

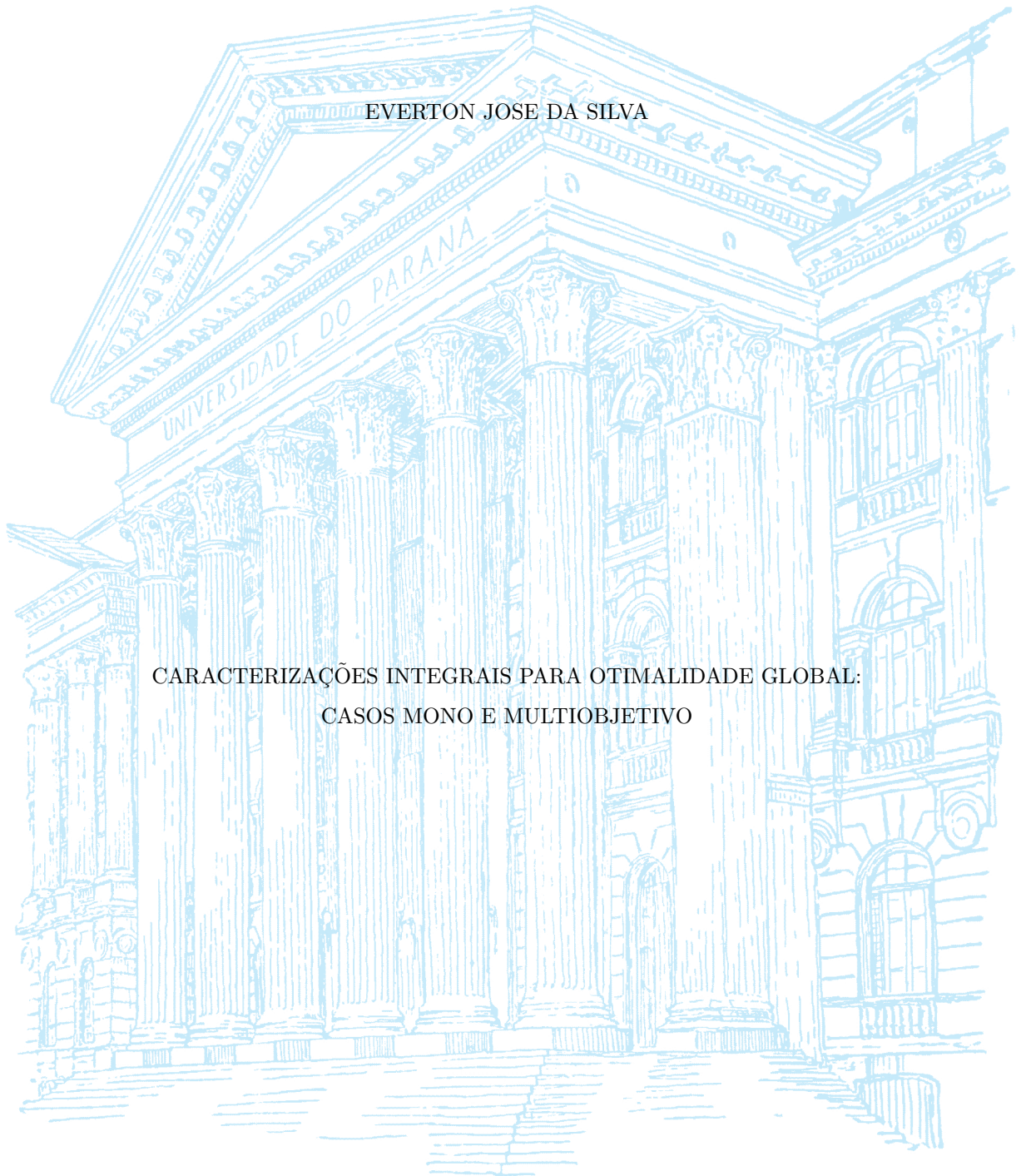
UNIVERSIDADE FEDERAL DO PARANÁ

EVERTON JOSE DA SILVA

CARACTERIZAÇÕES INTEGRAIS PARA OTIMALIDADE GLOBAL:
CASOS MONO E MULTI OBJETIVO

CURITIBA

2020



EVERTON JOSE DA SILVA

CARACTERIZAÇÕES INTEGRAIS PARA OTIMALIDADE GLOBAL:
CASOS MONO E MULTIOBJETIVO

Dissertação de Mestrado apresentada ao curso de Pós-Graduação em Matemática, Setor de Ciências Exatas, Universidade Federal do Paraná, como requisito parcial à obtenção do Título de Mestre em Matemática.

Orientadora: Profa. Dra. Elizabeth Wegner Karas.

Coorientadora: Profa. Dra. Lucelina Batista dos Santos.

CURITIBA

2020

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

S586c

Silva, Everton Jose da

Caracterizações integrais para otimalidade global: casos mono e multiobjetivo [recurso eletrônico] Everton Jose da Silva. – Curitiba, 2020.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Matemática, 2020.

Orientadora: Elizabeth Wegner Karas.

Coorientadora: Lucelina Batista dos Santos.

1. Algoritmos. 2. Otimização matemática. I. Universidade Federal do Paraná. II. Karas, Elizabeth Wegner. III. Santos, Lucelina Batista dos. IV. Título.

CDD: 518.1

Bibliotecária: Vanusa Maciel CRB- 9/1928

ATA DE SESSÃO PÚBLICA DE DEFESA DE MESTRADO PARA A OBTENÇÃO DO GRAU DE MESTRE EM MATEMÁTICA

No dia vinte e quatro de julho de dois mil e vinte às 14:00 horas, na sala virtual meet.google.com/ybk-ftox-hbc, em caráter remoto, foram instaladas as atividades pertinentes ao rito de defesa de dissertação do mestrando **EVERTON JOSE DA SILVA**, intitulada: **Caracterizações integrais para otimalidade global: casos mono e multiobjetivo**, sob orientação da Profa. Dra. ELIZABETH WEGNER KARAS. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: ELIZABETH WEGNER KARAS (UNIVERSIDADE FEDERAL DO PARANÁ), VALERIANO ANTUNES DE OLIVEIRA (UNIVERSIDADE EST. PAULISTA JÚLIO DE MESQUITA FILHO/SJR. PRETO), FERNANDA MARIA PEREIRA (LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA). A presidência iniciou os ritos definidos pelo Colegiado do Programa e, após exarados os pareceres dos membros do comitê examinador e da respectiva contra argumentação, ocorreu a leitura do parecer final da banca examinadora, que decidiu pela APROVAÇÃO. Este resultado deverá ser homologado pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais definidos pelo programa. A outorga de título de mestre está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, ELIZABETH WEGNER KARAS, lavrei a presente ata, que vai assinada por mim e pelos demais membros da Comissão Examinadora.

CURITIBA, 24 de Julho de 2020.

