514	0.1774	0.2056	0.2177

Tabela A.54: Resultados da seleção de objetivos: DTLZ6 com $m = 10$.

m	10	8	6	4	2
IGD ⁺	1.0906	1.1942	1.1931	0.4679	0.1313
p-valor	1	0.0013	0.0046	< 0.001	< 0.001
$I_{\epsilon+}$	1.4199	1.4398	1.4479	0.8037	0.4418
p-valor	1	0.6060	0.4734	< 0.001	< 0.001

Figura A.27: *Boxplots* dos IQs: DTLZ6 por seleção com $m = 10$.

Para o DTLZ6 com $m = 10$, a Tabela A.53 dispõe os escores calculados pelo LS para cada objetivo, mostrando que a ordem dos índices das funções foi quase completamente invertida. Além disso, os resultados da Tabela A.54 evidenciam que as reduções com $m = 8$ e $m = 6$ apresentaram IQs acima da média do original, com $I_{\epsilon+}$ estatisticamente igual, conforme o

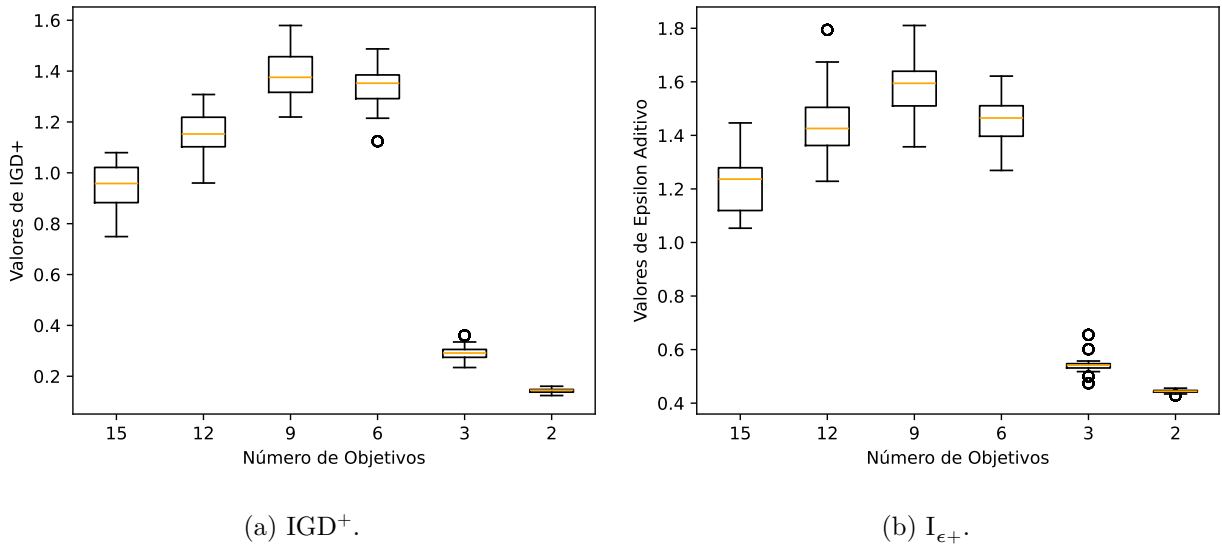
teste de hipóteses. Por outro lado, os cenários com $m = 4$ e $m = 2$ mostraram melhorias com evidências estatísticas para rejeição da hipótese nula, com grande destaque para o último. Pela Figura A.27, também é possível verificar esta análise pelas medianas dos *boxplots* e como a dispersão dos IQs de $m = 2$ está muito baixa.

Tabela A.55: Escores: DTLZ6 com $m = 15$.

f_{15}	f_{14}	f_{13}	f_{12}	f_{11}	f_{10}	f_9	f_7	f_8	f_5
0.0237	0.0267	0.0376	0.0632	0.0999	0.1181	0.1836	0.2203	0.2353	0.2831
f_3	f_1	f_6	f_4	f_2					
0.2893	0.3113	0.3141	0.3659	0.3736					

Tabela A.56: Resultados da seleção de objetivos: DTLZ6 com $m = 15$.

m	15	12	9	6	3	2
IGD ⁺	0.9548	1.1532	1.3770	1.3425	0.2951	0.1430
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	1.2159	1.4365	1.5799	1.4458	0.5433	0.4437
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.28: *Boxplots* dos IQs: DTLZ6 por seleção com $m = 15$.

As Tabelas A.55 e A.56 apresentam os resultados dos escores e dos IQs para este MaOP com $m = 15$, respectivamente. Ao analisá-las, observa-se que o padrão de ordenação dos índices foi mantido e as reduções mais incisivas ($m = 3$ e $m = 2$) obtiveram médias menores de IGD⁺ e $I_{\epsilon+}$ em comparação com o original. Além disso, similar ao cenário anterior,

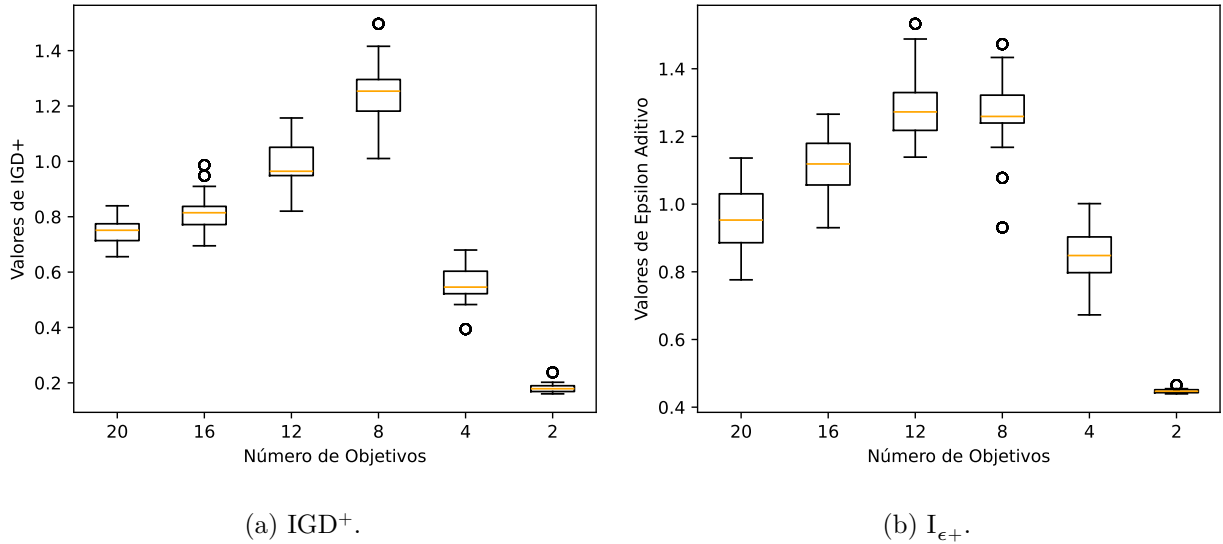
houve uma deterioração nos IQs médios para as demais reduções. Estas observações estão ilustradas também nas medianas da Figura A.28 e nas dispersões dos *boxplots* das últimas abordagens. É importante notar que todos os resultados dos IQs mostraram evidências estatísticas suficientes para rejeitar a hipótese nula.

Tabela A.57: Escores: DTLZ6 com $m = 20$.

f_{20}	f_{19}	f_{18}	f_{17}	f_{16}	f_{15}	f_{14}	f_{13}	f_{11}	f_8
0.0223	0.0278	0.0417	0.0647	0.108	0.1646	0.1726	0.2091	0.2692	0.2754
f_{12}	f_6	f_9	f_{10}	f_5	f_7	f_1	f_2	f_4	f_3
0.3052	0.311	0.3121	0.3684	0.3729	0.3751	0.3829	0.4098	0.42	0.4477

Tabela A.58: Resultados da seleção de objetivos: DTLZ6 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	0.7461	0.8131	0.9883	1.2486	0.5569	0.1820
p-valor	1	0.002	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.9705	1.1128	1.2802	1.2689	0.8516	0.4479
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.29: *Boxplots* dos IQs: DTLZ6 por seleção com $m = 20$.

Agora para o cenário com $m = 20$, a Tabela A.57 apresenta os escores calculados pelo LS, evidenciando o mesmo padrão na ordem dos índices das funções observado nos casos anteriores. Da mesma forma, a Tabela A.58 dos IQs mostra uma melhora de desempenho nas reduções para $m = 4$ e $m = 2$, com destaque especialmente para esta última, além de

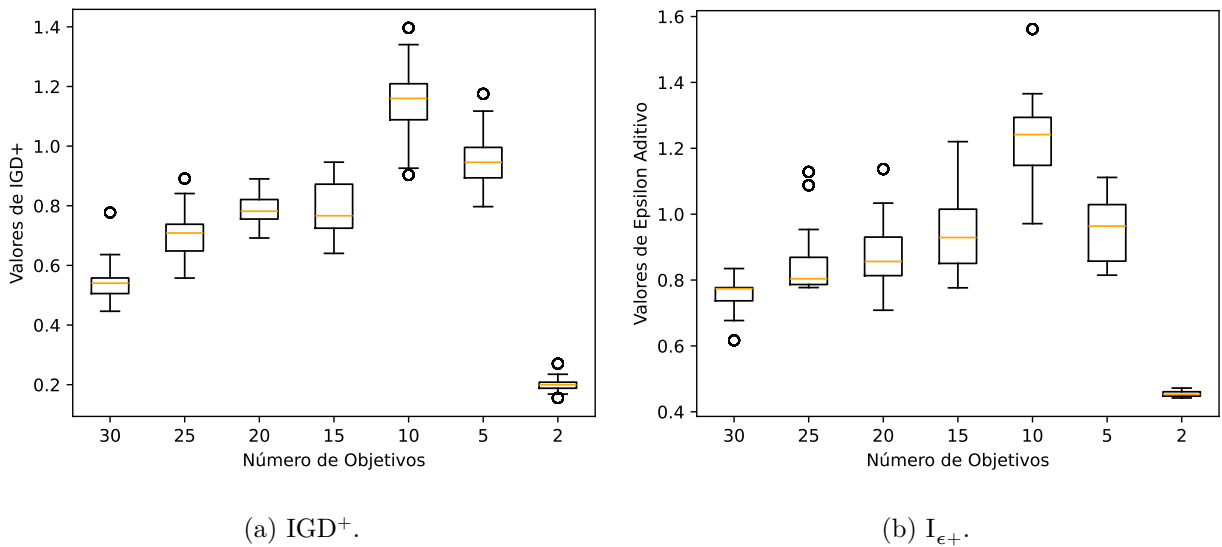
um p-valor $< 5\%$ em todos os casos. Observa-se também uma deterioração progressiva nos indicadores até o MaOP com $m = 8$. Analisando a Figura A.29, reforçam-se estas afirmações com adição das medianas, mas desta vez apenas $m = 2$ obteve distribuições nos resultados muito baixas.

Tabela A.59: Escores: DTLZ6 com $m = 30$.

f_{30}	f_{29}	f_{28}	f_{27}	f_{26}	f_{24}	f_{25}	f_{23}	f_{21}	f_{22}
0.0256	0.0295	0.0401	0.0773	0.1182	0.1647	0.1742	0.1837	0.2483	0.2875
f_{18}	f_{19}	f_{20}	f_{17}	f_4	f_9	f_2	f_{14}	f_7	f_{10}
0.2909	0.3363	0.3428	0.3479	0.3546	0.3574	0.366	0.3837	0.3903	0.3915
f_3	f_6	f_{15}	f_{12}	f_5	f_8	f_{13}	f_{11}	f_{16}	f_1
0.3977	0.4002	0.4052	0.4062	0.4203	0.4298	0.4367	0.4727	0.4747	0.5384

Tabela A.60: Resultados da seleção de objetivos: DTLZ6 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	0.5447	0.6938	0.7858	0.7939	1.1444	0.9613	0.2010
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	0.7564	0.8455	0.8757	0.9421	1.2168	0.9536	0.4537
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura A.30: *Boxplots* dos IQs: DTLZ6 por seleção com $m = 30$.

Por fim o DTLZ6 com $m = 30$, a Tabela A.59 revela que os escores seguiram novamente o padrão de inversão nos índices dos objetivos. A Tabela A.60 apresenta os IQs calculados, nos quais, assim como no cenário anterior, nota-se uma piora progressiva nas médias até

a redução para $m = 5$. Adicionalmente, desta vez a penúltima redução não melhorou o desempenho, somente para $m = 2$ que apresentou p-valores abaixo de 5% e demonstram evidências estatísticas para rejeição da hipótese nula. Os traços das medianas na Figura A.30 evidenciam as perdas de desempenho citadas, assim como o ganho da redução para $m = 2$ e sua baixa dispersão no *boxplot*.

A.7 DTLZ7

O problema DTLZ7, definido pela Equação 2.10, é utilizado para testar a capacidade de um MOEA em manter subpopulações distribuídas em diferentes regiões Pareto-ótimas. Os resultados mostram um padrão distinto, com a última função objetivo apresentando escores consistentemente mais baixos.

Tabela A.61: Escores: DTLZ7 com $m = 5$.

f_1	f_3	f_4	f_2	f_5
0.0342	0.035	0.0353	0.0379	0.1303

Tabela A.62: Resultados da seleção de objetivos: DTLZ7 com $m = 5$.

m	5	4	3	2
IGD ⁺	0.2475	0.2012	0.33515	0.5260
p-valor	1	0.0011	< 0.001	< 0.001
$I_{\epsilon+}$	0.8579	0.8816	0.9569	0.8652
p-valor	1	0.1552	0.0277	0.0152

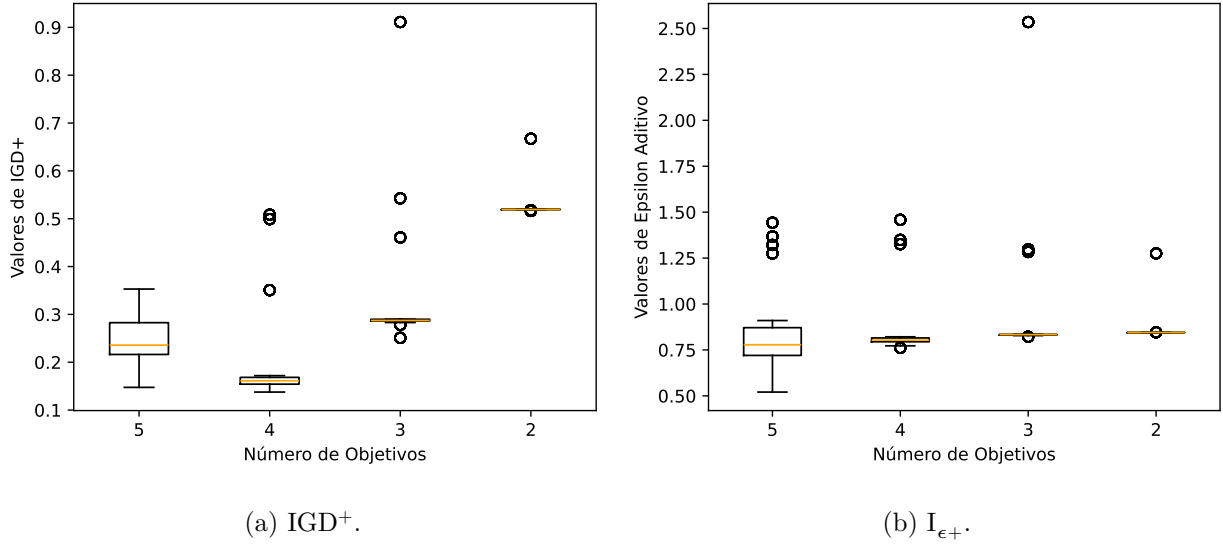


Figura A.31: *Boxplots* dos IQs: DTLZ7 por seleção com $m = 5$.

Como pode ser observado na Tabela A.61, a ordenação dos objetivos não parece ter seguido um padrão claro. Além disso, nota-se que a função f_5 está significativamente distante das demais em termos de magnitude. Conforme evidenciado na Tabela A.62 e na Figura A.31, a redução para $m = 4$ resultou em uma média e uma mediana de IGD^+ menores que as originais, juntamente com o p-valor $< 5\%$. Porém, não observa-se o mesmo para o $I_{\epsilon+}$, onde o p-valor sugere evidências estatísticas de ser igual ao problema original. Adicionalmente, houve uma deterioração gradual de desempenho nos demais casos, bem como a incidência de *outliers* em todos.

Tabela A.63: Escores: DTLZ7 com $m = 10$.

f_6	f_1	f_5	f_2	f_7	f_3	f_4	f_9	f_8	f_{10}
0.1373	0.1375	0.1396	0.1399	0.1437	0.144	0.146	0.1468	0.1491	0.4627

Tabela A.64: Resultados da seleção de objetivos: DTLZ7 com $m = 10$.

m	10	8	6	4	2
IGD^+	1.9907	1.0115	0.8131	0.9102	1.1674
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	6.4150	3.5018	2.2225	1.5663	0.9478
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

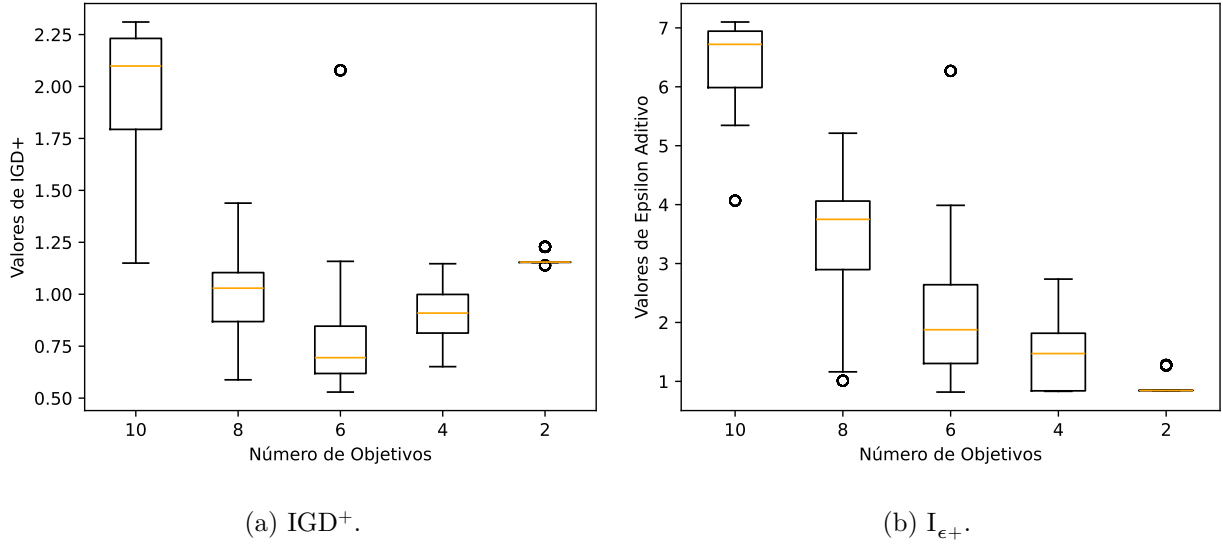


Figura A.32: *Boxplots* dos IQs: DTLZ7 por seleção com $m = 10$.

Para o DTLZ7 com $m = 10$, reforça-se a análise sobre a aleatoriedade na ordenação das funções feita no cenário anterior, conforme a Tabela A.63. Porém, desta vez houveram ganhos de desempenho significativos em todas as reduções, como pode ser observado nas médias e p-valores da Tabela A.64, e nas medianas dos *boxplots* na Figura A.32, além da dispersão quase nula dos IQs da dimensão $m = 2$. Destaca-se também a diferença da magnitude de escore para f_{10} , outra vez a função com maior índice do MaOP.