Assinatura Eletrônica

27/07/2020 17:50:52.0

ELIZABETH WEGNER KARAS

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

27/07/2020 16:38:04.0

VALERIANO ANTUNES DE OLIVEIRA

Avaliador Externo (UNIVERSIDADE EST. PAULISTA JÚLIO DE MESQUITA FILHO/SJR. PRETO)

Assinatura Eletrônica

27/07/2020 16:36:24.0

FERNANDA MARIA PEREIRA

Avaliador Externo (LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA)



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MATEMÁTICA -
40001016041P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **EVERTON JOSE DA SILVA** intitulada: **Caracterizações integrais para otimalidade global: casos mono e multiobjetivo**, sob orientação da Profa. Dra. ELIZABETH WEGNER KARAS, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 24 de Julho de 2020.

Assinatura Eletrônica

27/07/2020 17:50:52.0

ELIZABETH WEGNER KARAS

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

27/07/2020 16:38:04.0

VALERIANO ANTUNES DE OLIVEIRA

Avaliador Externo (UNIVERSIDADE EST. PAULISTA JÚLIO DE MESQUITA FILHO/SJR. PRETO)

Assinatura Eletrônica

27/07/2020 16:36:24.0

FERNANDA MARIA PEREIRA

Avaliador Externo (LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA)

Dedico esta dissertação à minha mãe, Maria.

AGRADECIMENTOS

Agradeço primeiramente a mim mesmo, por conseguir vencer todo dia uma batalha constante, por acreditar que sou capaz, por ter feito esse trabalho e nunca ter desistido daquilo que acredito. Agradeço também à Deus, por me dar forças diariamente para conseguir conquistar meus sonhos.

Aos meus pais e familiares, que foram incentivadores e que sempre acreditaram em meus sonhos. Principalmente à minha mãe por estar sempre ao meu lado, tornando possível este importante passo em minha vida.

Aos meus amigos da graduação que sempre estiveram comigo, mesmo que não presencialmente, por tornarem a universidade um lugar mais agradável durante a graduação. Agradeço em especial a Líbera, Lorayne, Isalene e Fernanda, pelos momentos e experiências compartilhados, com certeza, levarei para o resto de minha vida.

Aos colegas do PPGM que tornaram o ambiente menos hostil. Principalmente aos da sala 19, pelos conhecimentos compartilhados, amizade e companheirismo.

À minha amiga Flávia que me ajudou em momentos difíceis, de questionamentos e de muita apreensão. À minha amiga Carina pelos momentos de reflexão, apoio, incentivo e principalmente pela ajuda. Serei sempre grato.

À Luciane e ao Elias pelo apoio, companheirismo, compreensão e por estarem presentes nos momentos alegres e difíceis nestes meses de convivência no apartamento dos matemáticos.

Aos meus irmãos de orientação, Ana Paula e Emerson por todas as conversas, trocas de experiências e todo o apoio recebido.

Ao Programa de Pós Graduação em Matemática, do Setor de Ciências Exatas da Universidade Federal do Paraná, na pessoa de sua coordenadora Profa. Dra. Elizabeth Wegner Karas e também orientadora deste trabalho, pelos muitos momentos de discussões, apoio, tempo, esforço e dedicação empregados.

Aos professores do programa que foram grandes incentivadores de meu crescimento pessoal e profissional.

À professora Lucelina por ter coorientado este trabalho, obrigado pela dedicação, sugestões, correções, paciência e principalmente pela amizade que construímos ao longo dos anos de estudos.

Aos professores Valeriano e Fernanda por aceitarem o convite para participar da banca examinadora e pelas contribuições que ajudaram muito a melhorar a versão final deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

*“It is not paradox to say that
in our most theoretical moods
we may be nearest to our
most practical applications.”*

Alfred North Whitehead

RESUMO

Neste trabalho, estudamos uma caracterização para otimalidade global de problemas mono e multiobjetivos, que difere da tradicional pois se baseia na Teoria da Medida e Integração. Esta abordagem é interessante pois se aplica a uma certa classe de problemas cuja função objetivo não é necessariamente diferenciável (e nem mesmo contínua). Baseado nesta abordagem, Zheng propõe um algoritmo para a otimalidade global para problemas mono objetivo. A contribuição desta dissertação reside na extensão destes resultados para problemas multiobjetivos através do método da escalarização de Chebyshev. Propomos um algoritmo, baseado na caracterização integral para problemas multiobjetivos. Ilustramos e comprovamos a eficiência do algoritmo em uma coleção de problemas-teste mono e multiobjetivo.

Palavras-chave: Otimalidade Global Via Integração. Otimização Mono Objetivo. Otimização Multiobjetivo.

ABSTRACT

In this work, we study a characterization for the global optimality of mono and multi-objective problems, which differs from the traditional approach because it is based on the Theory of Measurement and Integration. This characterization is interesting since it can be applied to a certain class of problems whose objective function is not necessarily differentiable (or even continuous). Based on this approach, Zheng proposed an algorithm for global optimality of mono objective problems. The contribution of this dissertation resides in the extension of these results to multiobjective problems through the method of Chebyshev scalarization. We propose an algorithm, based on integral characterization for multiobjective problems. We illustrate and prove the efficiency of such algorithms in a collection of mono and multiobjective test problems.

Keywords: Integral Global Optimality. Mono-objective Optimization. Multi-objective Optimization.

LISTA DE TABELAS

3.1	Resultados obtidos com o Algoritmo 3.1 para $m = 100$ rodadas.	79
3.2	Resultados obtidos com o Algoritmo 3.2 para $m = 100$ rodadas.	84
3.3	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 2$, $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$. . .	96
3.4	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 3$, $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0274	96
3.5	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 4$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0169	97
3.6	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 5$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0141	97
3.7	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.7 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.0486e - 12$	98
3.8	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.8 com $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.3406e - 16$	99
3.9	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.9 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $5.1344e - 05$	101
3.10	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.10 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$	102
3.11	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.11 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 55.2526	103
3.12	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.12 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$	104
3.13	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.13 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $7.6898e - 13$	105
3.14	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.14 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.0988e - 11$	106
3.15	Parâmetros para o Problema 3.15 com $m = 4$ e $n = 3$	106
3.16	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.15, $N = 100$ e $t = 20$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0082	107
3.17	Parâmetros para o Problema 3.15 com $m = 4$ e $n = 6$	107

3.18	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.17, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0403.	108
3.19	Parâmetros para o Problema 3.16	108
3.20	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 5$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0161.	109
3.21	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 7$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $8.4984e - 04$	109
3.22	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 10$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0226.	110
3.23	Parâmetros para o Problema 3.17.	110
3.24	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.17 com os parâmetros dispostos na Tabela 3.23, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.1820e - 07$	111
3.25	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.5745e - 11$	111
3.26	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.4030e - 06$	112
3.27	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 4$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0012.	112
3.28	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $3.1435e - 11$	113
3.29	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $3.2227e - 11$	114
3.30	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$	115
3.31	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$	115
3.32	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.5735e - 11$	116
3.33	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 3$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.8601e - 11$	116

3.34	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.9471e - 11$	117
3.35	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $5.8077e - 11$	117
3.36	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.9937e - 11$	118
3.37	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 3$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $6.8880e - 11$	119
3.38	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 4$, $N = 20$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.5212e - 10$	119
3.39	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $4.1925e - 11$	120
3.40	Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.2501e - 09$	121