Tabela A.65: Escores: DTLZ7 com $m = 15$.

f_2	f_7	f_{11}	f_{10}	f_6	f_5	f_4	f_8	f_{14}	f_1
0.2255	0.2265	0.2292	0.2295	0.2299	0.2306	0.2316	0.2321	0.2342	0.2353
f_{12}	f_{13}	f_9	f_3	f_{15}					
0.2354	0.2368	0.2383	0.2436	0.5476					

Tabela A.66: Resultados da seleção de objetivos: DTLZ7 com $m = 15$.

m	15	12	9	6	3	2
IGD^+	5.9453	3.4795	1.8746	1.8224	1.9010	1.7413
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	15.7724	11.2937	6.9464	5.8868	4.0496	1.8453
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

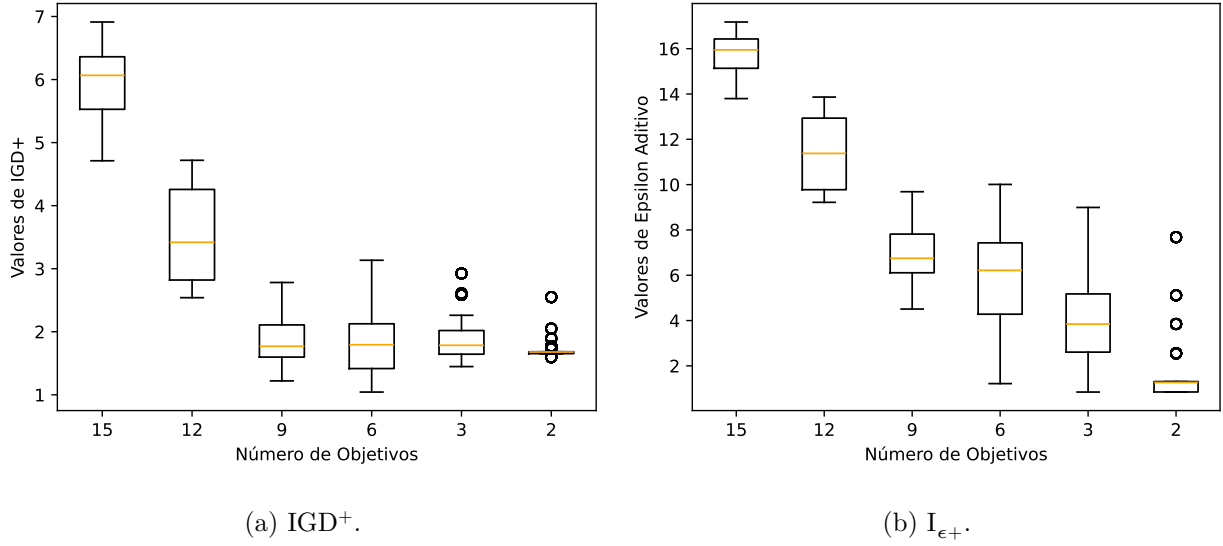


Figura A.33: *Boxplots* dos IQs: DTLZ7 por seleção com $m = 15$.

Agora para o cenário com $m = 15$, a Tabela A.65 dispõe os escores calculados pelo LS e nota-se mais uma vez a presença da aleatoriedade na ordenação, assim como a função com maior índice exibindo a maior magnitude. A Tabela A.66 e a Figura A.33 mostram as melhorias graduais alcançadas nos IQs, principalmente para o $I_{\epsilon+}$. Apesar dos vários *outliers*, destaca-se os resultados da redução para $m = 2$. Além disso, todas as dimensões apresentaram p-valor $< 5\%$, indicando evidências estatísticas para rejeição da hipótese nula.

Tabela A.67: Escores: DTLZ7 com $m = 20$.

f_{19}	f_{17}	f_{11}	f_8	f_{10}	f_{14}	f_7	f_{12}	f_5	f_9
0.2769	0.277	0.2776	0.2811	0.2823	0.2881	0.2882	0.2889	0.2894	0.2902
f_4	f_{18}	f_1	f_{15}	f_2	f_6	f_{13}	f_{16}	f_3	f_{20}
0.2912	0.2913	0.2929	0.293	0.2953	0.2973	0.298	0.2986	0.3002	0.6049

Tabela A.68: Resultados da seleção de objetivos: DTLZ7 com $m = 20$.

m	20	16	12	8	4	2
IGD ⁺	8.7401	6.6251	4.0595	2.1108	2.3697	2.2313
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	24.7895	21.2216	15.6623	9.0803	7.5661	3.1069
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

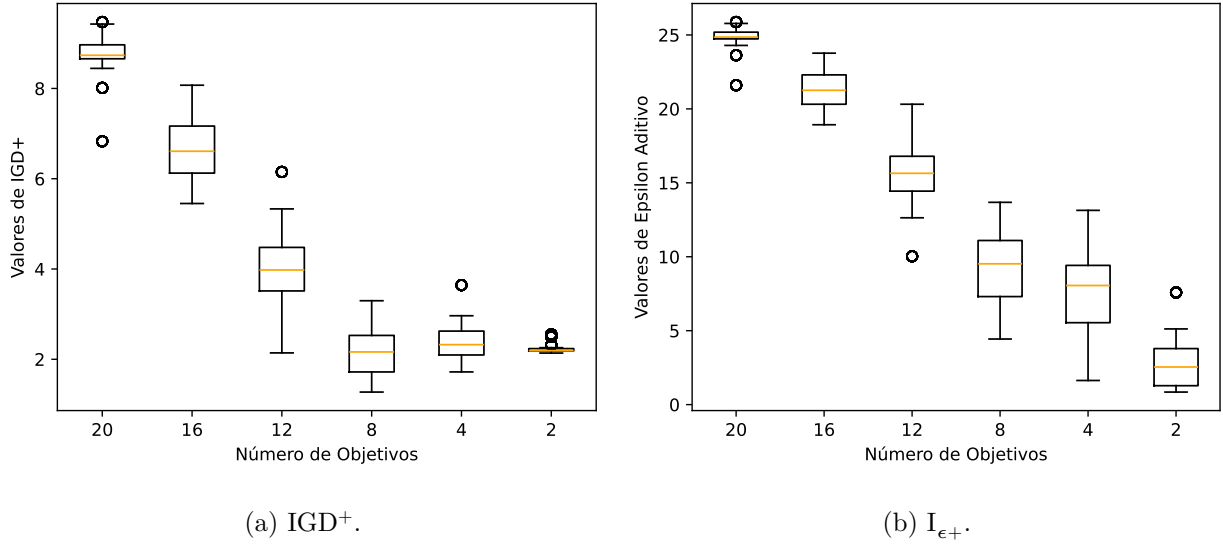


Figura A.34: *Boxplots* dos IQs: DTLZ7 por seleção com $m = 20$.

Seguindo para este MaOP com $m = 20$, os escores dispuseram a ordem das funções semelhante ao observado até aqui, destaca-se também a posição da última função do modelo (f_{20}) na lista, como pode ser visto na Tabela A.67. Outra vez, conforme a Tabela A.68 e a Figura A.34, as médias e medianas do $I_{\epsilon+}$ obtiveram valores gradativamente menores a medida que dimensionalidade era reduzida e o IGD^+ apresentou valores muito próximos nestas mesmas medidas em $m = 8$, $m = 4$ e $m = 2$. Evidencia-se também o ganho de desempenho do caso $m = 2$, bem como o p-valor $< 5\%$ para todas as reduções.

Tabela A.69: Escores: DTLZ7 com $m = 30$.

f_{26}	f_7	f_{10}	f_3	f_9	f_{14}	f_{12}	f_{15}	f_2	f_{23}
0.3142	0.3205	0.3217	0.3224	0.323	0.323	0.3239	0.3257	0.3261	0.3269
f_{20}	f_{11}	f_8	f_{29}	f_{22}	f_{21}	f_{28}	f_{18}	f_{27}	f_5
0.3282	0.3292	0.3303	0.3304	0.3307	0.3307	0.3309	0.331	0.3324	0.3329
f_6	f_{19}	f_{13}	f_{24}	f_1	f_{25}	f_{17}	f_4	f_{16}	f_{30}
0.333	0.3333	0.3338	0.3339	0.3351	0.3357	0.3361	0.3396	0.3448	0.5753

Tabela A.70: Resultados da seleção de objetivos: DTLZ7 com $m = 30$.

m	30	25	20	15	10	5	2
IGD ⁺	13.5028	11.3970	8.9990	6.7977	3.3994	2.9956	3.0352
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	43.3732	40.0334	35.3880	30.1097	18.3729	13.1472	4.9831
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

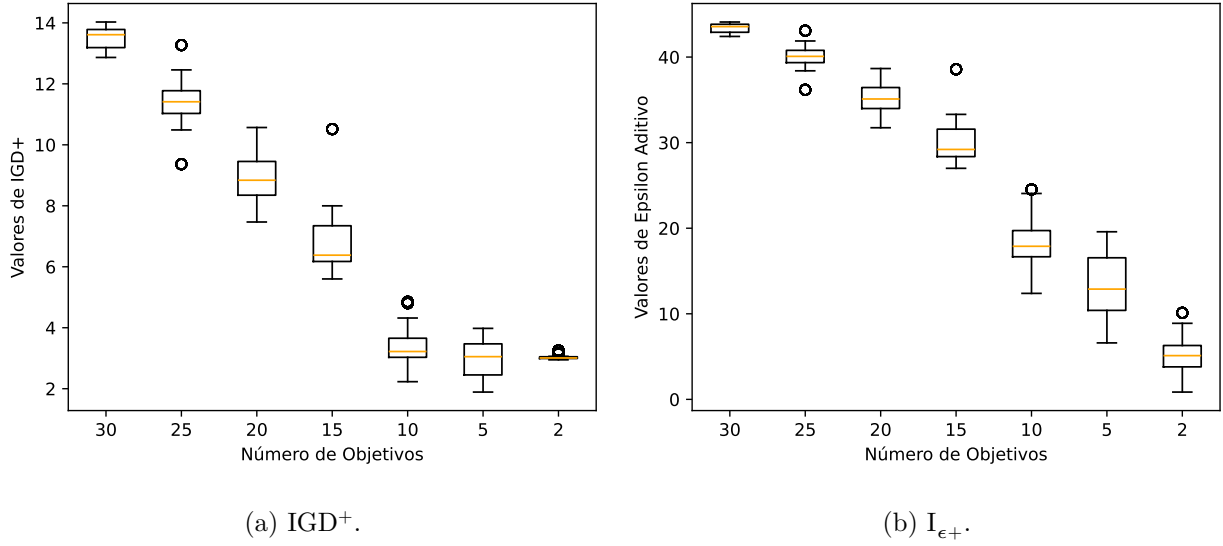


Figura A.35: *Boxplots* dos IQs: DTLZ7 por seleção com $m = 30$.

Por fim, o DTLZ7 com $m = 30$ alcançou os resultados dispostos nas Tabelas A.69 e A.70. Os escores calculados comportaram-se da mesma forma que os cenários anteriores, assim como o último objetivo (f_{30}) que obteve o menor *ranking*. Novamente, os IQs apresentaram ganhos de desempenho graduais concomitantemente com a redução de dimensionalidade, principalmente o I_{ϵ^+} . Analisando a Figura A.35, observa-se que os *boxplots* reforçam esta afirmação. Pela última vez, destaca-se o ganho de desempenho do caso $m = 2$, bem como o $p\text{-valor} < 5\%$ para todas as reduções, sugerindo evidências estatísticas para rejeição da hipótese nula.

Resultados Detalhados da Extração de Atributos: PCA

Este apêndice complementa a Seção 4.3 do Capítulo 4, apresentando os resultados detalhados dos experimentos computacionais realizados com o método de Análise de Componentes Principais (PCA) para redução de objetivos em MaOPs. As tabelas e gráficos aqui contidos correspondem às análises das funções de *benchmark* DTLZ1 a DTLZ7, fornecendo uma visão completa dos resultados obtidos para cada problema de teste.

B.1 DTLZ1

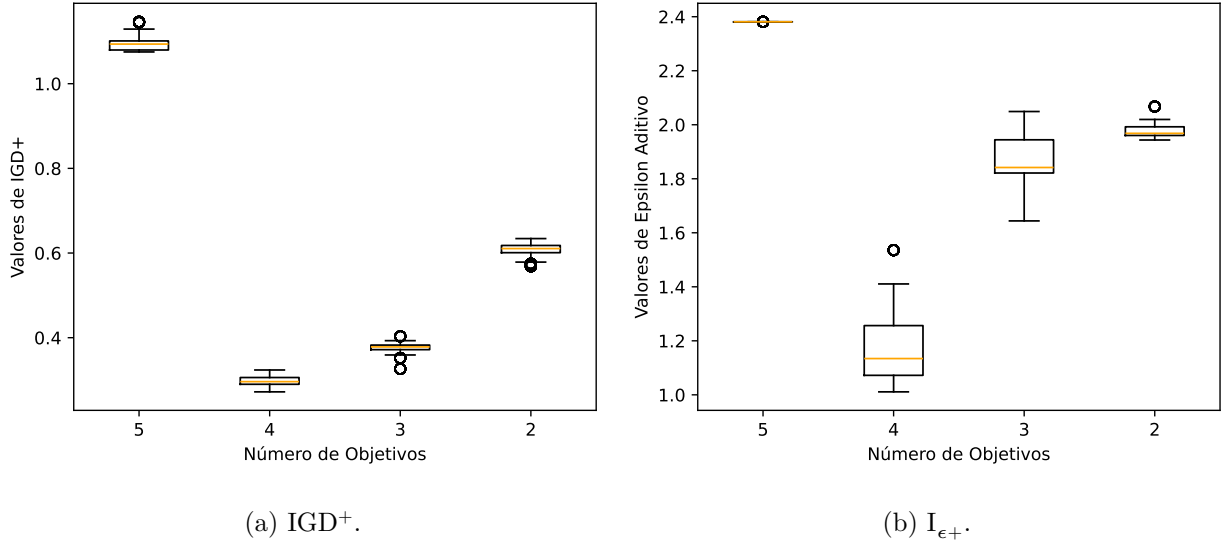
O problema DTLZ1, definido pela Equação 2.3, é utilizado para avaliar a capacidade de um MOEA em convergir para um hiperplano. Os resultados mostram que a PCA é eficaz em melhorar a convergência do NSGA-III, especialmente em cenários de alta dimensionalidade.

Tabela B.1: Porcentagens de variância total explicada: DTLZ1 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
36.1516	24.7583	17.9350	13.5371	7.6177

Tabela B.2: Resultados da extração de objetivos: DTLZ1 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	92.3823	78.8452	60.9102
IGD ⁺	1.0965	0.2985	0.3758	0.6074
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	2.3811	1.1665	1.8752	1.9782
p-valor	1	< 0.001	< 0.001	< 0.001

Figura B.1: *Boxplots* dos IQs: DTLZ1 por extração com $m = 5$.

Este MaOP com $m = 5$, as Tabelas B.1 e B.2 apresentam o decrescimento nos valores do IGD⁺ e do $I_{\epsilon+}$ para todas as abordagens em comparação ao problema original, principalmente para $m = 4$ que soma $\approx 92\%$ de variância total explicada. Verifica-se também que todos os conjuntos de IQs apresentaram o p-valor < 0.05 .

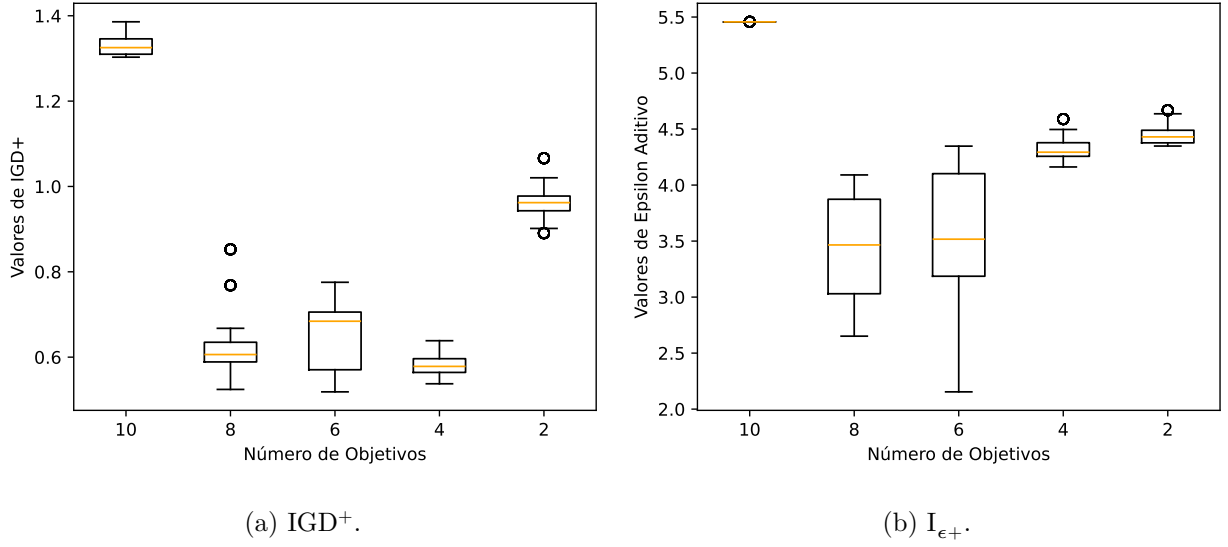
É importante lembrar que há uma perda de informação intrínseca a medida que as dimensões são reduzidas, porém também há uma redução na complexidade do MaOP, facilitando a busca do NSGA-III. Este ponto de equilíbrio pode ser visto nos *boxplots* da Figura B.1, precisamente em $m = 4$, uma vez que há um aumento nas medianas dos IQs assim que a variabilidade total explicada é reduzida para $\approx 79\%$.

Tabela B.3: Porcentagens de variância total explicada: DTLZ1 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
26.9329	14.9889	11.0800	8.8705	8.1169	7.6982	6.9999	6.0425	5.2790	3.9906

Tabela B.4: Resultados da extração de objetivos: DTLZ1 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	90.7304	77.688	61.8729	41.9224
IGD ⁺	1.3304	0.6209	0.6527	0.5812	0.9648
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	5.4567	3.4726	3.4896	4.3210	4.4522
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.2: *Boxplots* dos IQs: DTLZ1 por extração com $m = 10$.

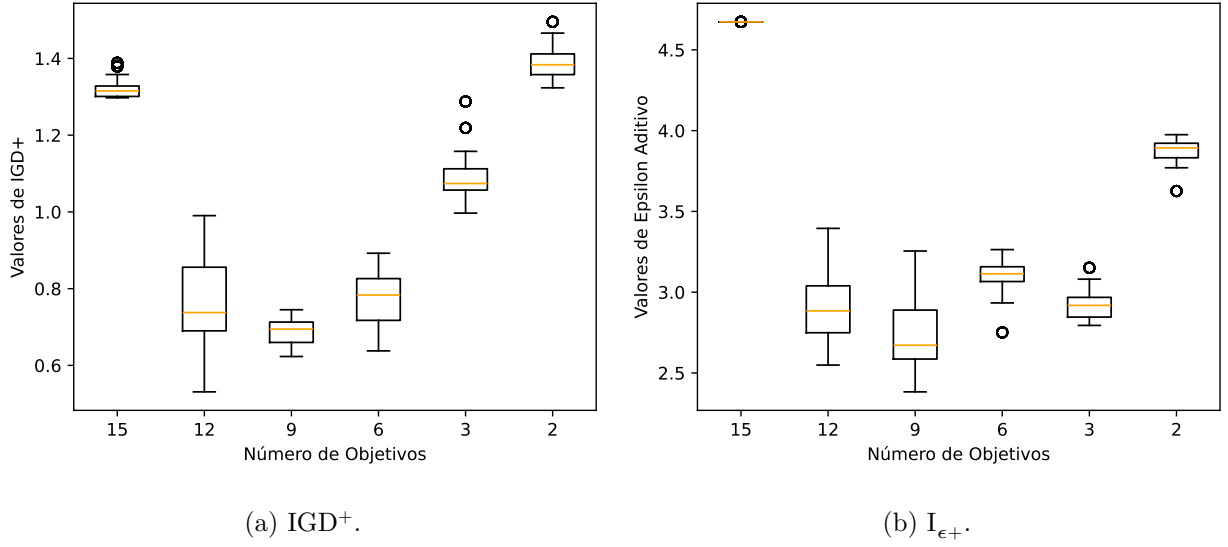
Prosseguindo, agora com dimensão igual a 10, as Tabelas B.3 e B.4 apresentam que todas as abordagens alcançaram ganhos de desempenho estatisticamente positivos em confronto com os IQs originais, inclusive para $m = 2$ com apenas $\approx 42\%$ de variância. Pela Figura B.2, é possível verificar uma oscilação nos resultados do IGD⁺ e também um ponto de equilíbrio entre $m = 8$ e $m = 6$ nas médias e medianas do $I_{\epsilon+}$, ou seja, no intervalo de variâncias totais $\approx [77\%, 90\%]$. Ademais, os *boxplots* ilustram como a dispersão para as duas últimas abordagens foi menor que as outras duas, bem como a original foi a menor entre todas.