LISTA DE FIGURAS

1.1	Minimizador global e local	23
1.2	Valor ótimo com tolerância ε	24
1.3	Função semicontínua inferior	24
1.4	Função semicontínua inferior	25
1.5	Função coerciva	27
1.6	Exemplos de Cone	28
1.7	Exemplos de Cone Polar	29
1.8	Exemplos de cones Viáveis Linearizados	30
1.9	Exemplos de Cones Tangentes	30
1.10	Ilustração Geométrica da condição de KKT	32
1.11	Interpretação geométrica da condição de KKT	33
1.12	Região viável e curvas de nível da função objetivo	36
1.13	Exemplos de pontos ótimos de Pareto.	40
1.14	Exemplos de pontos ótimos de Pareto Fraco.	40
1.15	Conjunto viável X	41
1.16	Exemplo 1.51	42
2.1	Gráfico das funções f e ξ	56
3.1	Comparação entre o método de integração por retângulos e por Monte Carlo	79
3.2	Comparação entre a média e o valor da integral para $N=100$	80
3.3	Comparação entre a média e o valor da integral para $N=1000$	81
3.4	Comparação entre a média e o valor da integral para $N=10000$	81
3.5	Comparação entre a média e o valor da integral para $N=100000$	82
3.6	Comparação entre a média e o valor da integral para $N=100$	84
3.7	Comparação entre a média e o valor da integral para $N=1000$	85
3.8	Comparação entre a média e o valor da integral para $N=2500$	85
3.9	Comparação entre a média e o valor da integral para $N=5000$	86
3.10	Cuboide inicial e discretização.	93
3.11	Malha e amostra.	93
3.12	Cuboide que contém os t menores valores de f	94
3.13	Problema 3.6	95

3.14 Problema 3.7	98
3.15 Problema 3.8	99
3.16 Problema 3.9	100
3.17 Problema 3.10	101
3.18 Problema 3.11	102
3.19 Problema 3.12	103
3.20 Problema 3.13	104
3.21 Problema 3.14	105
3.22 Problema 3.26	124
3.23 Problema 3.27	125
3.24 Problema 3.28	125
3.25 Problema 3.28	126
3.26 Problema 3.30	127
3.27 Problema 3.31	127
3.28 Problema 3.32	128
3.29 Problema 3.33	129
3.30 Problema 3.34	129
3.31 Problema 3.35	130
3.32 Problema 3.36	131
3.33 Problema 3.36	131
3.34 Problema 3.38	132
3.35 Problema 3.39	133
3.36 Problema 3.40	133
3.37 Problema 3.41	134
3.38 Problema 3.42	135
3.39 Problema 3.43	136
3.40 Problema 3.44	136
3.41 Problema 3.45	137
3.42 Problema 3.46	137
3.43 Problema 3.47	138

LISTA DE NOTAÇÕES

Os principais símbolos utilizados nesta dissertação são os seguintes:

X	Conjunto viável.
$\nabla f(x)$	Gradiente da função f em x .
$F = (f_1, f_2, \dots, f_r)$	Vetor de objetivos.
$y^1 \succ y^2$	y^1 é preferível a y^2 .
$y^1 \sim y^2$	y^1 é indiferente a y^2 .
$y^1 < y^2$	$y_\ell^1 < y_\ell^2$ para todo $\ell = 1, \dots, r$.
$y^1 \leqslant y^2$	$y_\ell^1 \leq y_\ell^2$ para todo $\ell = 1, \dots, r$.
$y^1 \leq y^2$	$y^1 \leqslant y^2$ mas $y^1 \neq y^2$.
$\ \cdot\ $	Norma euclidiana.
$\text{int}(A)$	Interior do conjunto A .
$\text{cl}(A)$	Fecho do conjunto A .
$B(\bar{x}, \varepsilon)$	Bola aberta centrada em $\bar{x}$ de raio ε , $\{x \in \mathbb{R}^n \mid \ x - \bar{x}\ < \varepsilon\}$.
$\mathcal{V}(A)$	Volume n -dimensional do conjunto $A \subseteq \mathbb{R}^n$.
$c_k \downarrow c$	Sequência decrescente que converge para c quando $k \rightarrow \infty$.
H_c	Conjunto de nível, $\{x \in \mathbb{R}^n \mid f(x) \leq c\}$.

Outras notações serão introduzidas ao longo do texto quando se fizerem necessárias.

SUMÁRIO

INTRODUÇÃO	19
1 Caracterização de Otimalidade Via Diferenciabilidade	22
1.1 Problema Mono Objetivo	22
1.1.1 Existência de Soluções	23
1.1.2 Condições Necessárias e Suficientes de Karush-Kuhn-Tucker	28
1.1.3 Condições de Otimalidade de Segunda Ordem	37
1.2 Problema Multiobjetivo	39
1.2.1 Existência de Soluções	42
1.2.2 Escalarização	44
1.2.3 Condições Necessárias e Suficientes de Karush-Kuhn-Tucker	46
1.2.4 Condições de Otimalidade de Segunda Ordem	47
2 Caracterização de Otimalidade Via Integração	52
2.1 Problema Mono Objetivo	52
2.1.1 Caso em que a função objetivo é contínua	53
2.1.2 Caso em que a função objetivo é descontínua	69
2.2 Problema Multiobjetivo	71
2.2.1 Caso em que as funções objetivo são contínuas	71
2.2.2 Caso em que as funções objetivo podem ser descontínuas	74
3 Algoritmos e Testes Numéricos	76
3.1 Método de Monte Carlo	76
3.1.1 Caso unidimensional	76
3.1.2 Caso multidimensional	80
3.2 Problema Mono objetivo	86
3.2.1 Problema irrestrito	86
3.2.2 Problema restrito	88
3.2.3 Testes Numéricos	95
3.3 Problema Multiobjetivo	121
3.3.1 Algoritmo	121
3.3.2 Testes Numéricos	123

Conclusão

139

REFERÊNCIAS

140

INTRODUÇÃO

Em problemas de diversas áreas do conhecimento, como em Física, Química, Engenharia, Economia e tantas outras, faz-se necessária a tomada de decisões, como por exemplo a minimização de custos e/ou maximização de lucros. Nessa linha de pensamento, surge o processo de otimização, que consiste em definir variáveis, quantificá-las através de relações, descrever seu comportamento e, por fim, determinar o objetivo a ser otimizado.

É muito comum depararmos com problemas em nosso cotidiano que tenham mais de um objetivo a ser alcançado. Por exemplo uma empresa deseja obter o maior lucro sobre a venda de determinado produto, porém é responsabilidade social das empresas tomar medidas para evitar todos os tipos de poluição, ou seja, tal empresa deve também reduzir o impacto ambiental causado pelo processo de fabricação do mesmo. Tipicamente, problemas de otimização multiobjetivos, ou seja, problemas que possuem mais de um objetivo a ser alcançado, surgem em áreas como Gerenciamento de Empresas, Administração de Recursos, Planejamento de Estratégias, etc. Assim é grande a aplicabilidade dessa classe de problemas. Daí nosso interesse em estudá-lo.

Existem inúmeras situações do dia a dia que precisam ser escritas como problemas multiobjetivos. A fim de simplificar a solução destes problemas os reescrevemos como um problema mono objetivo (i. e., com um único objetivo), utilizando-se técnicas de escalarização, a qual será de fundamental importância neste trabalho.