Tabela B.5: Porcentagens de variância total explicada: DTLZ1 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
25.5767	11.5681	8.8495	7.4274	6.3528	5.6927	5.0724	4.9644	4.6301	4.3735
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
3.6040	3.4144	3.0257	2.8011	2.6462					

Tabela B.6: Resultados da extração de objetivos: DTLZ1 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	91.527	80.1351	65.4682	45.9943	37.1448
IGD ⁺	1.3243	0.7471	0.6859	0.7720	1.0920	1.3902
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.6730	2.8901	2.7314	3.0888	2.9275	3.8761
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.3: *Boxplots* dos IQs: DTLZ1 por extração com $m = 15$.

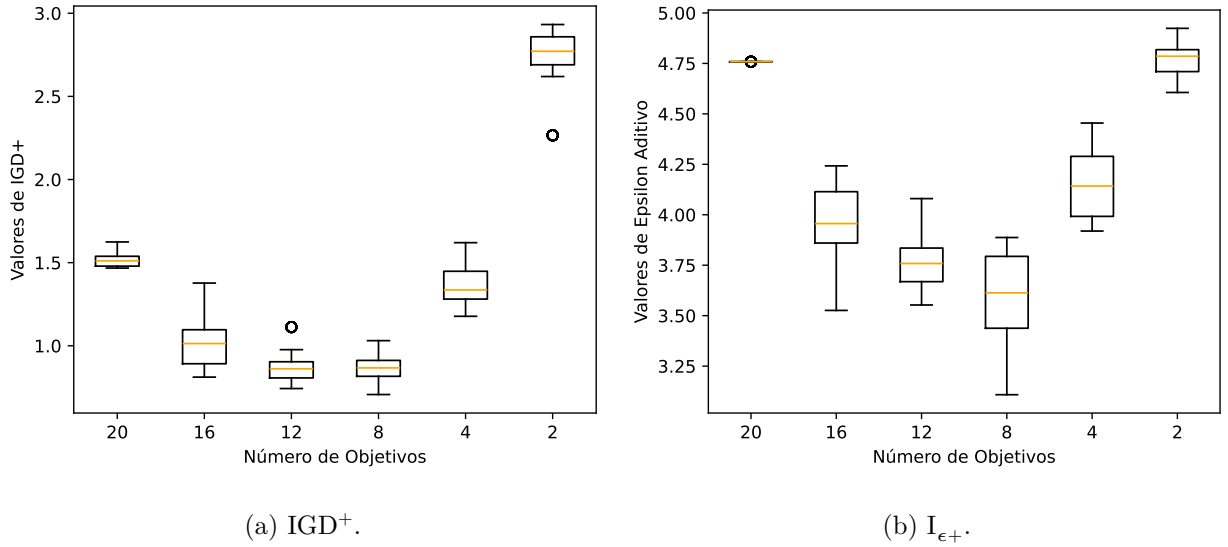
As Tabelas B.5 e B.6 dispõem os resultados para o DTLZ1 com $m = 15$, nota-se que quase todas as reduções obtiveram ganhos de desempenho, com exceção de $m = 2$ pelo IGD⁺, apesar de também indicar evidências estatísticas para rejeição da hipótese nula. Além disso, a abordagem para $m = 9$ apresentou os melhores resultados de IQs entre os demais, inclusive sendo o ponto de equilíbrio neste caso, com uma variância total de $\approx 80\%$. A respeito dos *boxplots* na Figura B.3, o comportamento das distribuições é semelhante ao cenário anterior, onde a original é quase nula.

Tabela B.7: Porcentagens de variância total explicada: DTLZ1 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
27.9259	12.4049	7.7235	5.8638	5.2673	4.7865	4.5448	4.3532	4.1977	3.7861
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
3.4188	3.0037	2.9233	2.2997	1.9441	1.9315	1.6960	0.9878	0.8175	0.1228

Tabela B.8: Resultados da extração de objetivos: DTLZ1 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	96.3748	87.2762	72.8699	53.9181	40.3308
IGD ⁺	1.5140	1.0255	0.8684	0.8696	1.3604	2.7520
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.7584	3.9514	3.7552	3.6012	4.1500	4.7660
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	0.7916

Figura B.4: *Boxplots* dos IQs: DTLZ1 por extração com $m = 20$.

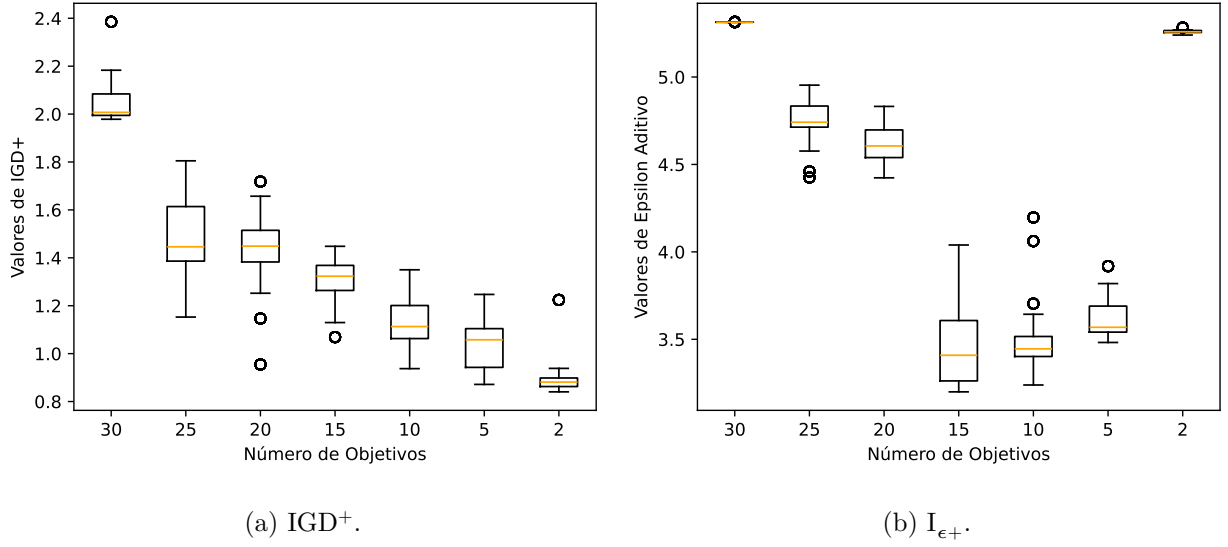
Seguindo para o cenário deste MaOP com dimensão igual a 20, as Tabelas B.7 e B.8 apresentam que quase todas as abordagens obtiveram ganhos de desempenho, exceto em $m = 2$. Especificamente, percebe-se um ponto de equilíbrio entre $m = 12$ e $m = 8$ que alcançaram os melhores valores de IQs, com variâncias totais no intervalo $\approx [72\%, 87\%]$. O problema $m = 2$ teve ambos os IQs ainda maiores que o original, mostrando a alta perda de informação para uma variância total explicada de $\approx 40\%$. Estas análises são também observáveis na Figura B.4.

Tabela B.9: Porcentagens de variância total explicada: DTLZ1 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
27.7287	11.8154	6.6513	5.5165	4.9913	4.1439	3.6157	3.3821	3.0302	2.8565
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
2.7150	2.6025	2.4001	2.2296	2.0356	1.8473	1.7605	1.6053	1.3692	1.3021
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
1.2583	1.1665	1.0363	0.9087	0.6580	0.5126	0.4737	0.2451	0.1077	0.0327

Tabela B.10: Resultados da extração de objetivos: DTLZ1 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	98.6266	93.5988	85.7144	73.7316	56.7032	39.5441
IGD ⁺	2.0548	1.4731	1.4358	1.3087	1.1284	1.0342	0.9023
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	5.3136	4.7426	4.6199	3.4471	3.5145	3.6123	5.2589
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.5: *Boxplots* dos IQs: DTLZ1 por extração com $m = 30$.

Por último, os resultados do DTLZ1 com $m = 30$ estão dispostos nas Tabelas B.9 e B.10, analisando-as verifica-se que as médias do IGD⁺ diminuíram gradualmente a medida que a dimensão era reduzida. Porém, conforme pode ser visto também na Figura B.5, o $I_{\epsilon+}$ apresentou um ponto de equilíbrio em $m = 15$ com variância total explicada de $\approx 85\%$. Este fato acarreta na evidente diferença entre os dois IQs na redução para $m = 2$, que obteve o menor valor de IGD⁺ com variância de $\approx 39\%$, porém não o menor valor de $I_{\epsilon+}$. Ademais, os p-valores de todas as distribuições foram abaixo de 5%.

B.2 DTLZ2

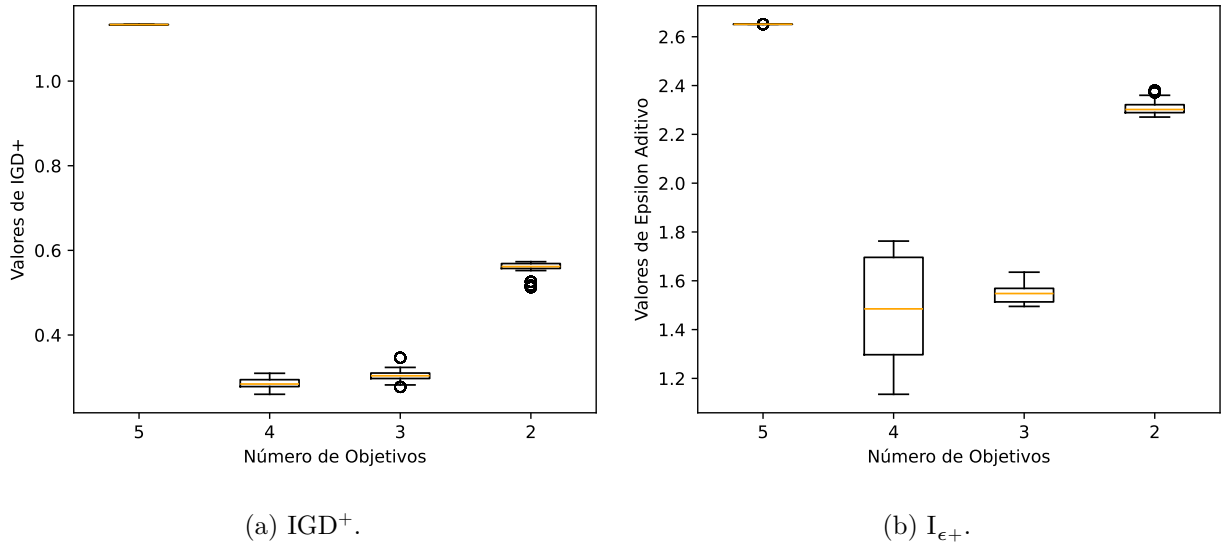
O problema DTLZ2, conforme a Equação 2.5, é projetado para testar a capacidade de um MOEA em lidar com um grande número de objetivos. A PCA mostrou-se consistente em melhorar o desempenho do MOEA em todos os cenários testados.

Tabela B.11: Porcentagens de variância total explicada: DTLZ2 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
40.1970	25.1863	17.4359	12.7052	4.4753

Tabela B.12: Resultados da extração de objetivos: DTLZ2 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	95.5244	82.8192	65.3833
IGD ⁺	1.1334	0.2846	0.3047	0.5575
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	2.6512	1.4839	1.5506	2.3081
p-valor	1	< 0.001	< 0.001	< 0.001

Figura B.6: *Boxplots* dos IQs: DTLZ2 por extração com $m = 5$.

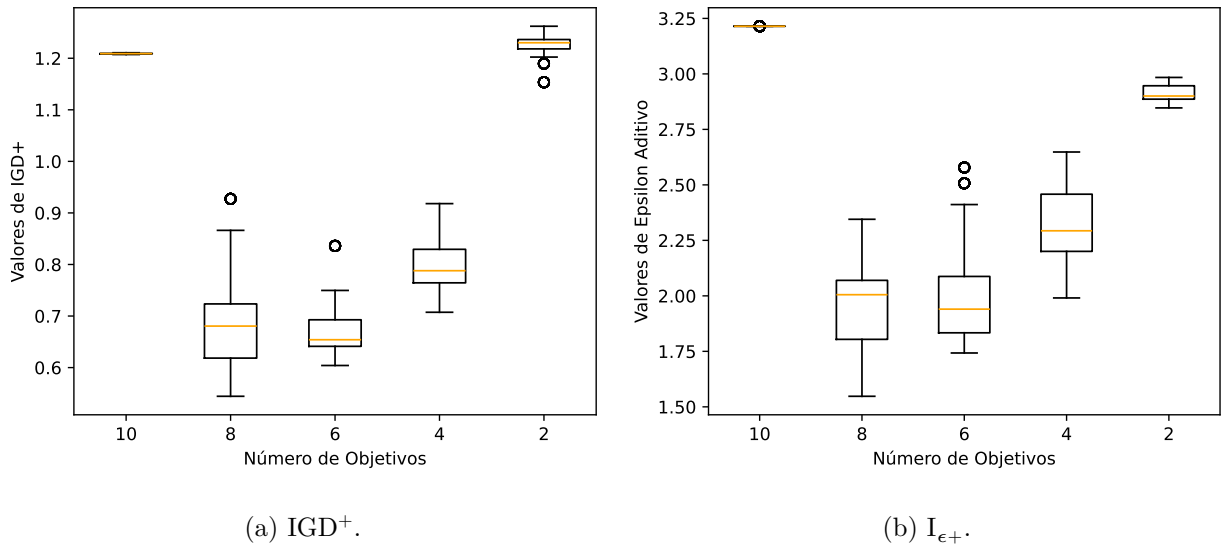
Analisando as Tabelas B.11 e B.12, verifica-se o decrescimento nos valores dos IQs para todas as abordagens em relação ao original, principalmente em $m = 4$ com variância total explicada de $\approx 95\%$. Ademais, todos os conjuntos de IQs evidenciaram estatísticas para rejeição da hipótese nula. O ponto de equilíbrio em $m = 4$ pode ser visto na Figura B.6, uma vez que há um aumento nas medianas dos IQs de $m = 3$ e $m = 2$. Além disso, esta redução para quatro objetivos dispõe a distribuição dos 21 $I_{\epsilon+}$ mais espaçada que os demais *boxplots*.

Tabela B.13: Porcentagens de variância total explicada: DTLZ2 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
33.8577	15.6006	10.9688	8.8338	6.9326	6.0544	5.8330	5.1636	4.4112	2.3439

Tabela B.14: Resultados da extração de objetivos: DTLZ2 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	93.2445	82.2479	69.2609	49.4583
IGD ⁺	1.2090	0.6876	0.6709	0.7952	1.2258
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.2146	1.9683	2.0070	2.3182	2.9089
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.7: *Boxplots* dos IQs: DTLZ2 por extração com $m = 10$.

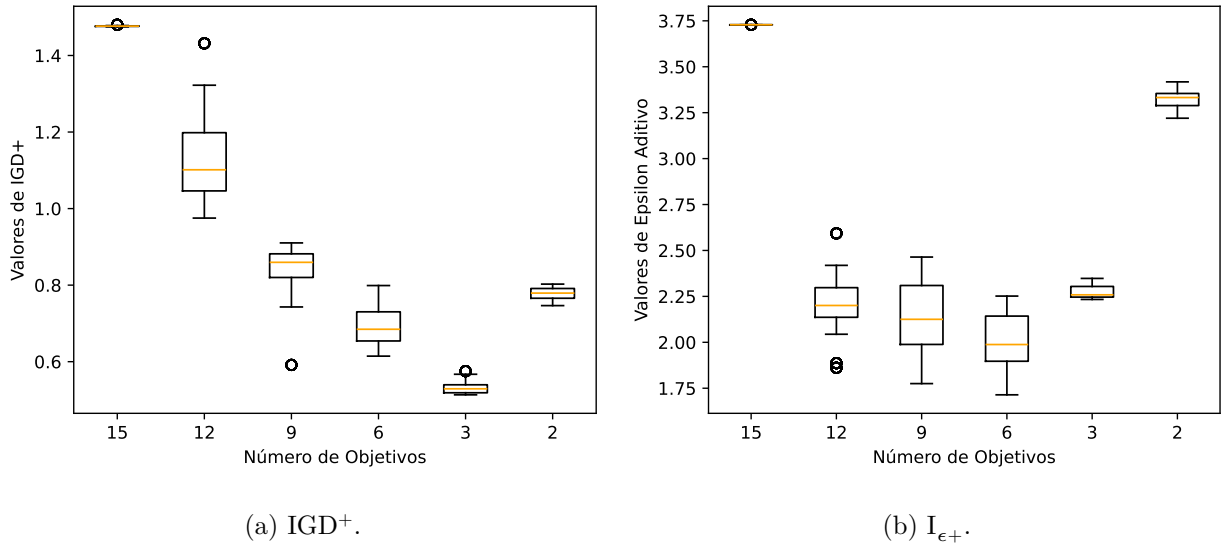
Seguindo para a análise deste MaOP com $m = 10$, nota-se nas Tabelas B.13 e B.14 que houveram ganhos de desempenho acompanhados dos p-valores $< 5\%$ em todas as reduções, segundo os IQs calculados, com exceção em $m = 2$. Por meio da Figura B.7, percebe-se o ponto de equilíbrio na abordagem para $m = 6$ que possui a variância explicada total de $\approx 82\%$. Ademais, o problema original e com dois objetivos apresentaram baixas distribuições em ambos os IQs.

Tabela B.15: Porcentagens de variância total explicada: DTLZ2 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
32.2095	13.3949	9.4570	7.3657	5.5016	5.0527	4.2223	4.1448	3.6326	3.4078
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
3.1635	2.8829	2.3235	1.7120	1.5284					

Tabela B.16: Resultados da extração de objetivos: DTLZ2 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	94.4353	84.9811	72.9814	55.0614	45.6044
IGD ⁺	1.4767	1.1339	0.8391	0.6909	0.5338	0.7778
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.7284	2.2123	2.1193	1.9951	2.2771	3.3237
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.8: *Boxplots* dos IQs: DTLZ2 por extração com $m = 15$.

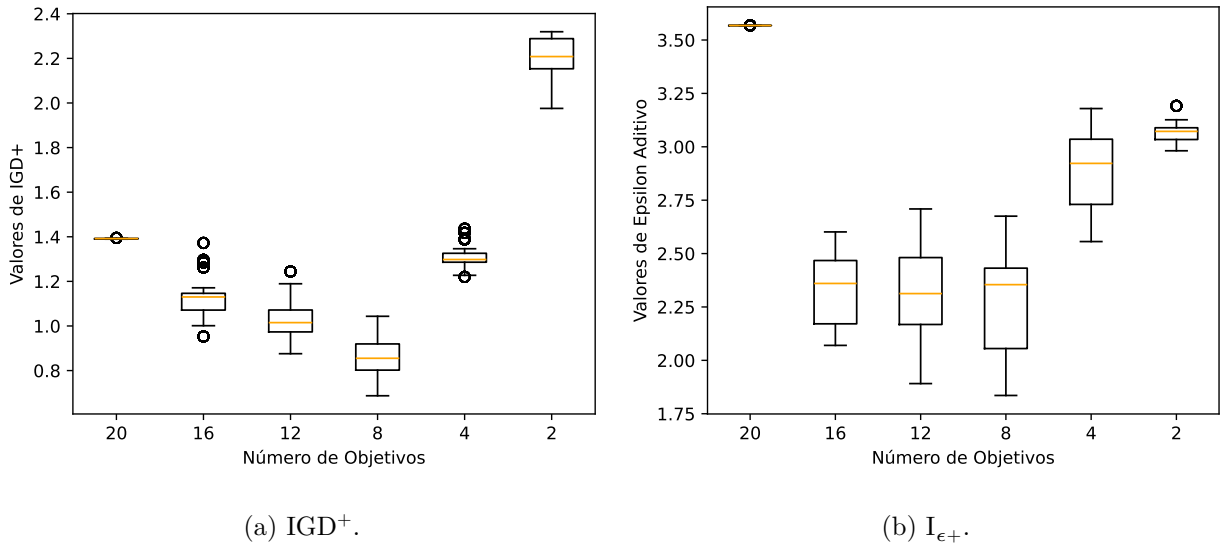
Avaliando os resultados deste DTLZ com dimensão igual a 15 pelas Tabelas B.15 e B.16, nota-se que todas as médias dos IQs reduzidos estão abaixo da média original e com p-valor menor que 5%. Além disto, a menor média de IGD⁺ entre as abordagens é para $m = 3$, porém a menor média de I_ε⁺ é para $m = 6$. Sendo assim, os pontos de equilíbrio também estão divididos entre os IQs pelos mesmos casos, onde o primeiro ocorre em $m = 3$ e o segundo em $m = 6$, resultando no intervalo de variância $\approx [55\%, 73\%]$. Os *boxplots* da Figura B.8 também mostram as dispersões menores nas reduções mais incisivas, além de sua quase inexistência no problema original.

Tabela B.17: Porcentagens de variância total explicada: DTLZ2 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
30.8962	13.1122	8.4518	6.4510	5.2530	4.5351	3.9716	3.5961	3.1856	2.9700
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
2.7600	2.5004	2.2420	2.0419	1.7673	1.6944	1.4389	1.2585	1.0873	0.7856

Tabela B.18: Resultados da extração de objetivos: DTLZ2 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	95.4286	87.683	76.267	58.9112	44.0084
IGD ⁺	1.3916	1.1322	1.0346	0.8653	1.3104	2.1997
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.5679	2.3311	2.3165	2.2726	2.8816	3.0668
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.9: *Boxplots* dos IQs: DTLZ2 por extração com $m = 20$.

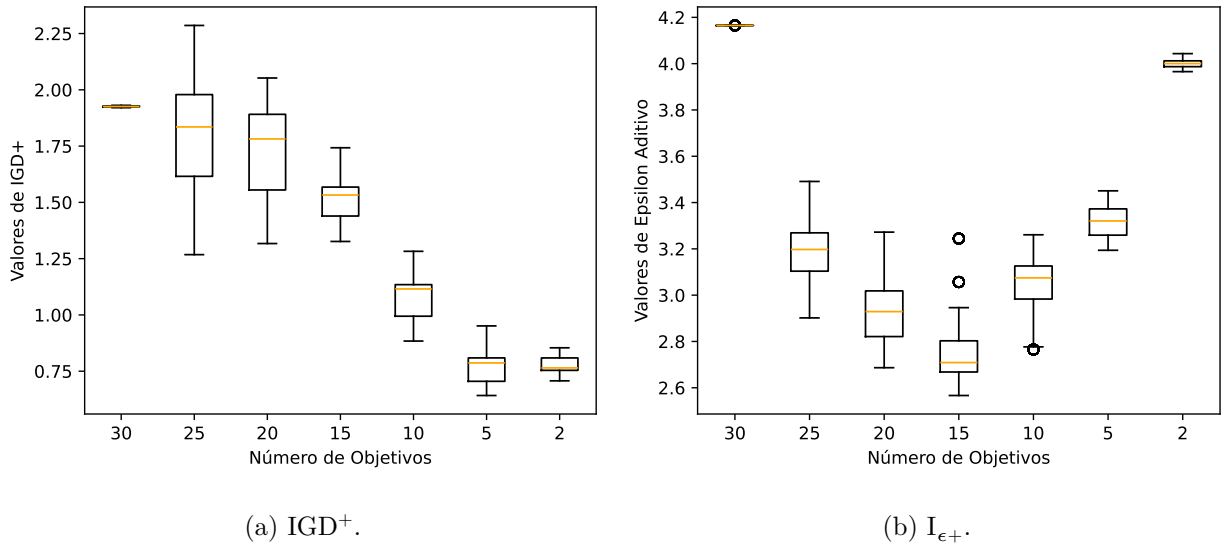
Seguindo para o DTLZ2 com $m = 20$, as Tabelas B.17 e B.18 mostram ganhos de desempenho em todos os IQs, exceto para o IGD⁺ em $m = 2$, com evidências estatísticas para rejeição da hipótese nula. O ponto de equilíbrio é o mesmo para ambos os indicadores: em $m = 8$ com variância $\approx 76\%$ as médias são as mais baixas, subindo a partir de $m = 4$. A Figura B.9 ilustra esse comportamento e, destacando as dispersões nos *boxplots*, desta vez os dois últimos casos não apresentaram as menores distribuições, novamente o original mostrou variação quase nula.

Tabela B.19: Porcentagens de variância total explicada: DTLZ2 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
30.8201	12.2872	7.0012	5.0483	3.9064	3.6296	3.2010	2.9326	2.7082	2.5251
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
2.3882	2.2075	2.1029	1.9499	1.8927	1.7718	1.7484	1.6227	1.4028	1.2726
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
1.1686	1.0849	1.0560	0.8976	0.8253	0.7032	0.6027	0.5287	0.4673	0.2447

Tabela B.20: Resultados da extração de objetivos: DTLZ2 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	97.4516	92.4192	84.6009	74.0597	59.0632	43.1073
IGD ⁺	1.9260	1.8136	1.7365	1.5157	1.0816	0.7665	0.7752
p-valor	1	0.1866	0.0036	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.1649	3.1811	2.9248	2.7667	3.0326	3.3217	4.0002
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.10: *Boxplots* dos IQs: DTLZ2 por extração com $m = 30$.

Pelos resultados do cenário $m = 30$ nas Tabelas B.19 e B.20, observa-se que à medida que a dimensão é reduzida, segue a atenuação do IGD⁺ até $m = 5$, porém o menor valor de $I_{\epsilon+}$ se dá em $m = 15$. E, pela primeira vez, uma das abordagens apresentou um p-valor acima de 5%: $m = 25$ pelo IGD⁺, mas as outras resultaram no contrário. Estas análises também estão representadas na Figura B.10, visto que os pontos de equilíbrio se dão em $m = 5$ para o IGD⁺ e em $m = 15$ para o $I_{\epsilon+}$, resultando no intervalo de variância $\approx [59\%, 84\%]$. Além

disso, percebe-se que a dispersão em ambos as métricas está menor para os últimos casos e quase inexistente para o original.

B.3 DTLZ3

O problema DTLZ3, definido pela Equação 2.6, compartilha as mesmas funções objetivo do DTLZ2, mas com um funcional modificado para testar a convergência para a PF global. A PCA demonstrou eficácia em melhorar a convergência do NSGA-III, especialmente em cenários de menor dimensionalidade.

Tabela B.21: Porcentagens de variância total explicada: DTLZ3 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
37.1841	22.2758	16.1345	13.1902	11.2150

Tabela B.22: Resultados da extração de objetivos: DTLZ3 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	88.7846	75.5944	59.4599
IGD ⁺	1.6041	0.4546	1.1598	1.2811
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	2.8362	1.6440	2.1266	2.2115
p-valor	1	< 0.001	< 0.001	< 0.001

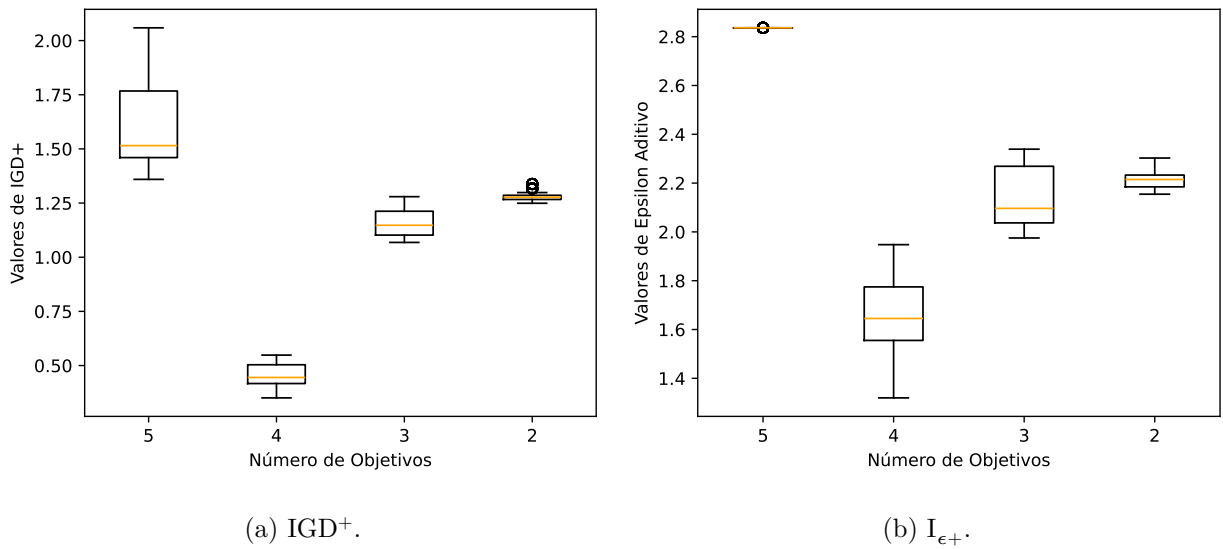


Figura B.11: *Boxplots* dos IQs: DTLZ3 por extração com $m = 5$.

Avaliando as Tabelas B.21 e B.22 do DTLZ3 com $m = 5$, nota-se o decrescimento nos valores do IGD^+ e do $I_{\epsilon+}$ para todas as abordagens em confronto ao problema original, com destaque para $m = 4$ que soma $\approx 89\%$ de variância total explicada. Verifica-se também que todos os conjuntos dos indicadores apresentaram o p-valor < 0.05 . E na Figura B.11, a respeito dos *boxplots* em ambos os IQs, a dispersão se mostra invertida entre o caso original e os reduzidos, isto é, enquanto o primeiro possui distribuição menor no IGD^+ , o segundo é no $I_{\epsilon+}$, e vice-versa.

Tabela B.23: Porcentagens de variância total explicada: DTLZ3 com $m = 10$.

$\lambda^1_{\%}$	$\lambda^2_{\%}$	$\lambda^3_{\%}$	$\lambda^4_{\%}$	$\lambda^5_{\%}$	$\lambda^6_{\%}$	$\lambda^7_{\%}$	$\lambda^8_{\%}$	$\lambda^9_{\%}$	$\lambda^{10}_{\%}$
33.2298	14.2347	10.3692	8.4389	7.1245	6.4620	6.1688	5.1369	4.9706	3.8642

Tabela B.24: Resultados da extração de objetivos: DTLZ3 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	91.1648	79.8591	66.2726	47.4645
IGD^+	2.0301	0.7213	0.6664	1.1449	1.6359
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	2.5643	1.6246	1.5945	2.2974	2.6363
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

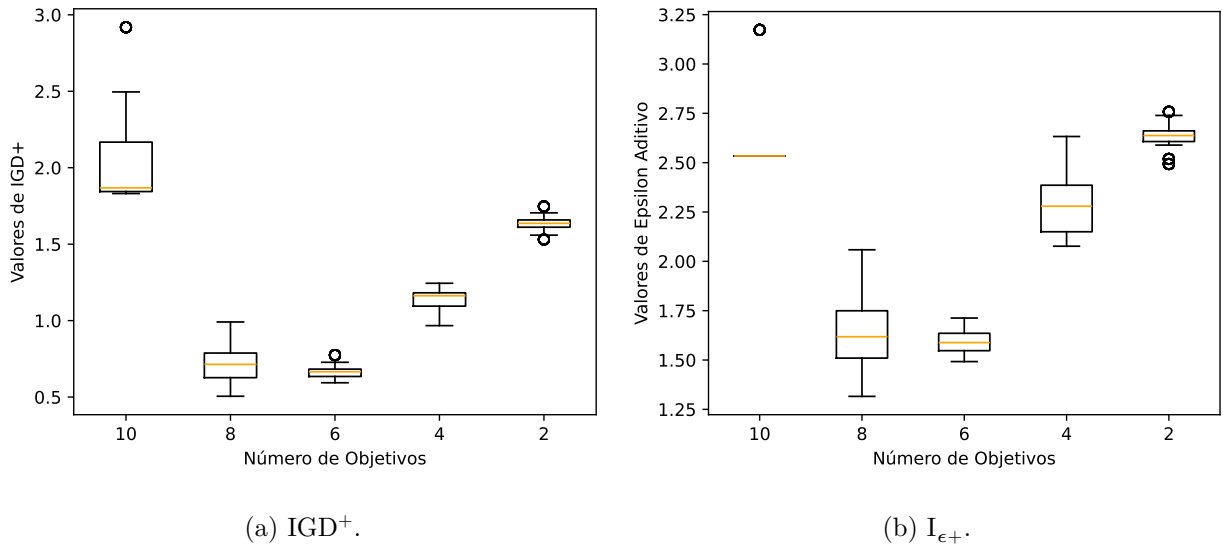


Figura B.12: *Boxplots* dos IQs: DTLZ3 por extração com $m = 10$.

Prosseguindo neste mesmo MaOP com $m = 10$, as Tabelas B.23 e B.24 mostram que todas as abordagens alcançaram ganhos de desempenho em confronto com os IQs originais,

exceto para $m = 2$ pelo $I_{\epsilon+}$, e todos os casos tiveram p-valor $< 5\%$. Analisando a Figura B.12, é possível verificar um ponto de equilíbrio em $m = 6$ nas médias e medianas dos IQs, com $\approx 80\%$ de variância total explicada. Ademais, os *boxplots* ilustram o mesmo comportamento da dispersão que no cenário anterior, mas desta vez percebe-se um *outlier* no DTLZ3 original.

Tabela B.25: Porcentagens de variância total explicada: DTLZ3 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
30.9756	13.2220	8.6805	7.1748	5.3835	4.7547	4.3512	4.1052	3.8166	3.5918
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
3.3910	3.3604	2.6482	2.5082	2.0357					

Tabela B.26: Resultados da extração de objetivos: DTLZ3 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	92.8073	82.4641	70.1911	52.8781	44.1976
IGD ⁺	2.1518	1.0861	0.9677	0.8814	0.7459	0.7316
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.5192	3.5659	3.1998	2.8122	3.7911	4.6823
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

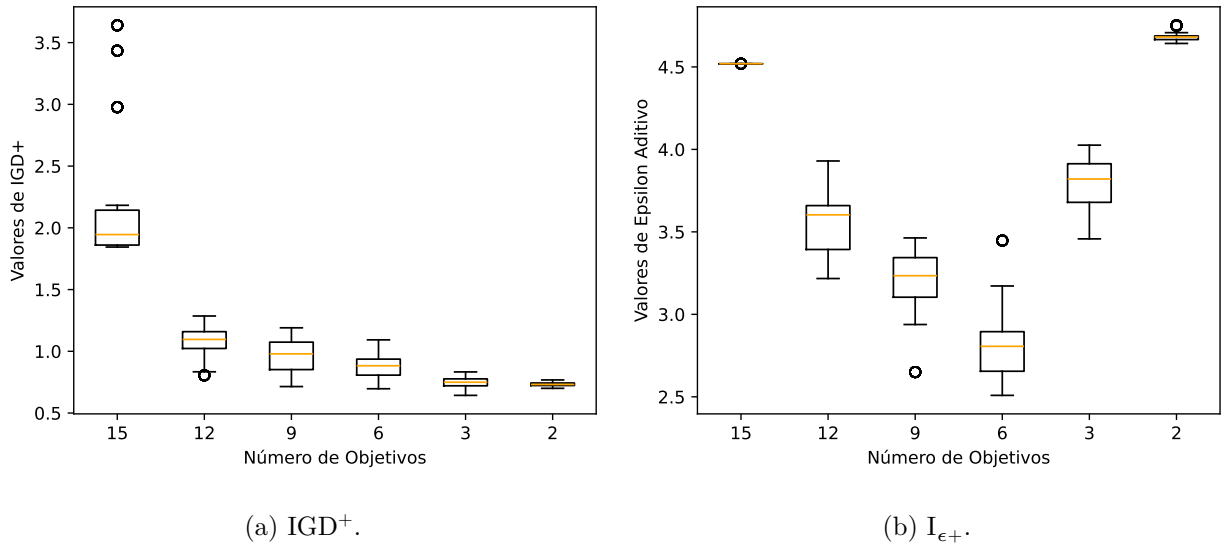


Figura B.13: *Boxplots* dos IQs: DTLZ3 por extração com $m = 15$.

As Tabelas B.25 e B.26 dispõem os resultados para o DTLZ3 com $m = 15$, percebe-se que as reduções obtiveram um decrescimento progressivo para o IGD⁺, porém, a partir de

$m = 6$ o $I_{\epsilon+}$ cresce até uma perda de desempenho em $m = 2$. Sendo assim, a abordagem para $m = 6$ aponta ser o ponto de equilíbrio neste cenário, com uma variância total de $\approx 70\%$. Além disso, todos indicaram evidências estatísticas para rejeição da hipótese nula. A respeito dos *boxplots* na Figura B.13, o comportamento das dispersões é semelhante aos dois cenários anteriores, somando a presença de três *outliers* para o IGD^+ do original.

Tabela B.27: Porcentagens de variância total explicada: DTLZ3 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
31.6026	13.5806	7.9381	6.2874	5.0157	4.4284	4.2000	3.6969	3.2709	3.1861
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
3.1804	2.8249	2.2784	2.0845	1.8180	1.3775	1.2509	0.8715	0.6737	0.4323

Tabela B.28: Resultados da extração de objetivos: DTLZ3 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	96.7704	89.212	76.7497	59.4087	45.1832
IGD^+	2.3202	1.2827	1.1978	1.1581	1.8701	3.2767
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.5275	3.3809	3.4604	2.7179	3.1264	4.4052
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

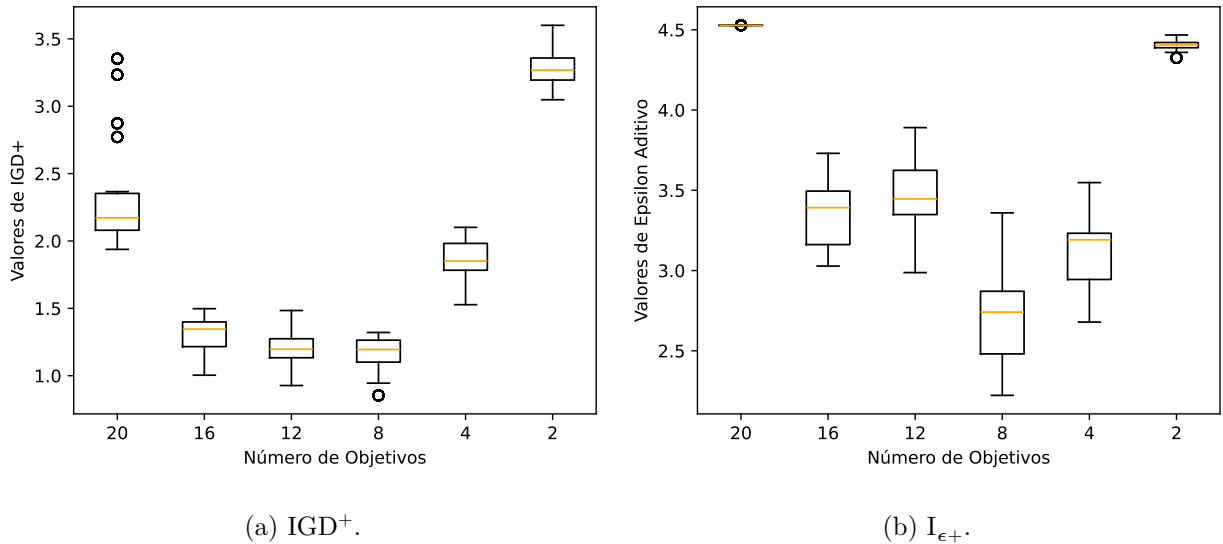


Figura B.14: *Boxplots* dos IQs: DTLZ3 por extração com $m = 20$.