Geralmente, não é possível minimizar simultaneamente todos os objetivos, o que leva diversas noções de otimalidade para esta classe de problemas. A noção mais aceita de ótimo de problemas multiobjetivo foi proposta originalmente por Francis Ysidro Edgeworth em 1881 e generalizada posteriormente por Vilfredo Pareto em 1896. Alguns autores chamam de ótimo Edgeworth-Pareto, porém é mais comum denominá-lo, simplesmente, de ótimo de Pareto. Uma solução de um problema de otimização multiobjetivo é chamada eficiente, ou Pareto ótima, quando não é possível melhorar nenhum objetivo sem que haja uma deterioração em outro, o conjunto de soluções é denominado conjunto de Pareto, e o conjunto das imagens das soluções é denominado Fronteira de Pareto. Esta terminologia é devida a Pareto, um dos pioneiros no campo da Otimização Multiobjetivo, que abordou o tema em seu notável trabalho “*Cours d’Economie Politique*” [45].

Depois da Segunda Guerra Mundial (época que coincide com o auge da Pesquisa Operacional) um grande número de trabalhos surgiram nesta área, devido à importância de sua aplicabilidade. A principal concentração dos estudos foram nas condições necessárias e suficientes para determinação de pontos ótimos tanto para problemas mono objetivos quanto problemas multiobjetivos. Desta época se destacam os trabalhos de Fritz John [26], Harold W. Kuhn [31], Albert W. Tucker [32], Charnes e

Cooper [10], Gale [20] e de Karlin [28].

As primeiras condições de otimalidade para tratar de problemas mono objetivos foram estabelecidas por Fritz John [26] em 1948 e alguns anos depois por Kuhn-Tucker [32] em 1951. Mais tarde foi descoberto que tais condições já haviam sido estabelecidas por W. Karush em 1939 na sua dissertação, porém nunca foi publicada, mas resultados similares foram obtidos por Kuhn [31]. Assim as condições de Kuhn-Tucker passaram a ser chamadas de condições de Karush-Kuhn-Tucker ou simplesmente KKT.

Para desenvolver as condições KKT para problemas de otimização não linear, é necessário estudar o comportamento da função objetivo e a estrutura do conjunto viável numa vizinhança de uma solução do problema, ou seja, caracterizamos a otimalidade para problemas mono objetivos em termos dos cones aproximantes (cone linearizado e cone tangente). As condições de otimalidade KKT são necessárias somente se uma condição específica é satisfeita. Esta condição, chamada de *condição de qualificação da restrição*, impõe uma restrição sobre as funções de restrição de um problema não linear, com o objetivo de eliminar certas irregularidades da fronteira da região viável, as quais invalidariam as condições KKT se os valores ótimos ocorressem nas mesmas, ou seja, condições de qualificação são exigências feitas sobre as funções que definem o conjunto viável para que um minimizador satisfaça as condições KKT e assim possa ser encontrado por diversos métodos de otimização existentes. Não se pode garantir que as condições KKT sejam válidas sem que se cumpra uma condição de qualificação. Ainda, é conhecido que condições de otimalidade, desempenham um papel importante na teoria da otimização. Além disso, existem inúmeros estudos sobre como caracterizar a otimalidade de um problema. No entanto, quase todas essas tentativas são de natureza local, e ainda tendem a exigir vários graus de diferenciabilidade, a menos que seja imposta alguma hipótese de convexidade ou convexidade generalizada. A busca de condições necessárias e suficientes para a otimalidade global sem exigir diferenciabilidade e/ou convexidade é um esforço importante e necessário. Para fins mais aplicados, é importante estabelecer condições que garantam a suficiência para a otimalidade. Em geral, estas condições exigem que as funções sejam convexas (ou convexas generalizadas) ou duas vezes continuamente diferenciáveis [1, 6, 7, 8, 21, 22, 36, 38, 44, 53]. De qualquer modo, tais hipóteses são bastante restritivas.

Todavia podemos nos deparar com problemas onde a função objetivo (uma ou mais, no caso multiobjetivo) sejam descontínuas, onde os métodos tradicionais de otimalidade não se aplicam, ver [11]. Nosso interesse se concentra em estudar algum método no qual não sejam impostas “muitas” hipóteses sob a função objetivo.

A abordagem apresentada por Falk [16] estabelece uma condição de otimalidade global a problemas mono objetivo recorrendo à teoria da medida e integração. As hipóteses do problema são simples, onde são exigidas apenas compacidade do conjunto viável com interior não vazio e continuidade da função objetivo. Esta abordagem é mais teórica, porém fornece uma maneira simples de verificar se um ponto é, ou não, um ponto de ótimo global.

A abordagem apresentada nos trabalhos [13, 23, 30, 55, 56, 57, 58, 60, 61] também recorrem à teoria de medida e integração, porém com um caráter mais aplicado. É interessante como o estudo das caracterizações de pontos ótimos globais são obtidas nestes trabalhos, pois em alguns deles não exigem nem mesmo continuidade da função objetivo, englobando uma classe ainda maior de problemas.

Todas as caracterizações que recorrem à Teoria de Medida e Integração citadas anteriormente foram propostas para problemas de otimização mono objetivo. A contribuição desta dissertação é estender tais caracterizações integrais para otimalidade global de problemas multiobjetivos. Até onde sabemos este tema ainda não havia sido tratado na literatura. Além dos resultados teóricos, propomos um algoritmo para resolver problemas multiobjetivos e apresentamos resultados numéricos preliminares.

Este trabalho está dividido em três capítulos:

No Capítulo 1, são apresentados conceitos básicos de soluções, existência de soluções, condições necessárias e suficientes de Karush-Kuhn-Tucker (KKT) e condições de segunda ordem para problemas de programação não linear, tanto para problemas mono objetivo quanto para problemas multiobjetivos.

O Capítulo 2 contém as contribuições teóricas desta dissertação. Neste capítulo é estudada a condição de otimalidade global para problemas mono objetivo via integração, proposta por Falk [16]. Com base em [23], são apresentados os conceitos de valor médio, variância e variância modificada de uma função contínua, possivelmente descontínua, sobre seus conjuntos de nível para caracterizar a otimalidade global de um problema sem e com restrições. Através de técnicas de escalarização, propomos as devidas extensões das caracterizações integrais para otimalidade global de problemas multiobjetivos. Diferentemente das condições de otimalidade apresentadas no Capítulo 1, as quais utilizavam algum grau de diferenciabilidade, estas apelam para o uso da Teoria de Medida e Integração.

No Capítulo 3 são apresentados experimentos numéricos para problemas mono e multiobjetivos. Inicialmente, discutimos um algoritmo baseado nas caracterizações integrais, vistas no Capítulo 2, para otimalidade global de problemas mono objetivo. As integrais envolvidas nesse algoritmo são calculadas pelo método de Monte Carlo. O algoritmo é então estendido para obtenção de pontos Pareto ótimos de problemas multiobjetivos. Esta extensão representa uma contribuição numérica dessa dissertação. Resultados numéricos preliminares destes algoritmos, implementados em Matlab, ilustram sua eficiência.