Seguindo para o cenário deste MaOP com dimensão igual a 20, as Tabelas B.27 e B.28 apresentam uma inversão do caso anterior, do ponto de vista que a última redução obteve

ganho de desempenho no $I_{\epsilon+}$ e não no IGD^+ . É possível verificar também um ponto de equilíbrio em $m = 8$, que alcançou os melhores valores de IQs com variância total explicada de $\approx 77\%$], e todas as abordagens obtiveram p-valor $< 5\%$. Analisando os *boxplots* de ambos os IQs na Figura B.14, nota-se que o padrão da observado nos cenários anteriores não se manifestou para $m = 2$. Além disso, apesar da mediana desta redução estar maior que a do problema original, este último ainda apresentou uma média inferior devido aos quatro *outliers*.

Tabela B.29: Porcentagens de variância total explicada: DTLZ3 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
29.1771	13.2736	7.2771	4.8480	4.2707	3.7146	3.0971	2.9015	2.7065	2.3726
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
2.2036	2.1889	2.1394	1.9810	1.8710	1.7906	1.6978	1.6347	1.5600	1.4468
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
1.2955	1.2433	1.1228	0.9978	0.8192	0.7552	0.5802	0.4690	0.3035	0.2596

Tabela B.30: Resultados da extração de objetivos: DTLZ3 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	97.6312	92.1526	84.0227	73.6388	58.8465	42.4507
IGD^+	3.2168	1.8036	1.6354	1.3362	1.1549	0.9578	1.3438
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.6415	3.4012	3.3251	3.4977	3.4436	3.8645	4.4414
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

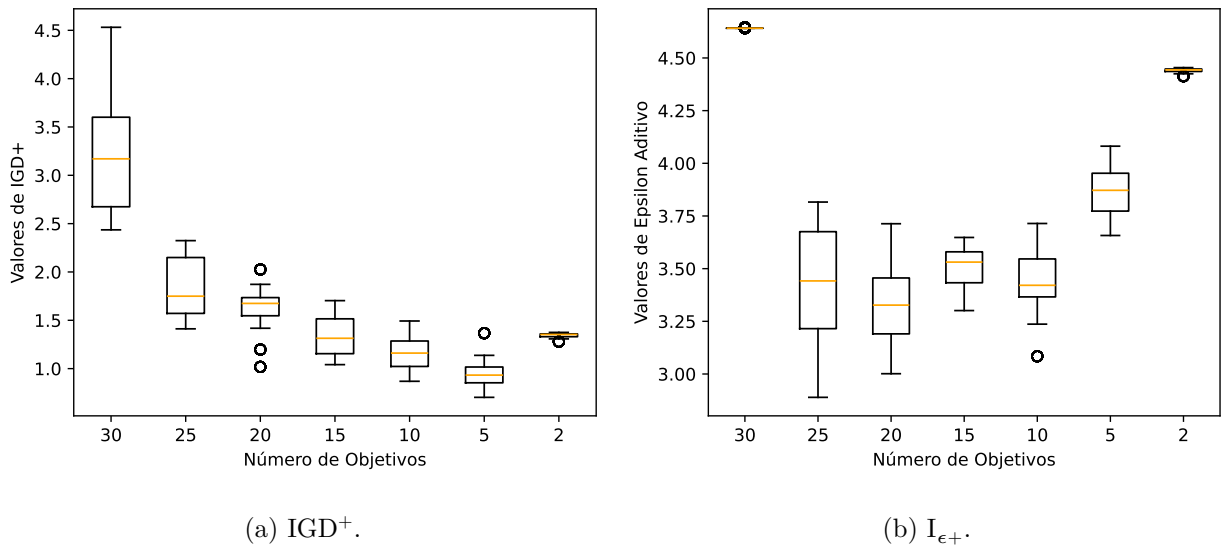


Figura B.15: *Boxplots* dos IQs: DTLZ3 por extração com $m = 30$.

Por último, os resultados do DTLZ3 com $m = 30$ estão dispostos nas Tabelas B.29 e B.30, analisando-as verifica-se que as médias do IGD^+ diminuíram gradualmente até subir na redução $m = 2$. A respeito do $I_{\epsilon+}$, conforme pode ser visto também na Figura B.15, percebe-se um comportamento diferente, onde há uma oscilação que termina com o problema bi-objetivo sendo o pior entre as demais abordagens e $m = 20$ com o menor valor para este IQ. Assim, o ponto de equilíbrio se torna mais difícil de ser definido, apesar disso, destaca-se a redução $m = 10$ com variância total explicada de $\approx 74\%$. Adicionalmente, todas as reduções obtiveram ganhos de desempenho com evidências estatísticas para rejeição da hipótese nula.

B.4 DTLZ4

O problema DTLZ4, conforme a Equação 2.7, é uma variação do DTLZ2 com um parâmetro α modificado, resultando em um conjunto denso de soluções. A PCA mostrou-se eficaz em capturar as direções mais relevantes, melhorando a convergência e a diversidade das soluções.

Tabela B.31: Porcentagens de variância total explicada: DTLZ4 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
31.2004	20.6877	20.0357	19.7939	8.2820

Tabela B.32: Resultados da extração de objetivos: DTLZ4 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	91.7177	71.9238	51.8881
IGD^+	1.4881	0.6691	0.6610	0.9330
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	8.7674	4.5325	3.8712	4.2952
p-valor	1	< 0.001	< 0.001	< 0.001

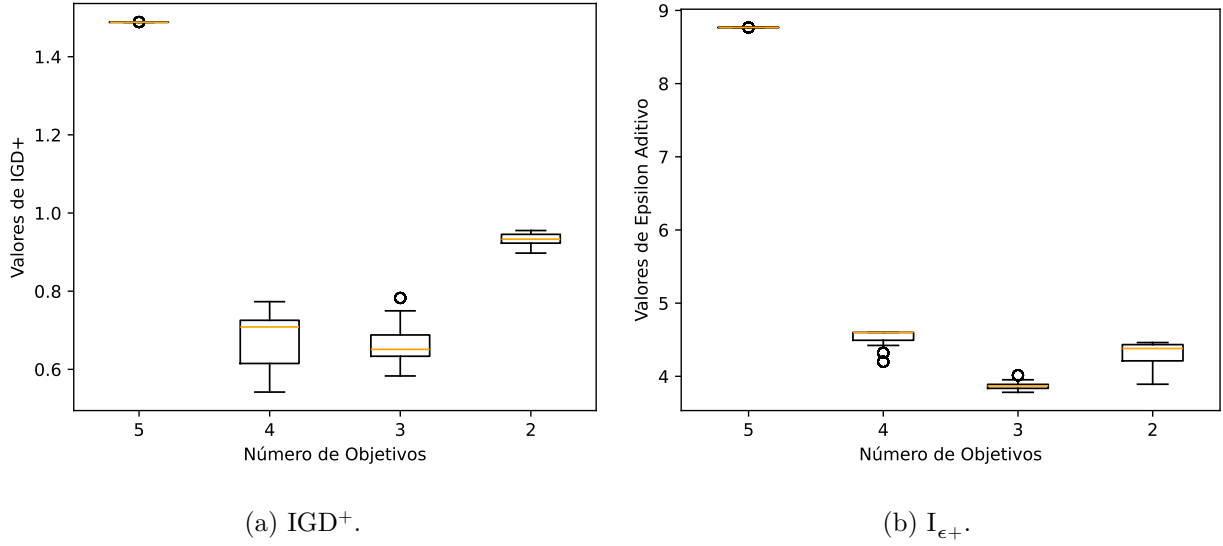


Figura B.16: *Boxplots* dos IQs: DTLZ4 por extração com $m = 5$.

Analisando as Tabelas B.31 e B.32, nota-se o decrescimento nos valores dos IQs para todas as abordagens em relação ao original, com todos os conjuntos dece IQs evidenciando estatísticas para rejeição da hipótese nula. Desta vez, os menores em $m = 3$ com variância total explicada de $\approx 72\%$, diferentemente do visto nos DTLZ anteriores neste mesmo cenário. O ponto de equilíbrio em $m = 3$ pode ser visto na Figura B.16, uma vez que há um aumento nas medianas dos IQs de $m = 2$. Além disso, a distribuição dos *boxplots* é praticamente nula para o problema original.

Tabela B.33: Porcentagens de variância total explicada: DTLZ4 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
16.4106	10.6843	10.5092	10.3326	10.1292	9.9899	9.8304	9.5692	9.4222	3.1219

Tabela B.34: Resultados da extração de objetivos: DTLZ4 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	87.4554	68.0558	47.9367	27.0949
IGD^+	1.4275	1.0824	1.0858	1.0556	1.3203
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	7.1592	4.1115	4.0020	3.8423	4.8025
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

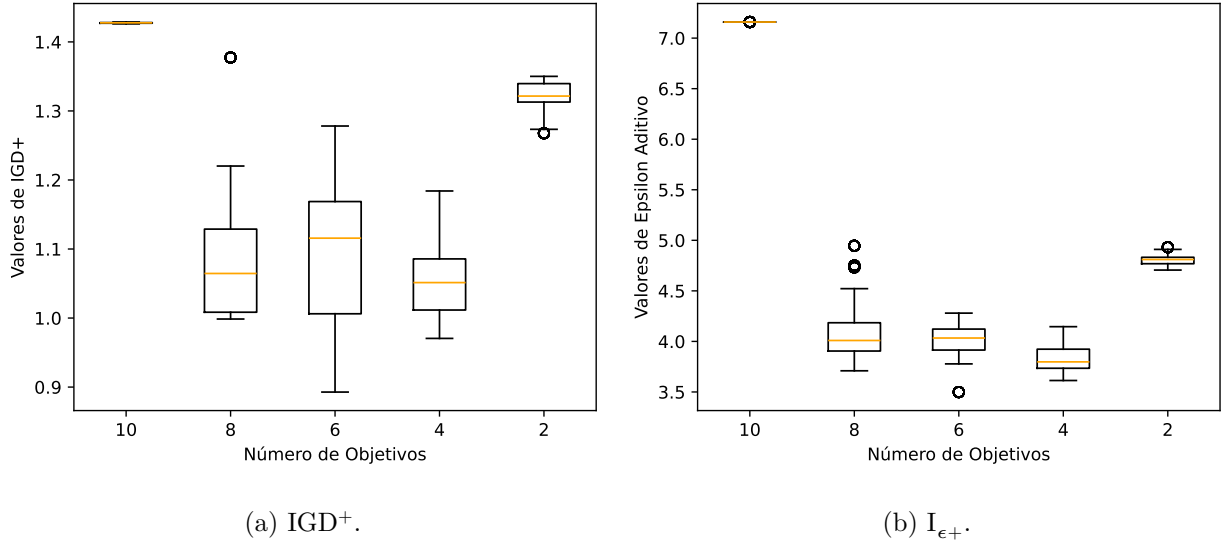


Figura B.17: *Boxplots* dos IQs: DTLZ4 por extração com $m = 10$.

Seguindo para a análise deste MaOP com $m = 10$, verifica-se nas Tabelas B.33 e B.34 que houveram ganhos de desempenho com p-valores $< 5\%$ em todas as reduções, segundo os IQs calculados, inclusive para a abordagem $m = 2$ com variância de apenas $\approx 27\%$. Por meio da Figura B.17, percebe-se o ponto de equilíbrio na abordagem para $m = 4$ que possui a variância explicada total de $\approx 48\%$. Ademais, as dispersões dos *boxplots* do IGD^+ foram mais expressivas que as do I_{ϵ^+} , exceto em $m = 4$ e no original que é quase nula.

Tabela B.35: Porcentagens de variância total explicada: DTLZ4 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
11.0798	7.1793	7.0154	6.9886	6.8610	6.8520	6.7821	6.7109	6.6666	6.6276
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
6.5233	6.4809	6.4364	6.2366	1.5588					

Tabela B.36: Resultados da extração de objetivos: DTLZ4 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	85.7675	66.1357	45.9761	25.2745	18.2591
IGD^+	1.5427	1.4624	1.2026	1.1116	1.4009	1.6782
p-valor	1	0.0036	< 0.001	< 0.001	< 0.001	< 0.001
I_{ϵ^+}	5.2665	3.6676	3.5703	3.8280	4.3534	4.7627
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

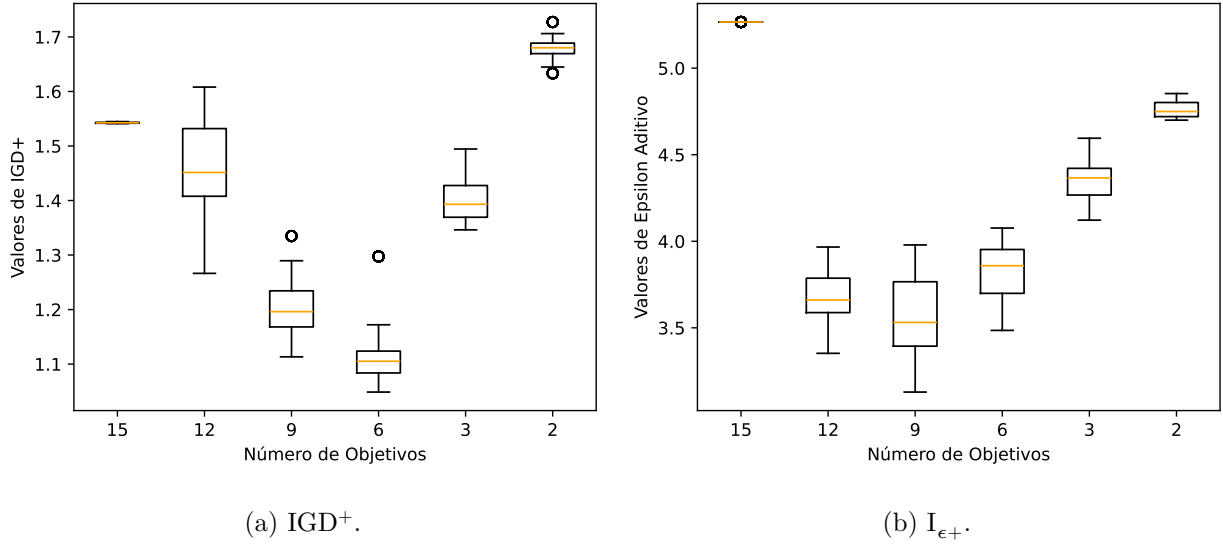


Figura B.18: *Boxplots* dos IQs: DTLZ4 por extração com $m = 15$.

Avaliando os resultados deste DTLZ com $m = 15$ pelas Tabelas B.35 e B.36, nota-se que apenas as médias dos IQs da abordagem $m = 2$ estão acima da média original, com variância de apenas $\approx 18\%$. Além disto, todos os casos apresentaram p-valores abaixo de 5% e a menor média de IGD^+ entre as abordagens se deu em $m = 6$, porém a menor média de I_{ϵ^+} se deu em $m = 9$, resultando no intervalo de variância $\approx [46\%, 66\%]$. Os *boxplots* da Figura B.18 também mostram medianas mais próximas do original no IGD^+ e, outra vez, a dispersão praticamente nula.

Tabela B.37: Porcentagens de variância total explicada: DTLZ4 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
8.3035	5.5603	5.4485	5.3441	5.3272	5.2719	5.1505	5.1400	5.0953	5.0296
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
4.9712	4.9622	4.9198	4.8766	4.8054	4.7737	4.6841	4.6709	4.5966	1.0675

Tabela B.38: Resultados da extração de objetivos: DTLZ4 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	84.9798	65.6043	45.5460	24.6564	13.8638
IGD^+	1.6690	1.7525	1.5522	1.3190	1.6423	1.8631
p-valor	1	< 0.001	< 0.001	< 0.001	0.0644	< 0.001
I_{ϵ^+}	5.0659	4.2672	4.1594	3.9306	4.5657	5.3522
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

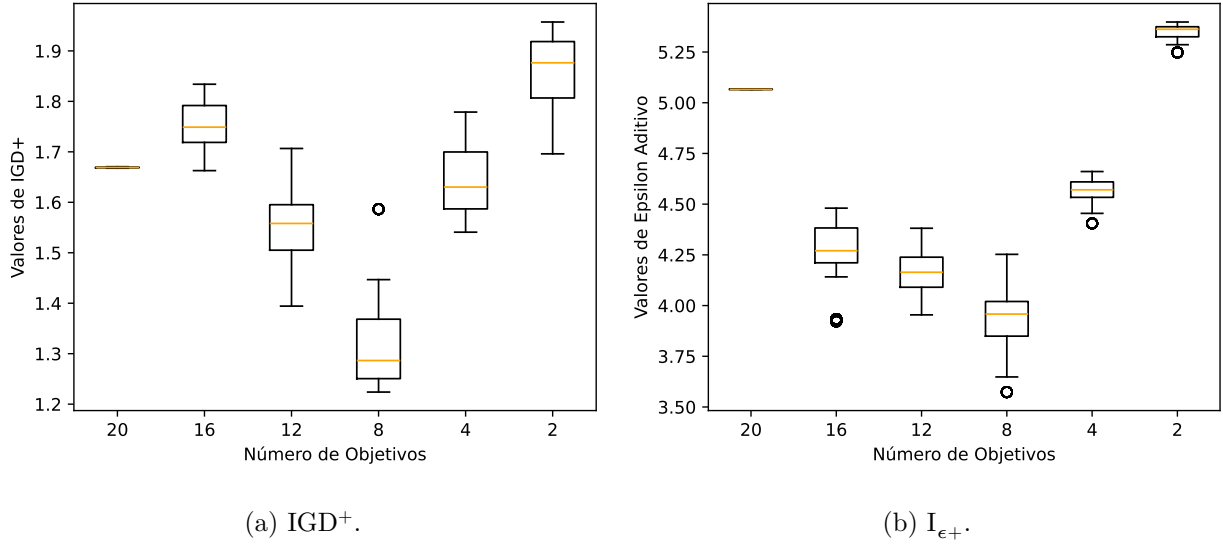


Figura B.19: *Boxplots* dos IQs: DTLZ4 por extração com $m = 20$.

Seguindo para o DTLZ4 com $m = 20$, as Tabelas B.37 e B.38 mostram ganhos de desempenho em menos cenários desta vez, sendo o IGD^+ de $m = 16$ e $m = 2$ maiores que o original, este último com $I_{\epsilon+}$ maior também. Ademais, todos apresentaram evidências estatísticas para rejeição da hipótese nula. O ponto onde as médias são as mais baixas é o mesmo para ambos os indicadores: em $m = 8$ com variância $\approx 46\%$. A Figura B.19 ilustra esse comportamento, além de dispersões mais expressivas para o IGD^+ nos *boxplots*, e novamente o original mostrou variação quase nula.