Por fim, apresentamos algumas conclusões e possibilidades de trabalhos futuros.

Capítulo 1

CARACTERIZAÇÃO DE OTIMALIDADE VIA DIFERENCIABILIDADE

Neste capítulo apresentamos alguns conceitos básicos de soluções, existência de soluções e condições necessárias e suficientes de otimalidade para problemas mono objetivo e multiobjetivo.

1.1 Problema Mono Objetivo

Nesta seção, discutimos a existência de soluções para problemas mono objetivo, que consiste em minimizar certa função objetivo f sujeita a restrições de igualdade e desigualdade. Os resultados desta seção podem ser encontrados em [2, 25, 47]. Consideremos o seguinte problema:

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{Sujeito a } g_i(x) \leq 0, \quad i \in I := \{1, \dots, m\}, \\ &\quad h_j(x) = 0, \quad j \in J := \{1, \dots, p\}. \end{aligned} \tag{P}$$

onde $f, g_i, h_j : \mathbb{R}^n \rightarrow \mathbb{R}$ são funções continuamente diferenciáveis. Consideremos

$$X := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad i = 1, \dots, m \text{ e } h_j(x) = 0, \quad j = 1, \dots, p\}.$$

Neste caso, X é chamado de conjunto viável (ou factível) de (P) e os pontos de X serão chamados de pontos viáveis (ou factíveis).

Ainda, diremos que uma restrição de desigualdade é ativa em $\bar{x} \in X$ se $g_i(\bar{x}) = 0$. Denotamos $I(\bar{x})$ como o conjunto de índices das restrições de desigualdades ativas em um ponto viável $\bar{x}$, isto é, $I(\bar{x}) = \{i \in I \mid g_i(\bar{x}) = 0\}$.

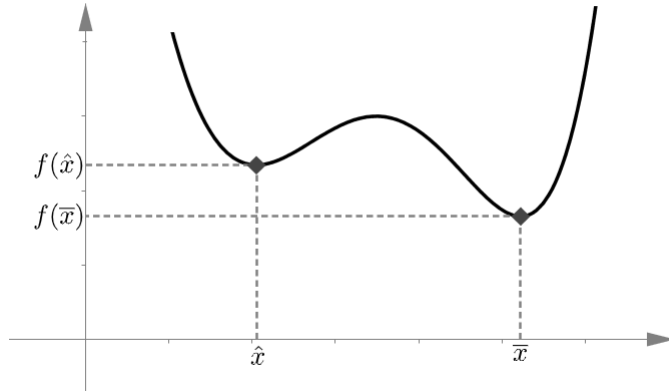
Definição 1.1. Considere uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e $\bar{x} \in X \subset \mathbb{R}^n$. Dizemos que $\bar{x}$ é um **minimizador local** de f em X quando existe $\delta > 0$, tal que $f(\bar{x}) \leq f(x)$ para todo $x \in B(\bar{x}, \delta) \cap X$. Caso $f(\bar{x}) \leq f(x)$, para todo $x \in X$, $\bar{x}$ é dito **minimizador global** de f em X .

Definição 1.2. Se define valor ótimo de (P) como:

$$v(P) = \inf_{x \in X} f(x)$$

onde $v(P) = +\infty$, se $X = \emptyset$.

A Figura 1.1 representa o gráfico de uma função real f decrescente à esquerda de $\hat{x}$ e crescente à direita de $\bar{x}$. Neste caso, o ponto $\hat{x}$ é ponto minimizador local e $\bar{x}$ é ponto minimizador global do problema de minimizar f em $\mathbb{R}$. O valor $y = f(\bar{x})$ é o valor mínimo do problema representado.



Fonte: O autor

Figura 1.1: Minimizador global e local

Definição 1.3. Definimos o conjunto das soluções ótimas de (P) com tolerância ε , $\hat{x}(\varepsilon)$, como:

$$\hat{x}(\varepsilon) = \{x \in X \mid f(x) \leq v(P) + \varepsilon\}.$$

Em particular, $f(\hat{x}(0)) = v(P)$.

Exemplo 1.4. Considere a função $f : [0, \infty) \rightarrow \mathbb{R}$, dada por $f(x) = x^2$. No caso da Figura 1.2, $\hat{x}(\varepsilon) = \{x \in \mathbb{R} \mid 0 \leq x \leq \sqrt{\varepsilon}\}$.

1.1.1 Existência de Soluções

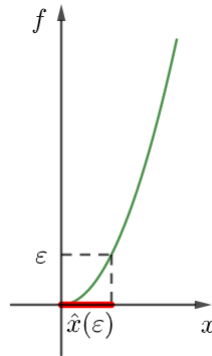
Veremos agora algumas condições de existência de soluções ótimas para o problema (P) . Para isso, necessitamos do seguinte conceito:

Definição 1.5 (Def. 5, [2]). 1. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é chamada **semicontínua inferior** em $\bar{x} \in \mathbb{R}^n$ se qualquer que seja $\varepsilon \geq 0$

$$f(x) > f(\bar{x}) - \varepsilon, \quad \forall x \in B(\bar{x}, \varepsilon). \quad (1.1)$$

2. Além disto, diremos que f é **semicontínua inferior** se for **semicontínua inferior** em todo ponto.

Observação 1.6. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se diz **semicontínua superior** em um ponto $\bar{x} \in \mathbb{R}^n$, se a função $(-f)$ é **semicontínua inferior** neste ponto.

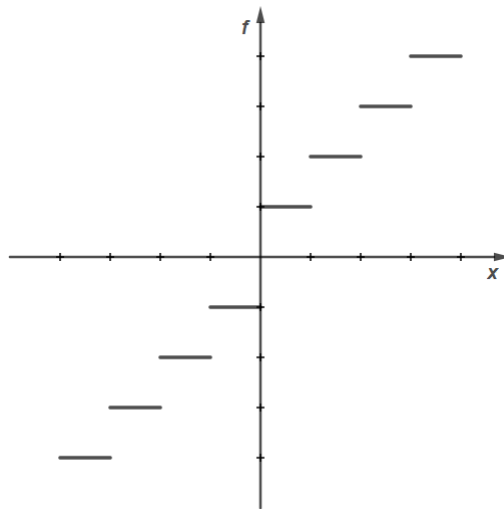


Fonte: O autor

Figura 1.2: Valor ótimo com tolerância ε

Vejamos dois exemplos de funções semicontínuas inferiores.

Exemplo 1.7. *Função teto de x , dada por $f(x) = \lceil x \rceil$, onde $\lceil x \rceil$ denota o menor inteiro ou igual a x . A Figura 1.3 exibe o gráfico de f .*



Fonte: O autor

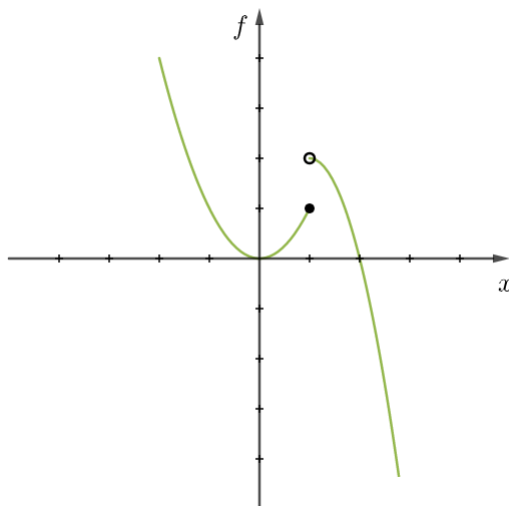
Figura 1.3: Função semicontínua inferior

O exemplo a seguir mostra que semi-continuidade inferior não implica continuidade.