Tabela B.39: Porcentagens de variância total explicada: DTLZ4 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
5.4625	3.7647	3.7160	3.6566	3.6305	3.5933	3.5766	3.5463	3.5187	3.5027
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
3.4694	3.4364	3.4072	3.3891	3.3646	3.3281	3.3104	3.2823	3.2636	3.2372
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
3.2114	3.1808	3.1651	3.1506	3.1416	3.0762	3.0436	3.0088	2.9419	0.6222

Tabela B.40: Resultados da extração de objetivos: DTLZ4 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	87.3057	71.4562	55.0346	37.9679	20.2303	9.2272
IGD^+	1.9172	2.2537	2.1233	1.8148	1.6132	1.7478	2.2670
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.8025	3.4755	3.3607	3.2401	3.2404	3.5761	4.1651
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

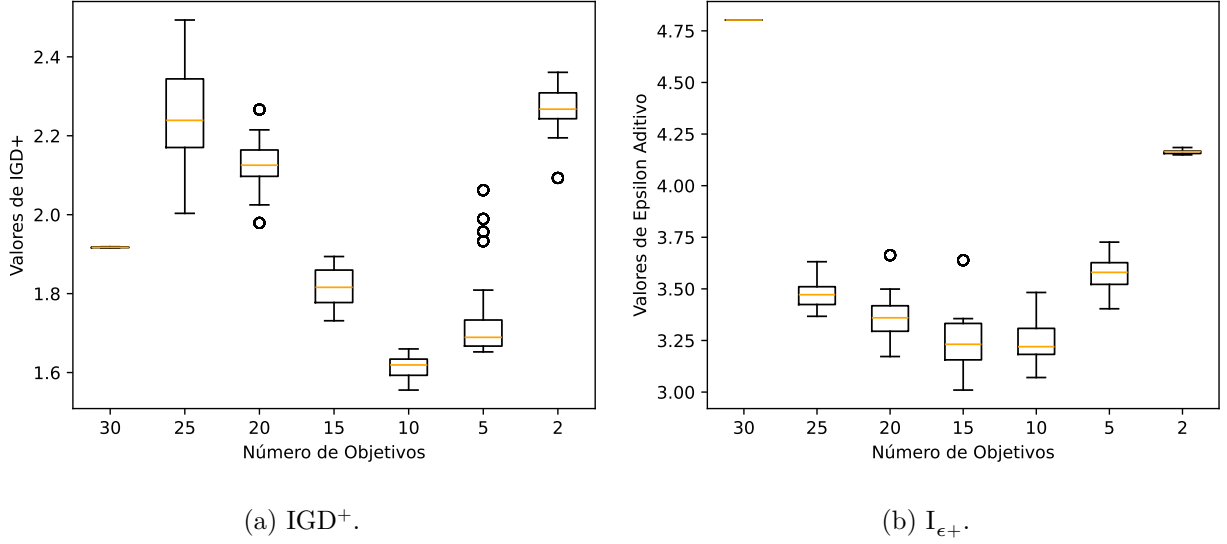


Figura B.20: *Boxplots* dos IQs: DTLZ4 por extração com $m = 30$.

Pelos resultados do cenário $m = 30$ nas Tabelas B.39 e B.40, observa-se perdas de desempenho no IGD⁺ para $m = 25$, $m = 20$ e $m = 2$, porém todos foram abaixo do original no I_ε⁺. E, apesar de $m = 15$ possuir o menor I_ε⁺ entre os demais, a abordagem $m = 10$ mostrou um resultado muito próximo e por isso é plausível considerar este como o ponto de equilíbrio com variância total de $\approx 38\%$. Estas análises também estão representadas nas medianas da Figura B.20, percebe-se também que a dispersão nos *boxplots* das métricas está maior para os casos que houveram perdas, bem como a maior presença de *outliers*. Por fim, o problema original teve a distribuição dos IQs nas 21 execuções quase nula também neste cenário e todos apresentaram um p-valor abaixo de 5%.

B.5 DTLZ5

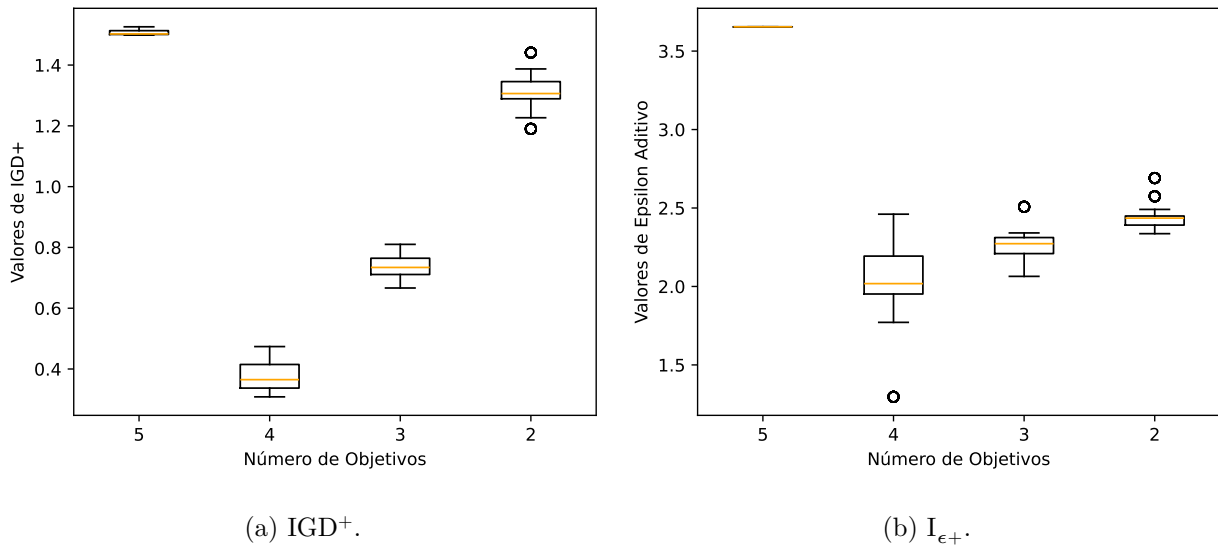
O problema DTLZ5, definido pela Equação 2.8, é utilizado para avaliar a capacidade de um MOEA em convergir para uma curva específica. A PCA mostrou-se eficaz em melhorar o desempenho do NSGA-III, com alta concentração de variância na primeira componente principal.

Tabela B.41: Porcentagens de variância total explicada: DTLZ5 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
81.1006	7.1403	4.5088	3.9670	3.2830

Tabela B.42: Resultados da extração de objetivos: DTLZ5 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	96.7167	92.7497	88.2409
IGD ⁺	1.5062	0.3816	0.7328	1.3093
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.6546	2.0337	2.2603	2.4389
p-valor	1	< 0.001	< 0.001	< 0.001

Figura B.21: *Boxplots* dos IQs: DTLZ5 por extração com $m = 5$.

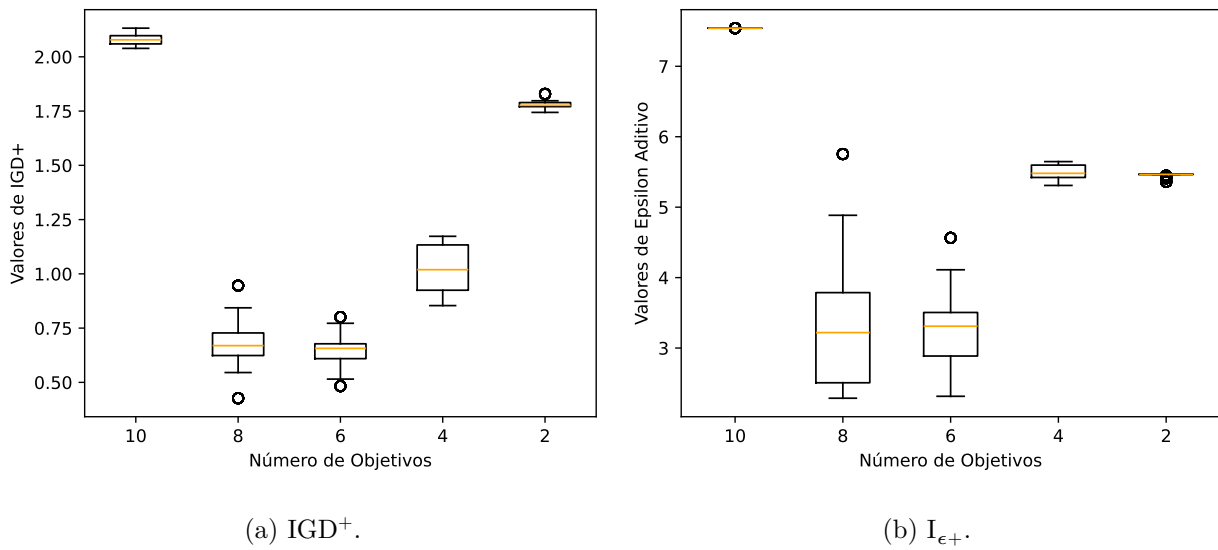
Avaliando as Tabelas B.41 e B.42 do DTLZ5 com $m = 5$, nota-se o alto decaimento nos valores de ambos IQs para a redução $m = 4$ com $\approx 97\%$ de variância total e depois o crescimento das médias e medianas para as demais abordagens. Verifica-se também que todos os conjuntos dos indicadores apresentaram o p-valor < 0.05 . E na Figura B.21, as dispersões se mostram semelhantes nos *boxplots* do IGD⁺ e do $I_{\epsilon+}$, com destaque para o problema original neste quesito, outro detalhe está na ocorrência de mais *outliers* em $I_{\epsilon+}$.

Tabela B.43: Porcentagens de variância total explicada: DTLZ5 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
76.4209	6.9465	3.5430	2.4795	2.1456	2.1270	1.8786	1.6825	1.4379	1.3380

Tabela B.44: Resultados da extração de objetivos: DTLZ5 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	97.2236	93.6625	89.3899	83.3674
IGD ⁺	2.0781	0.6856	0.6456	1.0203	1.7788
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	7.5412	3.3261	3.2581	5.4904	5.4521
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.22: *Boxplots* dos IQs: DTLZ5 por extração com $m = 10$.

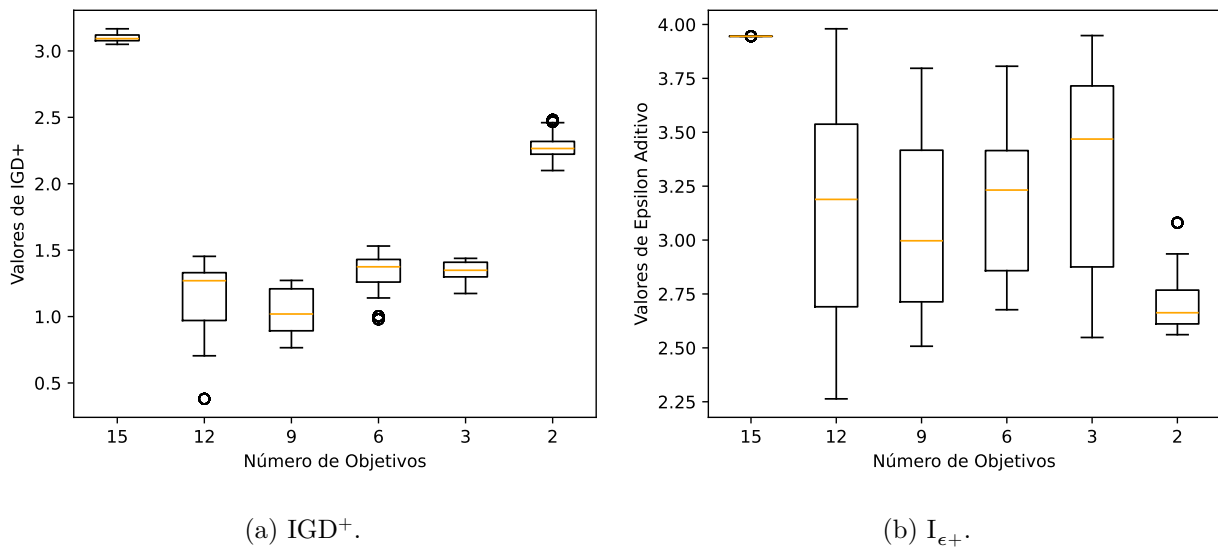
Prosseguindo neste mesmo MaOP com $m = 10$, as Tabelas B.43 e B.44 mostram um comportamento semelhante ao cenário anterior, em que as primeiras reduções obtiveram os melhores resultados e todas as abordagens alcançaram ganhos de desempenho com p-valor $< 5\%$ em confronto com os IQs originais. Analisando a Figura B.22, é possível verificar a paridade entre $m = 8$ e $m = 6$, porém este primeiro apresentou uma dispersão mais acentuada que o segundo. Sendo assim, é admissível adotar um ponto de equilíbrio em $m = 6$ nas médias e medianas dos IQs, com $\approx 94\%$ de variância total explicada. Ademais, os *boxplots* ilustram novamente as distribuições mais espaçadas do $I_{\epsilon+}$ nos dois primeiros casos e quase nulas no $m = 2$ e no DTLZ5 original.

Tabela B.45: Porcentagens de variância total explicada: DTLZ5 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
73.0911	8.6893	3.5526	2.0602	1.7443	1.4765	1.3854	1.2700	1.1853	1.0712
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
1.0560	0.9702	0.9085	0.8176	0.7209					

Tabela B.46: Resultados da extração de objetivos: DTLZ5 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	97.5521	94.4547	90.614	85.333	81.7804
IGD ⁺	3.0984	1.1369	1.0374	1.3317	1.3433	2.2821
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.9451	3.1756	3.0461	3.1774	3.3064	2.7002
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.23: *Boxplots* dos IQs: DTLZ5 por extração com $m = 15$.

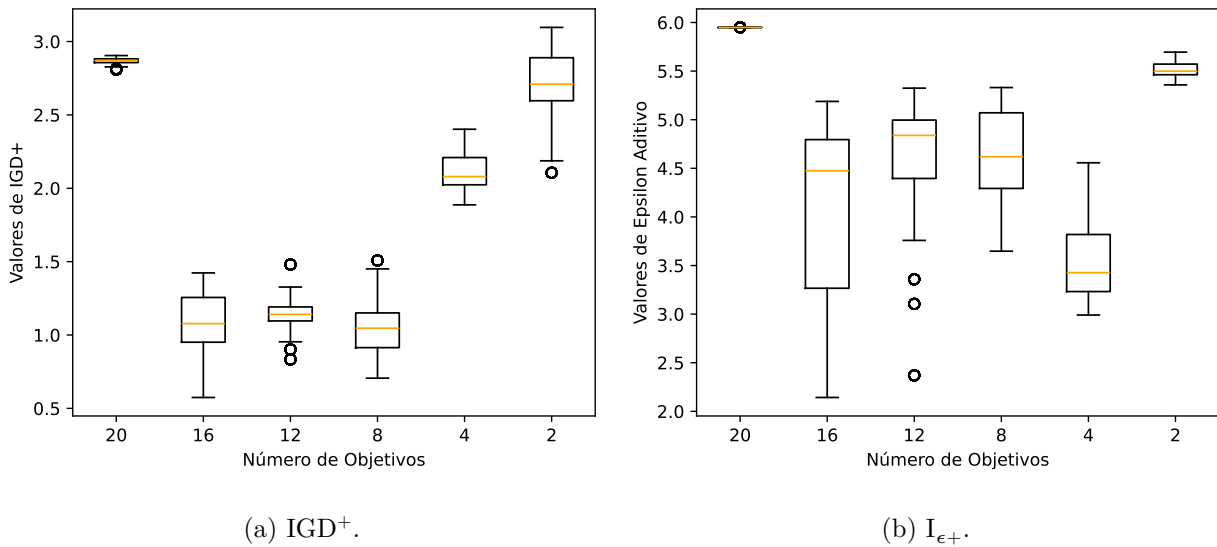
As Tabelas B.45 e B.46 dispõem os resultados para o DTLZ5 com $m = 15$, percebe-se que este cenário também apresentou as primeiras reduções com os melhores resultados do IGD⁺ em comparação com os IQs originais e todas as abordagens com evidências estatísticas para rejeição da hipótese nula. Analisando em paralelo a Figura B.23, destaca-se a média e a mediana do IGD⁺ para o caso $m = 9$ com variância total de $\approx 94\%$. Curiosamente, estes valores para o $I_{\epsilon+}$ foram superiores em $m = 2$ com $\approx 82\%$ de variância, tornando assim mais difícil a tarefa de decisão para um ponto de equilíbrio. A respeito das dispersões nos *boxplots*, observa-se um comportamento muito discrepante dos cenários anteriores e entre os próprios IQs, onde o $I_{\epsilon+}$ mostrou uma distância entre máximo-mínimo bem acentuada, exceto para a última redução.

Tabela B.47: Porcentagens de variância total explicada: DTLZ5 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
71.3930	9.7233	3.6218	2.0161	1.5297	1.1481	1.1298	1.0105	0.9806	0.8989
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
0.8350	0.8303	0.7737	0.7171	0.6460	0.6366	0.6157	0.5434	0.5326	0.4167

Tabela B.48: Resultados da extração de objetivos: DTLZ5 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	97.8905	95.1171	91.5723	86.7542	81.1163
IGD ⁺	2.8663	1.0792	1.1371	1.0640	2.1164	2.6917
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	0.0295
$I_{\epsilon+}$	5.9492	4.0812	4.5460	4.6388	3.5970	5.5134
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.24: *Boxplots* dos IQs: DTLZ5 por extração com $m = 20$.

Seguindo para este MaOP com dimensão igual a 20, as Tabelas B.47 e B.48 apresentam uma concordância com o observado no cenário $m = 15$, do ponto de vista que as primeiras reduções alcançaram os melhores ganhos de desempenho para o IGD⁺. Porém, a última abordagem desta não obteve bons resultados em nenhum dos IQs, destacando-se desta vez $m = 4$, com a menor média e mediana para o $I_{\epsilon+}$ e variância total explicada de $\approx 87\%$. Além disso, constata-se que todas as abordagens obtiveram p-valor $< 5\%$. Analisando os *boxplots* na Figura B.24, percebe-se o padrão observado nos cenários $m = 5$ e $m = 10$, onde

as dispersões foram piores no segundo IQ, exceto para $m = 2$ e o problema original que é quase nula.

Tabela B.49: Porcentagens de variância total explicada: DTLZ5 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
68.3612	11.6163	4.1978	2.0078	1.4001	1.0410	0.8998	0.7894	0.7134	0.7079
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
0.6317	0.6292	0.5997	0.5668	0.5417	0.5306	0.4969	0.4646	0.4317	0.4082
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
0.3859	0.3574	0.3502	0.3247	0.3160	0.3073	0.2776	0.2473	0.2210	0.1753

Tabela B.50: Resultados da extração de objetivos: DTLZ5 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	98.7700	97.0358	94.7038	91.7347	87.5832	79.9775
IGD ⁺	3.2504	1.2483	1.1007	0.9998	0.9416	0.8921	0.9031
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	8.3450	6.7710	6.7514	6.9031	6.5938	5.6376	6.9028
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

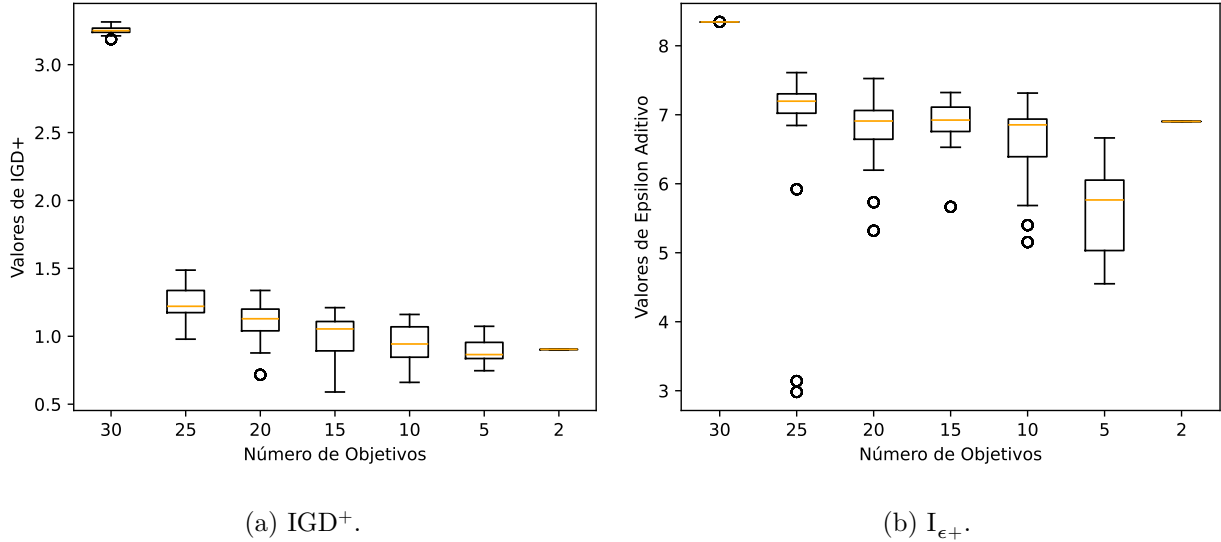


Figura B.25: *Boxplots* dos IQs: DTLZ5 por extração com $m = 30$.

Por último, os resultados do DTLZ5 com $m = 30$ estão dispostos nas Tabelas B.49 e B.50, analisando-as nota-se que as médias do IGD⁺ diminuíram gradualmente até crescer na redução $m = 2$. A respeito do $I_{\epsilon+}$, como pode ser visualizado também na Figura B.25,

percebe-se um comportamento diferente, onde todas as abordagens obtiveram médias e medianas próximas entre si e com pequenas oscilações, exceto para $m = 5$ com variância de $\approx 88\%$, que apresentou os menores valores de IQs de maneira geral. Adicionalmente, todas as reduções obtiveram ganhos de desempenho com evidências estatísticas para rejeição da hipótese nula. Os *boxplots* comportaram-se diferente dos casos anteriores, havendo a manifestação de maior número de *outliers* no $I_{\epsilon+}$.

B.6 DTLZ6

O problema DTLZ6, conforme a Equação 2.9, é uma variação do DTLZ5 projetada para dificultar a convergência do MOEA. A PCA mostrou-se eficaz em melhorar o desempenho do NSGA-III, mesmo com o aumento na dificuldade de convergência.

Tabela B.51: Porcentagens de variância total explicada: DTLZ6 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
50.0482	21.9895	14.7614	10.7778	2.4228

Tabela B.52: Resultados da extração de objetivos: DTLZ6 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	97.5769	86.7991	72.0377
IGD ⁺	2.2116	0.2176	0.2389	0.4985
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	2.1004	0.8600	1.1567	1.6516
p-valor	1	< 0.001	< 0.001	< 0.001

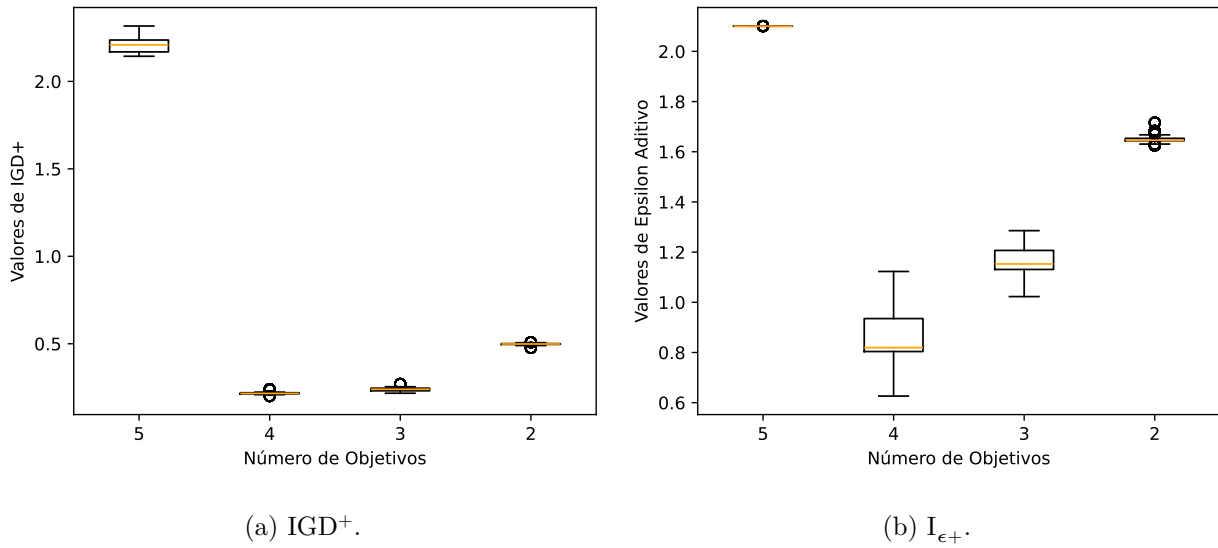


Figura B.26: *Boxplots* dos IQs: DTLZ6 por extração com $m = 5$.

Analisando as Tabelas B.51 e B.52, nota-se o decrescimento expressivo nos valores dos IQs para todas as abordagens em relação ao original, com todos os conjuntos de IQs evidenciando estatísticas para rejeição da hipótese nula. Sendo os menores em $m = 4$ com variância total explicada de $\approx 98\%$. O ponto de equilíbrio nesta redução pode ser visto na Figura B.26, uma vez que há um aumento nas medianas dos IQs de $m = 3$ e posteriormente de $m = 2$. A respeito dos *boxplots*, a dispersão se mostra invertida entre o caso original e os reduzidos, isto é, enquanto o primeiro possui distribuição menor no IGD^+ , o segundo é no I_{ϵ^+} , e vice-versa.

Tabela B.53: Porcentagens de variância total explicada: DTLZ6 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
40.7446	16.2441	9.7821	7.4709	6.3339	5.5691	4.7441	4.4903	3.4060	1.2145

Tabela B.54: Resultados da extração de objetivos: DTLZ6 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	95.3791	86.1447	74.2417	56.9887
IGD^+	2.7887	0.7016	0.5401	0.4989	0.6291
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
I_{ϵ^+}	3.5961	1.9357	1.7191	1.6182	2.3157
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

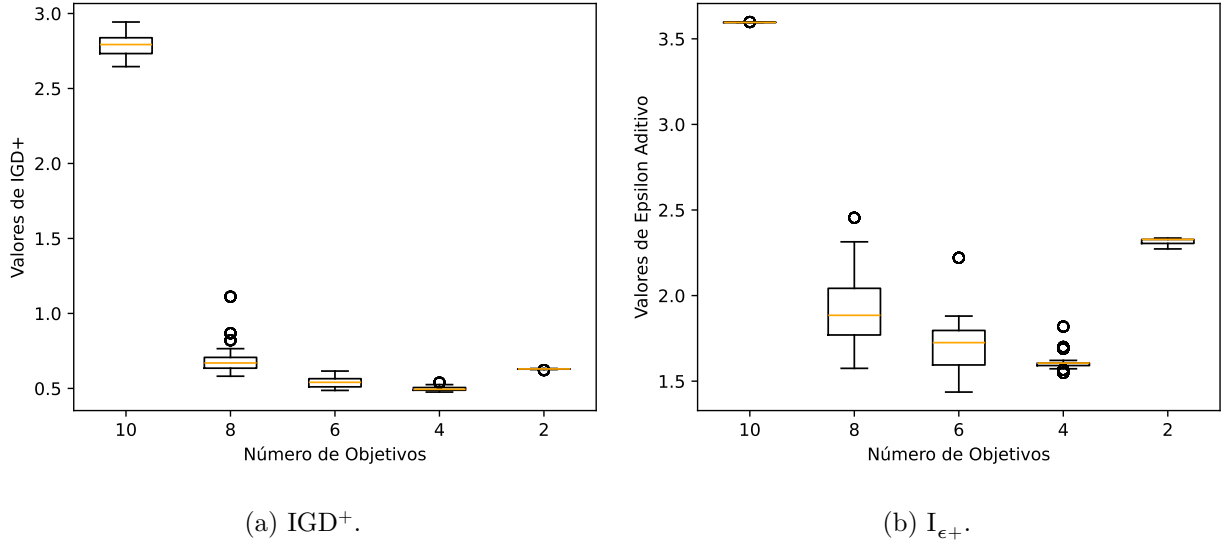


Figura B.27: *Boxplots* dos IQs: DTLZ6 por extração com $m = 10$.

Seguindo para a análise deste MaOP com $m = 10$, verifica-se nas Tabelas B.53 e B.54 que houveram ganhos de desempenho com p-valores $< 5\%$ em todas as reduções, segundo os IQs calculados. Por meio da Figura B.27, percebe-se o ponto de equilíbrio na abordagem para $m = 4$ que possui a variância explicada total de $\approx 74\%$. Ademais, as dispersões dos *boxplots* se comportaram da mesma maneira que no cenário anterior, sendo que as do $I_{\epsilon+}$ foram mais expressivas que as do IGD^+ nas reduções, e o contrário no problema original.

Tabela B.55: Porcentagens de variância total explicada: DTLZ6 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
39.3527	14.1165	8.6825	6.2316	5.0627	4.4641	3.9234	3.2730	3.0469	2.7580
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
2.5560	2.2528	2.1192	1.3753	0.7846					

Tabela B.56: Resultados da extração de objetivos: DTLZ6 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	95.7202	88.1534	77.9101	62.1517	53.4692
IGD^+	2.5953	0.8959	0.8073	0.7065	0.8748	1.4746
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	3.7933	2.1413	2.1152	2.3088	2.7961	2.9509
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

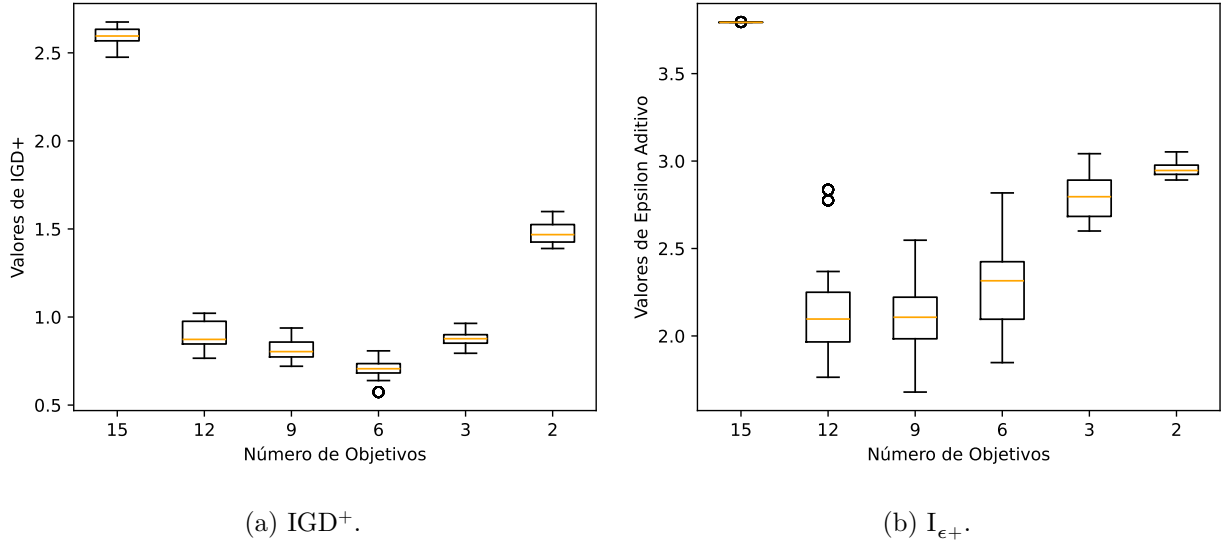


Figura B.28: *Boxplots* dos IQs: DTLZ6 por extração com $m = 15$.

Avaliando os resultados deste DTLZ com $m = 15$ pelas Tabelas B.55 e B.56, nota-se que todas as médias dos IQs reduzidos estão abaixo da média original. Além disto, todos os casos apresentaram p-valores abaixo de 5% e a menor média de IGD^+ entre as abordagens se deu em $m = 6$, porém a menor média de I_{ϵ^+} se deu em $m = 9$, resultando no intervalo de variância $\approx [78\%, 88\%]$. Os *boxplots* da Figura B.28 mostram as dispersões com o mesmo padrão dos cenários passados, com exceção em $m = 2$.

Tabela B.57: Porcentagens de variância total explicada: DTLZ6 com $m = 20$.

$\lambda^1_{\%}$	$\lambda^2_{\%}$	$\lambda^3_{\%}$	$\lambda^4_{\%}$	$\lambda^5_{\%}$	$\lambda^6_{\%}$	$\lambda^7_{\%}$	$\lambda^8_{\%}$	$\lambda^9_{\%}$	$\lambda^{10}_{\%}$
39.5875	13.3165	8.0203	5.3740	4.1098	3.7115	3.0936	2.8535	2.6627	2.5253
$\lambda^{11}_{\%}$	$\lambda^{12}_{\%}$	$\lambda^{13}_{\%}$	$\lambda^{14}_{\%}$	$\lambda^{15}_{\%}$	$\lambda^{16}_{\%}$	$\lambda^{17}_{\%}$	$\lambda^{18}_{\%}$	$\lambda^{19}_{\%}$	$\lambda^{20}_{\%}$
2.3539	2.0691	1.9134	1.7028	1.4970	1.4374	1.2114	1.1675	0.7649	0.6268

Tabela B.58: Resultados da extração de objetivos: DTLZ6 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	96.2283	89.6777	80.0667	66.2983	52.9040
IGD^+	2.6305	1.1859	0.9491	0.8371	1.0202	1.8835
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
I_{ϵ^+}	3.2512	2.2969	2.1587	2.1323	2.4618	2.8381
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

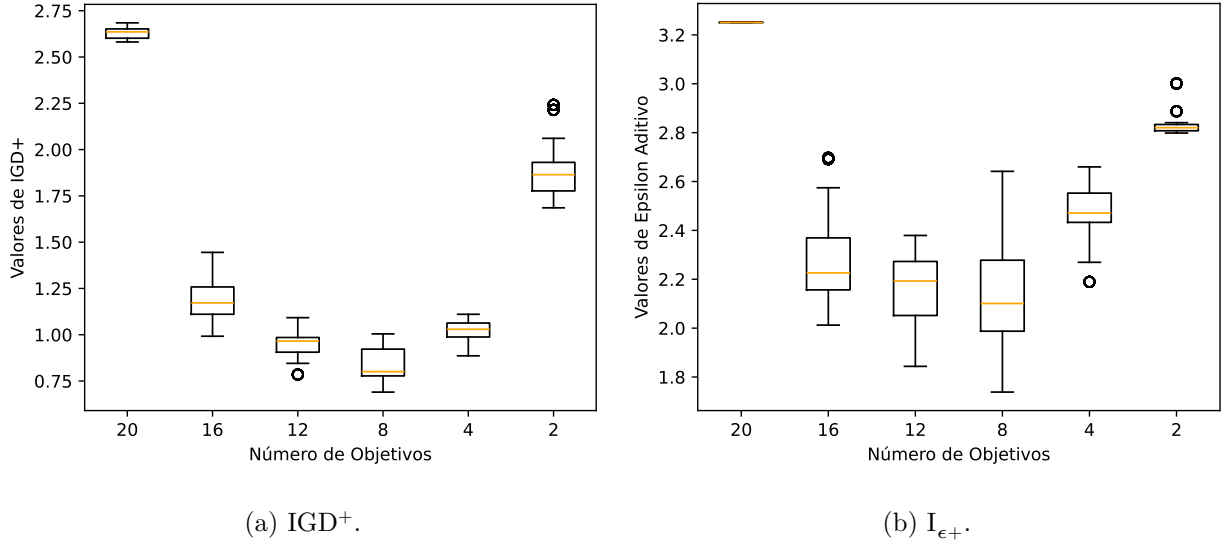


Figura B.29: *Boxplots* dos IQs: DTLZ6 por extração com $m = 20$.

Seguindo para o DTLZ6 com $m = 20$, as Tabelas B.57 e B.58 mostram ganhos de desempenho em todos os cenários mais uma vez, com todos apresentando evidências estatísticas para rejeição da hipótese nula. O ponto onde as médias são as mais baixas é o mesmo para ambos os indicadores: em $m = 8$ com variância $\approx 80\%$. A Figura B.29 ilustra esse comportamento, além das dispersões com o comportamento muito semelhante ao cenário $m = 15$.

Tabela B.59: Porcentagens de variância total explicada: DTLZ6 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
37.2726	11.7773	7.5042	5.1321	3.8617	3.2107	2.9144	2.6143	2.4727	2.0719
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
1.8863	1.8329	1.7246	1.6312	1.5529	1.3666	1.3310	1.2650	1.1252	1.0817
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
0.9905	0.8912	0.8085	0.7194	0.6212	0.5511	0.5408	0.4809	0.4163	0.3491

Tabela B.60: Resultados da extração de objetivos: DTLZ6 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	97.6601	93.6293	87.4598	78.8319	65.5479	49.0499
IGD ⁺	3.0736	1.4937	1.4076	1.1760	1.0012	0.8456	1.0873
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	4.3688	2.9055	2.8195	2.8640	2.7181	2.8824	3.2728
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

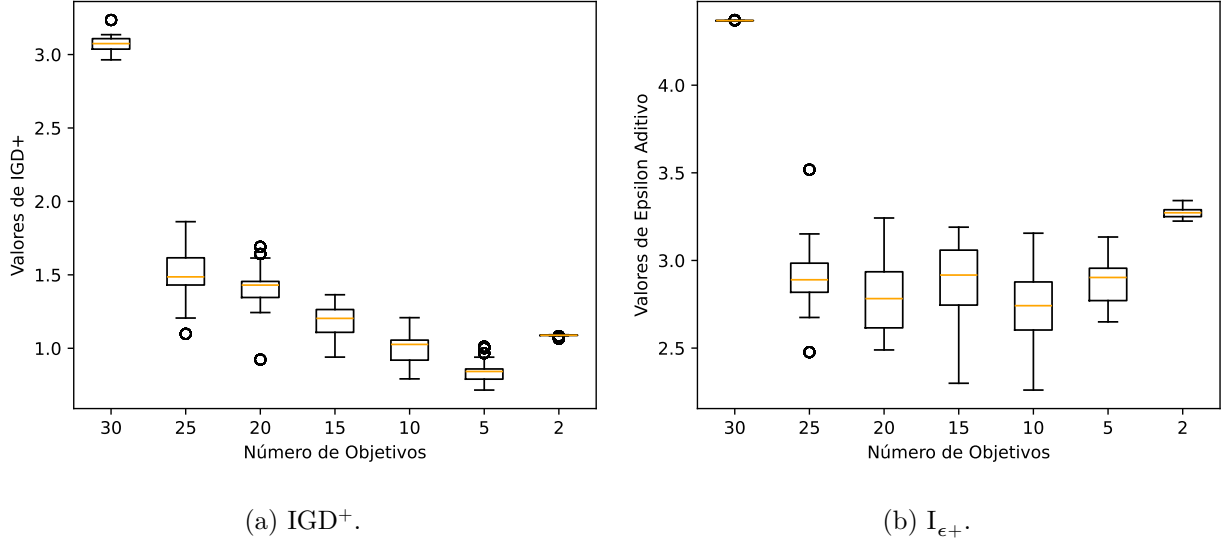


Figura B.30: *Boxplots* dos IQs: DTLZ6 por extração com $m = 30$.

Pelos resultados do cenário $m = 30$ nas Tabelas B.59 e B.60, observa-se outra vez os ganhos de desempenho nos IQs reduzidos com p-valor abaixo de 5%, sendo a menor média de IGD^+ em $m = 5$ e $I_{\epsilon+}$ em $m = 10$. Portanto, é plausível considerar o ponto de equilíbrio no intervalo $\approx [66\%, 79\%]$ de variância total explicada. Estas análises também estão representadas nas medianas da Figura B.30, percebe-se também que a dispersão nos *boxplots* das métricas está sob o mesmo padrão que os outros cenários.