Exemplo 1.8. *Seja a função $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por*

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

Temos que f é semicontínua inferior em $x = 1$, entretanto, pelo gráfico dado na Figura 1.4, podemos



Fonte: O autor

Figura 1.4: Função semicontínua inferior

observar que o limite lateral inferior tende a 1 e o superior a 2, logo f é descontínua em $x = 1$.

Claramente toda função contínua é, em particular, semicontínua inferior. Consequentemente, uma função contínua também é semicontínua superior. Veremos agora uma caracterização de funções semicontínuas inferiores utilizando sequências.

Proposição 1.9 (Prop. 1, [2]). *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é semicontínua inferior em $\bar{x} \in \mathbb{R}^n$, se, e somente se, para qualquer sequência x^k , convergente a $\bar{x}$, tem-se $f(\bar{x}) \leq \liminf_{k \rightarrow \infty} f(x^k)$.*

Vejamos outra caracterização a qual é determinada pelos conjuntos de nível.

Proposição 1.10 (Prop. 2, [2]). *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é semicontínua inferior se e somente se para qualquer $c \in \mathbb{R}$ o conjunto de nível $H_c = \{x \in \mathbb{R}^n | f(x) \leq c\}$ é fechado (eventualmente vazio) em $\mathbb{R}^n$.*

Algumas consequências interessantes da Proposição 1.10 podem ser observadas: Seja C um conjunto fechado e não vazio de $\mathbb{R}^n$ ($C \subseteq \mathbb{R}^n$). Então:

1. Se o conjunto viável X de (P) está definido como $X := \{x \in C | g_i(x) \leq 0\}$ onde as funções g_i são semicontínuas inferiores então X é um conjunto fechado do $\mathbb{R}^n$.
2. Se o conjunto viável X de (P) está definido como $X := \{x \in C | h_j(x) = 0\}$, onde as funções h_j são contínuas, então X é um conjunto fechado do $\mathbb{R}^n$.

3. Se o conjunto viável X de (P) está definido como $X := \{x \in C | g_i(x) \leq 0, h_j(x) = 0\}$, onde as funções g_i são semicontínuas inferiores e as funções h_j são contínuas, então X é um conjunto fechado do $\mathbb{R}^n$.

Propriedade 1.11 (Propr. 2, [2]). *Se a função objetivo f de (P) é semicontínua inferior (ou contínua) e se o conjunto viável X de (P) é fechado, o conjunto $\hat{x}(\varepsilon)$ de soluções ótimas de (P) com tolerância ε é fechado, qualquer que seja $\varepsilon \geq 0$.*

A seguir temos a definição de conjunto convexo.

Definição 1.12. *Seja $C \subset \mathbb{R}^n$ dizemos que C é convexo se*

$$x, y \in C \text{ e } t \in [0, 1] \Rightarrow ((1-t)x + ty) \in C,$$

ou seja, C é convexo se dados dois pontos quaisquer de C , o segmento que une esses pontos está contido em C .

Vejamos a definição de quase-convexidade de uma função.

Definição 1.13. *Se C é um conjunto convexo de $\mathbb{R}^n$, então uma função $\theta : C \rightarrow \mathbb{R}$ se diz estritamente quase-convexa sobre C , se:*

$$x, y \in C, x \neq y; z \in]x, y[^1 \Rightarrow f(z) < \max\{f(x), f(y)\}. \quad (1.2)$$

Teorema 1.14 (Propr. 7, [2]). *Considere o problema (P) com X convexo e f estritamente quase-convexa sobre X . Se $\bar{x}$ é ponto minimizador local de (P) , então $\bar{x}$ é ponto minimizador global único de (P) .*

Note que funções convexas são, em particular, quase-convexas. Um problema de otimização quase convexa pode ter soluções ótimas locais que não são soluções ótimas globais, isso não ocorre em problemas convexas.

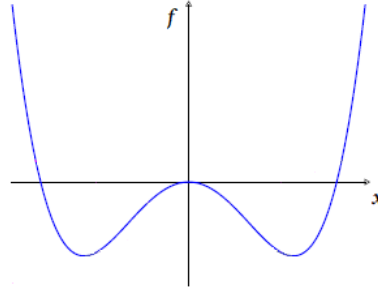
Observação 1.15. *Suponha que X é convexo e f é convexa em X . Se $\bar{x}$ é minimizador local de (P) , então é minimizador global de (P) .*

A seguir temos a definição de uma classe de funções para a qual é possível garantir a existência de minimizador global, sem a suposição de compacidade.

Definição 1.16. *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ se diz coerciva sobre X se para toda sequência de pontos viáveis $x^k \in X$ tal que $\lim_{k \rightarrow \infty} \|x^k\| = \infty$ então $\lim_{k \rightarrow \infty} f(x^k) = \infty$.*

¹ $z \in]x, y[\Leftrightarrow z = (t-1)y + tx$, para $t \in]0, 1[$.

Exemplo 1.17. A função $f : \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(x) = x^4 - x^2$ é claramente coerciva.



Fonte: O autor

Figura 1.5: Função coerciva

Teorema 1.18 (Teo. 1, [2]). Se a função objetivo f de (P) é semicontínua inferior e se o conjunto viável X de (P) é fechado não vazio, então, sob qualquer uma das seguintes hipóteses

- X é limitado;
- f coerciva sobre X ;

o conjunto das soluções ótimas de (P) é compacto não vazio de $\mathbb{R}^n$.

Observação 1.19. Note que o Teorema de Weierstrass é consequência imediata do Teorema 1.18. Além disso, temos também como consequência imediata do Teorema 1.18, que se existe $c \in \mathbb{R}$ tal que o conjunto H_c é compacto não vazio, então f tem um minimizador global.

A seguir veremos que, coercividade pode ser caracterizada pelo conjunto de curvas de nível da função objetivo.

Proposição 1.20 (Prop. 4, [2]). Uma função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é coerciva sobre X se, e somente se, H_c é limitado, qualquer que seja $c \in \mathbb{R}$, onde

$$H_c = \{x \in X \subset \mathbb{R}^n \mid f(x) \leq c\}.$$

As próximas propriedades relacionam coercividade com a limitação do conjunto viável e o conjunto de soluções ótimas de (P) .

Propriedade 1.21 (Prop.5, [2]). Se o conjunto viável X de (P) está definido como

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i \in I\}$$

onde ao menos uma das funções g_i é coerciva, então X é limitado.

Propriedade 1.22 (Propr. 6, [2]). *Se a função objetivo f de (P) é coerciva sobre X , então qualquer que seja $\varepsilon \geq 0$, o conjunto $\hat{x}(\varepsilon)$ das soluções ótimas de (P) , com tolerância ε , é limitado (eventualmente vazio).*

O próximo teorema garante a existência de minimizador global em $\mathbb{R}^n$, sem supor compacidade, apenas supondo coercividade e continuidade da função.

Teorema 1.23 (Teo. 2.6, [47]). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função contínua e coerciva. Então, f tem um minimizador global.*

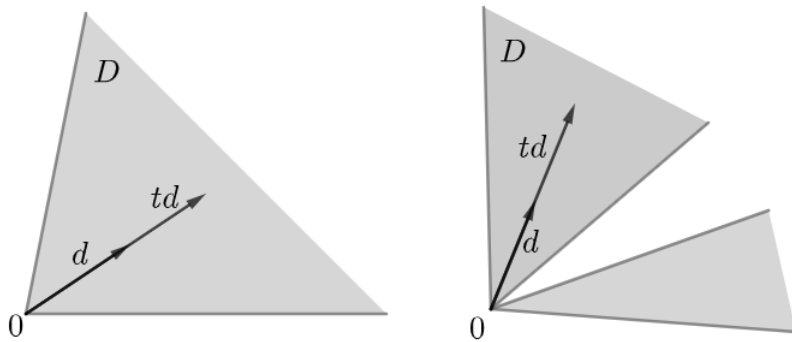
1.1.2 Condições Necessárias e Suficientes de Karush-Kuhn-Tucker

Nesta seção apresentaremos as condições necessárias e suficientes do tipo Karush-Kuhn-Tucker (KKT) para que um ponto viável seja minimizador local. Veremos que, sob certas condições chamadas condições de qualificação, podemos garantir que todo minimizador local é um ponto KKT. Além disso, veremos critérios suficientes de otimalidade, ou seja, através de certas condições podemos garantir que alguns pontos ou todos os pontos satisfazendo as condições necessárias são ou não ótimos. Tais resultados serão obtidos através das propriedades de convexidade (ou de convexidade generalizada) nas funções do problema (P) .

Iniciamos com as definições de cone e cone polar.

Definição 1.24. *Seja $C \subset \mathbb{R}^n$ um subconjunto não vazio, dizemos que C é um cone quando para todo $t \geq 0$ e $d \in C$ tem-se que $td \in C$.*

De maneira informal podemos dizer que um cone é um conjunto de direções. A Figura 1.6 mostra exemplos de cone convexo e não convexo, respectivamente.



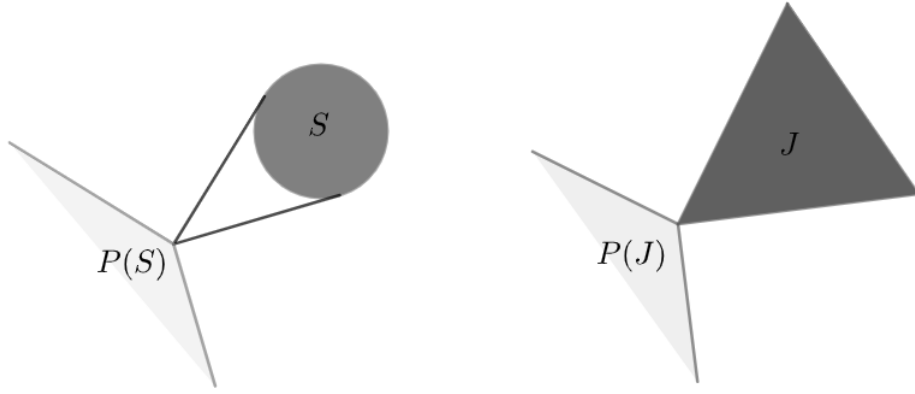
Fonte: O autor

Figura 1.6: Exemplos de Cone

Definição 1.25. Seja $S \subset \mathbb{R}^n$ um conjunto, definimos o **polar** de S como

$$\mathcal{P}(S) = \{p \in \mathbb{R}^n \mid p^T x \leq 0, \forall x \in S\}.$$

Podemos observar que o polar de um certo conjunto é o cone formado por todas as direções que formam ângulo reto ou obtuso com qualquer elemento do próprio conjunto. Na Figura 1.7 temos dois exemplos de cones polares.



Fonte: O autor

Figura 1.7: Exemplos de Cone Polar

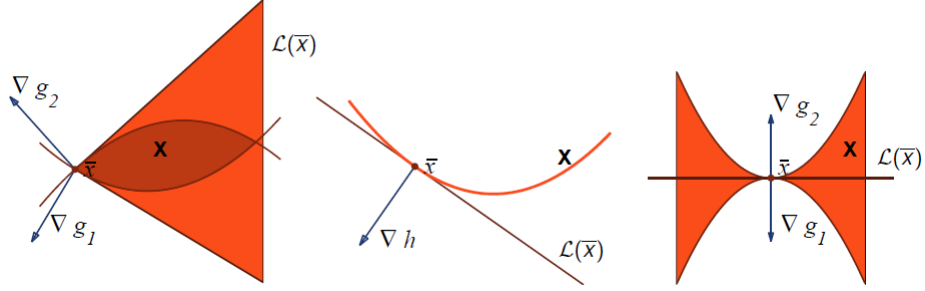
Agora discutiremos as principais aproximações do conjunto viável. A primeira forma de aproximar o conjunto viável X que veremos é o cone viável linearizado que se baseia nos gradientes das restrições, dado na seguinte definição:

Definição 1.26. Dado $\bar{x} \in X$, denotamos por $\mathcal{L}(\bar{x})$ o cone linearizado de X em torno de $\bar{x}$, por

$$\mathcal{L}(\bar{x}) = \{d \in \mathbb{R}^n \mid \nabla h_j(\bar{x})^T d = 0, \text{ se } j = 1, \dots, p \text{ e } \nabla g_i(\bar{x})^T d \leq 0, \text{ se } i \in I(\bar{x})\}.$$

Na Figura 1.8 temos três exemplos de cones linearizados. No primeiro, temos restrições de desigualdades, e os gradientes ativos são linearmente independentes. Podemos então dizer que o conjunto viável é de certa forma “bem” aproximado por $\mathcal{L}(\bar{x})$ em uma vizinhança de $\bar{x}$, a independência dos gradientes ativos confere uma certa regularidade ao conjunto viável. No segundo, temos apenas uma restrição de igualdade e também podemos dizer que, localmente, $\mathcal{L}(\bar{x})$ é uma boa aproximação para o conjunto viável X na vizinhança de $\bar{x}$. Já no último exemplo, temos que o cone é uma reta, porém o conjunto viável é uma região do plano. Note que, neste caso, os gradientes ativos são paralelos, ou seja, linearmente dependentes.

Outra forma de aproximar o conjunto viável X , é o cone que considera os vetores que tangenciam ou “penetram” em X .