B.7 DTLZ7

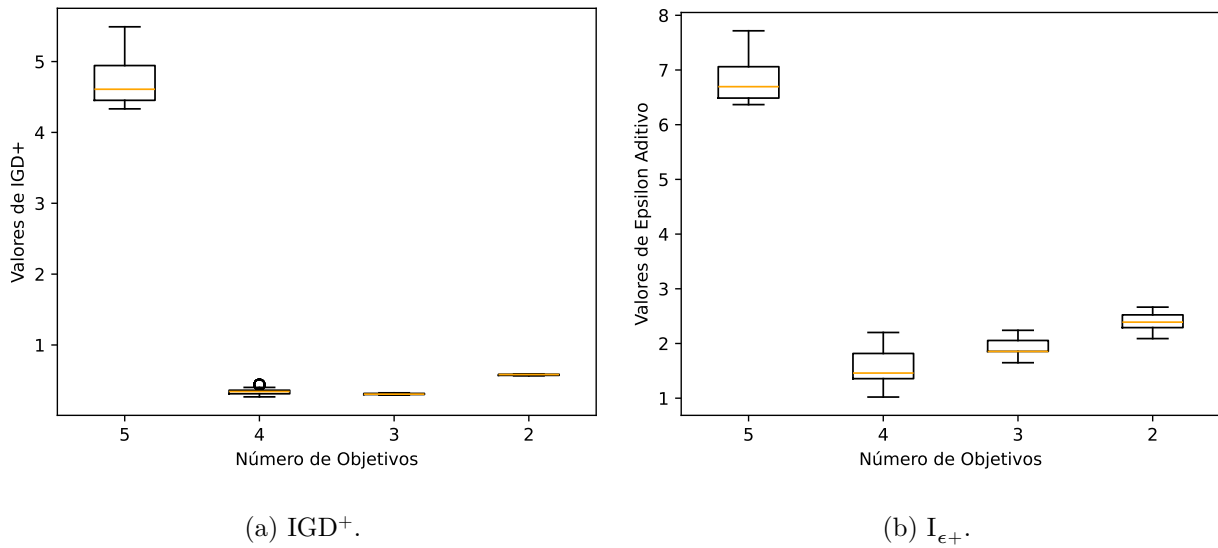
O problema DTLZ7, definido pela Equação 2.10, é utilizado para testar a capacidade de um MOEA em manter subpopulações distribuídas em diferentes regiões Pareto-ótimas. A PCA mostrou-se eficaz em melhorar o desempenho do NSGA-III, embora a concentração das soluções em uma única região Pareto-ótima possa introduzir vieses nos resultados.

Tabela B.61: Porcentagens de variância total explicada: DTLZ7 com $m = 5$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$
22.8194	20.4352	20.0512	19.0299	17.6640

Tabela B.62: Resultados da extração de objetivos: DTLZ7 com $m = 5$.

m	5	4	3	2
$\lambda_{\%}$	100	82.3357	63.3058	43.2546
IGD ⁺	4.7595	0.3439	0.3073	0.5803
p-valor	1	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	6.8523	1.5426	1.9447	2.3944
p-valor	1	< 0.001	< 0.001	< 0.001

Figura B.31: *Boxplots* dos IQs: DTLZ7 por extração com $m = 5$.

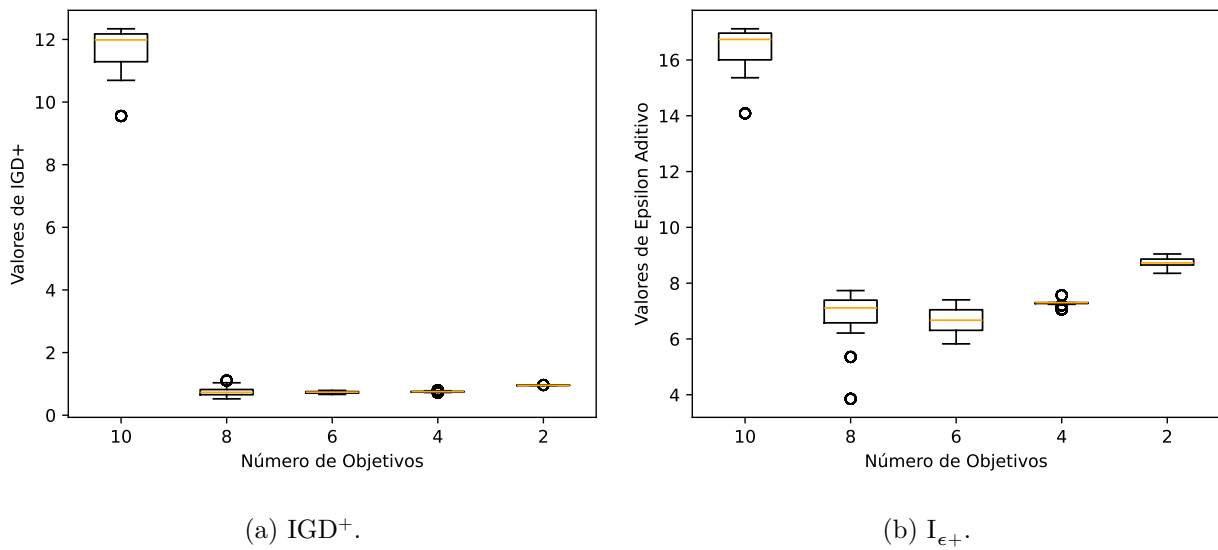
Analisando as Tabelas B.61 e B.62 do DTLZ7 com $m = 5$, observa-se o decrescimento nos valores dos IQs para todas as abordagens em comparação ao problema original, sendo o menor valor de IGD⁺ em $m = 3$ e de $I_{\epsilon+}$ em $m = 4$, resultando no intervalo $\approx [63\%, 82\%]$ de variância total explicada. Verifica-se também que todos os conjuntos de IQs apresentaram o p-valor < 0.05 . Além disto, pela Figura B.31, nota-se que as distribuições do $I_{\epsilon+}$ estão maiores que as do IGD⁺ e o problema original é praticamente idêntico entre os dois.

Tabela B.63: Porcentagens de variância total explicada: DTLZ7 com $m = 10$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
11.3934	10.9454	10.4998	10.3099	10.0522	9.8581	9.6185	9.4961	9.2638	8.5624

Tabela B.64: Resultados da extração de objetivos: DTLZ7 com $m = 10$.

m	10	8	6	4	2
$\lambda_{\%}$	100	82.1734	63.0588	43.1485	22.3388
IGD ⁺	11.6889	0.7429	0.7333	0.7534	0.9535
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	16.4323	6.7992	6.6355	7.2683	8.7412
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.32: *Boxplots* dos IQs: DTLZ7 por extração com $m = 10$.

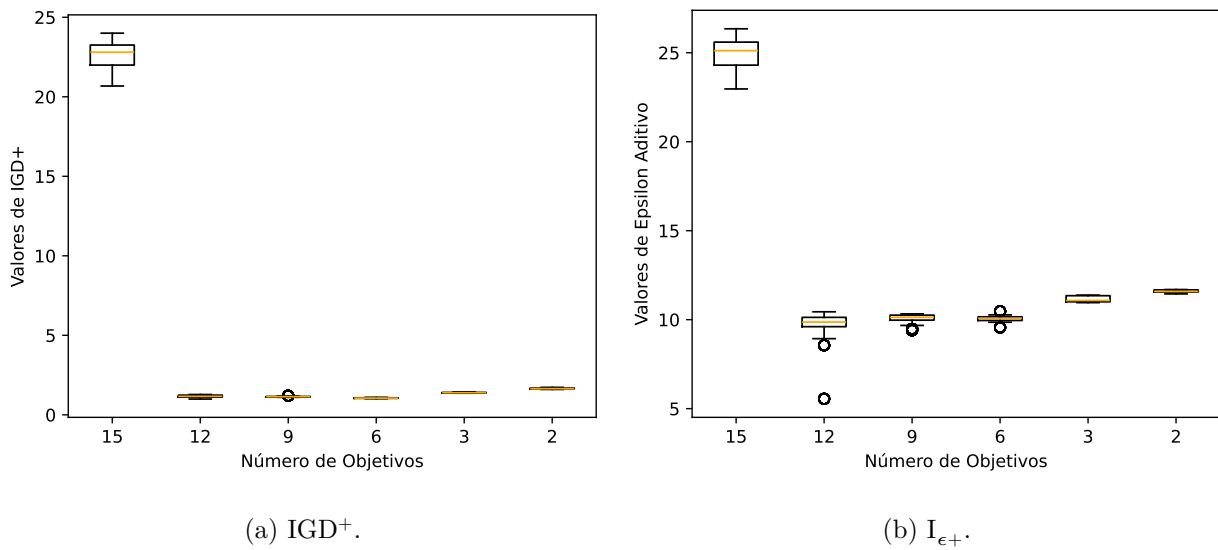
Prosseguindo neste mesmo MaOP agora com dimensão igual a 10, as Tabelas B.63 e B.64 apresentam que todas as abordagens alcançaram ganhos de desempenho estatisticamente positivos em confronto com os IQs originais, inclusive para $m = 2$ com apenas $\approx 22\%$ de variância. Ademais, as menores médias manifestaram-se na redução para $m = 6$, com variância total explicada de $\approx 63\%$. Pela Figura B.32, é possível verificar o mesmo comportamento dos *boxplots* que no cenário anterior, inclusive a igualdade entre eles no problema original.

Tabela B.65: Porcentagens de variância total explicada: DTLZ7 com $m = 15$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
7.7162	7.4513	7.2571	7.0837	7.0398	6.7734	6.6523	6.5894	6.5300	6.4727
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$					
6.4029	6.2072	6.1654	5.9692	5.6887					

Tabela B.66: Resultados da extração de objetivos: DTLZ7 com $m = 15$.

m	15	12	9	6	3	2
$\lambda_{\%}$	100	82.1760	63.0932	43.3215	22.4246	15.1675
IGD ⁺	22.6123	1.1664	1.1443	1.0552	1.3954	1.6576
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	24.9410	9.6108	10.0291	10.0639	11.1414	11.6041
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.33: *Boxplots* dos IQs: DTLZ7 por extração com $m = 15$.

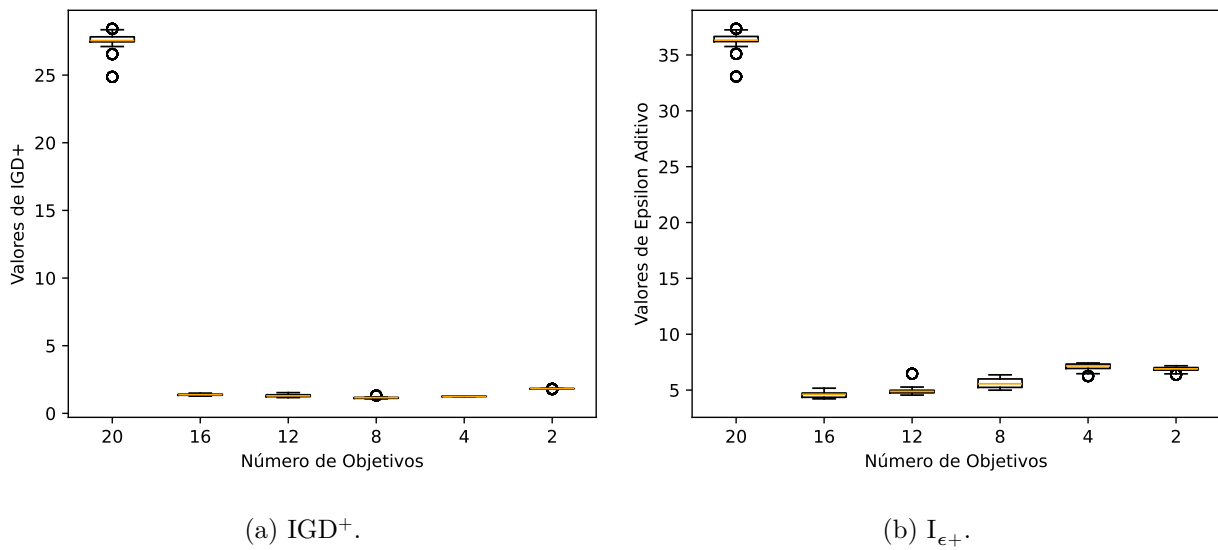
As Tabelas B.65 e B.66 apresentam os resultados para o DTLZ7 com $m = 15$. Todas as reduções mostraram ganhos de desempenho, inclusive em $m = 2$, apesar da baixa variância de aproximadamente 15%, com evidências estatísticas suficientes para rejeitar a hipótese nula. A abordagem com $m = 6$ apresentou o melhor resultado para o IGD⁺, enquanto $m = 12$ se destacou no $I_{\epsilon+}$, dificultando a escolha de um ponto de equilíbrio, já que esses resultados abrangem um intervalo de variância total entre $\approx [63\%, 82\%]$. Os *boxplots* na Figura B.33 destacam a expressiva diferença entre as reduções e o problema original, que novamente exibiu *boxplots* idênticos para ambos os IQs.

Tabela B.67: Porcentagens de variância total explicada: DTLZ7 com $m = 20$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
6.0713	5.7064	5.6868	5.4759	5.4197	5.2371	5.1807	5.1153	5.0059	4.9720
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
4.8862	4.8756	4.8174	4.7338	4.6997	4.5630	4.5499	4.4148	4.3443	4.2431

Tabela B.68: Resultados da extração de objetivos: DTLZ7 com $m = 20$.

m	20	16	12	8	4	2
$\lambda_{\%}$	100	82.4468	63.6329	43.8932	22.9404	11.7777
IGD ⁺	27.5110	1.3730	1.3101	1.1588	1.2570	1.8251
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	36.2567	4.5892	4.9556	5.6450	7.0788	6.8685
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.34: *Boxplots* dos IQs: DTLZ7 por extração com $m = 20$.

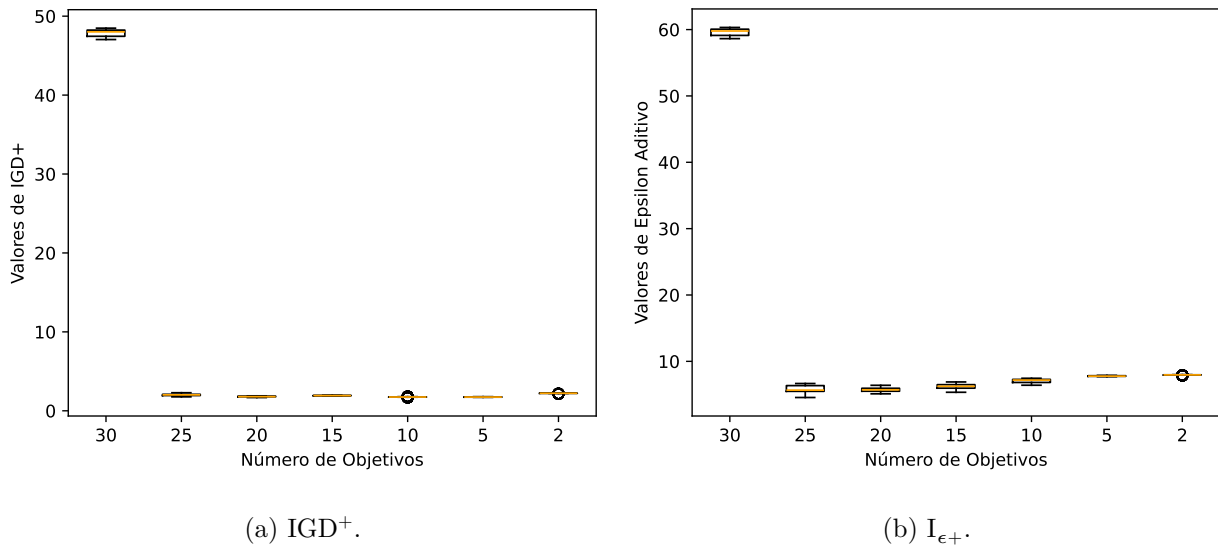
Seguindo para este MaOP com dimensão igual a 20, as Tabelas B.67 e B.68 mostram que todas as abordagens apresentaram ganhos de desempenho, com p-valor < 5%. Especificamente, o melhor valor de IGD⁺ foi obtido em $m = 8$ e de $I_{\epsilon+}$ em $m = 16$. Como no cenário anterior, o intervalo de variância total é amplo, variando entre 44% e 82%. Essas análises também podem ser observadas na Figura B.34, que evidencia ainda mais a diferença nas médias e medianas entre as reduções e o problema original, de forma mais expressiva que no cenário anterior. Além disso, observa-se novamente a semelhança entre os *boxplots* do problema original, reforçando o padrão identificado nos cenários antecessores.

Tabela B.69: Porcentagens de variância total explicada: DTLZ7 com $m = 30$.

$\lambda_{\%}^1$	$\lambda_{\%}^2$	$\lambda_{\%}^3$	$\lambda_{\%}^4$	$\lambda_{\%}^5$	$\lambda_{\%}^6$	$\lambda_{\%}^7$	$\lambda_{\%}^8$	$\lambda_{\%}^9$	$\lambda_{\%}^{10}$
4.0091	3.9357	3.8520	3.8239	3.7848	3.7316	3.7179	3.6521	3.5895	3.5135
$\lambda_{\%}^{11}$	$\lambda_{\%}^{12}$	$\lambda_{\%}^{13}$	$\lambda_{\%}^{14}$	$\lambda_{\%}^{15}$	$\lambda_{\%}^{16}$	$\lambda_{\%}^{17}$	$\lambda_{\%}^{18}$	$\lambda_{\%}^{19}$	$\lambda_{\%}^{20}$
3.4982	3.4328	3.3976	3.3403	3.3158	3.3082	3.2753	3.2272	3.2169	3.1768
$\lambda_{\%}^{21}$	$\lambda_{\%}^{22}$	$\lambda_{\%}^{23}$	$\lambda_{\%}^{24}$	$\lambda_{\%}^{25}$	$\lambda_{\%}^{26}$	$\lambda_{\%}^{27}$	$\lambda_{\%}^{28}$	$\lambda_{\%}^{29}$	$\lambda_{\%}^{30}$
3.1329	3.0981	3.0476	2.9556	2.9202	2.8921	2.8609	2.7944	2.7663	2.7311

Tabela B.70: Resultados da extração de objetivos: DTLZ7 com $m = 30$.

m	30	25	20	15	10	5	2
$\lambda_{\%}$	100	85.9536	70.7992	54.5948	37.6101	19.4055	7.9448
IGD ⁺	47.8578	2.0170	1.8011	1.9253	1.7473	1.7443	2.2199
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
$I_{\epsilon+}$	59.5984	5.8095	5.7124	6.2049	7.0644	7.7828	7.9541
p-valor	1	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Figura B.35: *Boxplots* dos IQs: DTLZ7 por extração com $m = 30$.

Por último, os resultados do DTLZ7 com $m = 30$ estão apresentados nas Tabelas B.69 e B.70. Ao analisá-las, verifica-se que as médias dos IQs foram significativamente menores que as médias do problema original, com p-valores abaixo de 5%. No entanto, como mostrado na Figura B.35, o IGD⁺ encontrou um ponto de equilíbrio em $m = 5$ e o $I_{\epsilon+}$ em $m = 20$, ampliando ainda mais o intervalo de variância total comparado aos casos anteriores. Além disso, a diferença nas médias e medianas entre as reduções e o problema original foi ainda

mais acentuada que no cenário anterior. Observa-se também, outra vez, a semelhança entre os *boxplots* do problema original.

Destaca-se a semelhança no comportamento da variância total explicada deste DTLZ com o DTLZ4, onde ela é mais distribuída entre as componentes. Neste MaOP, a identificação de um ponto de equilíbrio comum para todos os cenários tornou-se ainda mais desafiadora, dado que a inconsistência nas médias foi mais pronunciada em quatro cenários distintos, exceto em $m = 10$.