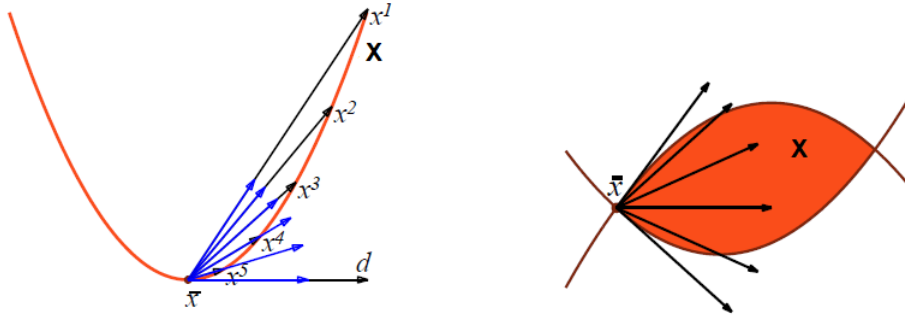
Fonte: Adaptado de [47]

Figura 1.8: Exemplos de cones Viáveis Linearizados

Definição 1.27. Dado $\bar{x} \in X$, denotamos por $\mathcal{T}(\bar{x})$ o cone tangente de X em torno de $\bar{x}$, por

$$\mathcal{T}(\bar{x}) = \{0\} \cup \left\{ d \in \mathbb{R}^n \mid \exists x^k \in X, x^k \rightarrow \bar{x}, \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \rightarrow \frac{d}{\|d\|} \right\}.$$

Na Figura 1.9 é ilustrado este conceito. Na esquerda o conjunto viável é uma curva definida por uma restrição de igualdade, na qual é representada uma direção tangente d e a convergência indicada na definição. Na outra ilustração o conjunto viável é uma região determinada por duas restrições de desigualdade. Nesta figura aparecem algumas direções tangentes. Note que uma direção que “penetra” no conjunto viável também satisfaz a Definição 1.27.



Fonte: Adaptado de [47]

Figura 1.9: Exemplos de Cones Tangentes

É imediato da definição que o conjunto das direções tangentes em $\bar{x}$ é, de fato, um cone com vértice em 0. O cone tangente $\mathcal{T}(\bar{x})$ é frequentemente usado como aproximação para o conjunto viável em torno de um ponto, de tal forma que ele é caracterizado como o conjunto de limites de sequências de direções pelas quais é possível dar passos sem sair do conjunto viável.

O próximo lema estabelece a relação entre os cones $\mathcal{T}(\bar{x})$ e $\mathcal{L}(\bar{x})$.

Lema 1.28 (Lema 6.19, [47]). *Dado $\bar{x} \in X$, temos $\mathcal{T}(\bar{x}) \subset \mathcal{L}(\bar{x})$.*

Uma condição necessária de otimalidade é que em um minimizador toda direção tangente é de não decrescimento, como pode ser visto no teorema a seguir.

Teorema 1.29 (Teo. 6.20, [47]). *Se $\bar{x} \in X$ é um minimizador local do problema (P), então*

$$\nabla^T f(\bar{x})d \geq 0,$$

para todo $d \in \mathcal{T}(\bar{x})$.

Para que KKT seja uma condição de otimalidade é preciso assumir alguma hipótese adicional às restrições do problema. Qualquer hipótese com essa propriedade é chamada de condição de qualificação. Consideremos a seguinte condição de qualificação:

Definição 1.30 (Condição de Qualificação de Guignard (GCQ)). *A condição de qualificação de Guignard é satisfeita em $\bar{x}$ quando ocorre a igualdade*

$$\mathcal{P}(\mathcal{T}(\bar{x})) = \mathcal{P}(\mathcal{L}(\bar{x})). \quad (1.3)$$

Além disto, pode ser demonstrado que esta é a condição de qualificação mais geral. Veja Gould e Tolle [21].

Teorema 1.31 (Teo. 6.21, [47]). *Seja $\bar{x} \in X$ um minimizador local do problema (P), e suponha que (GCQ) é válida, então existem $\mu_i, \lambda_j \in \mathbb{R}$ tais que*

$$\nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^p \lambda_j \nabla h_j(\bar{x}) = 0 \quad (1.4)$$

$$\mu_i \geq 0, \quad i = 1, \dots, m \quad (1.5)$$

$$\mu_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m. \quad (1.6)$$

Os pontos viáveis $\bar{x} \in X$ que cumprem (1.4)-(1.6) são chamados pontos estacionários (ou pontos KKT) do problema (P) e são os principais candidatos à solução quando o problema cumpre uma condição de qualificação e μ_i e λ_j são ditos multiplicadores de KKT.

Exemplo 1.32 (Exemplo 6.22, [47]). *Consideremos o seguinte problema:*

$$\begin{aligned} \text{Minimizar } f(x) &= (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{Sujeito a } g_1(x) &= x_1 + x_2 - 2 \leq 0 \\ g_2(x) &= x_1^2 - x_2 \leq 0. \end{aligned} \quad (\text{P}_1)$$

Note primeiro que o conjunto viável dado por $X = \{x \in \mathbb{R}^2 \mid g_i(x) \leq 0, i = 1, 2\}$ é compacto.

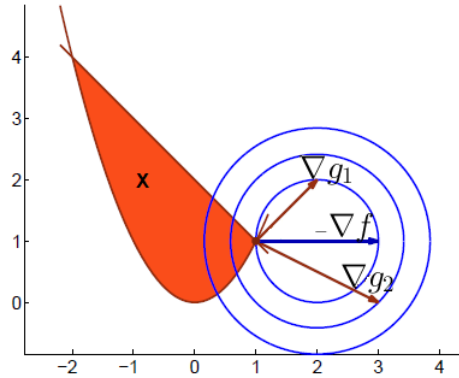
Como

$$x_1^2 \leq x_2 \leq 2 - x_1,$$

temos $x_1^2 + x_1 - 2 \leq 0$. Portanto $-2 \leq x_1 \leq 1$ e $0 \leq x_2 \leq 4$. Além disso, temos $\mathcal{T}(x) = \mathcal{L}(x)$ para todo ponto viável x . Assim, o minimizador deve satisfazer

$$-2 \begin{pmatrix} x_1 - 2 \\ x_2 - 1 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} \quad (1.7)$$

além de $\mu_i \geq 0$ e $\mu_i g_i(x) = 0, i = 1, 2$. Como nenhum ponto de X cumpre $x_1 = 2$, pelo menos um dos multiplicadores deve ser não nulo. Veremos agora que os dois são não nulos. De fato, se fosse $\mu_1 = 0$ e $\mu_2 > 0$, teríamos $x_1^2 - x_2 = 0$ (restrição ativa) e $x_2 > 1$. Assim, $x_1 \leq -1$, o que contradiz (1.7). Por outro lado, se $\mu_1 > 0$ e $\mu_2 = 0$, então $x_1 + x_2 = 2$ (restrição ativa) e $x_1 - 2 = x_2 - 1$ relação (1.7). Assim, $x_1 = \frac{3}{2}$, o que também é uma contradição. Agora fica fácil resolver o sistema KKT, pois $x_1 + x_2 = 2$ e $x_1^2 = x_2$ fornecem $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $\tilde{x} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$. Como $\tilde{x}$ não satisfaz (1.7) para $\mu_i \geq 0$, a solução é x^* com multiplicador $\mu^* = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix}$. A próxima Figura 1.10 ilustra este exemplo.



Fonte: Adaptado de [47]

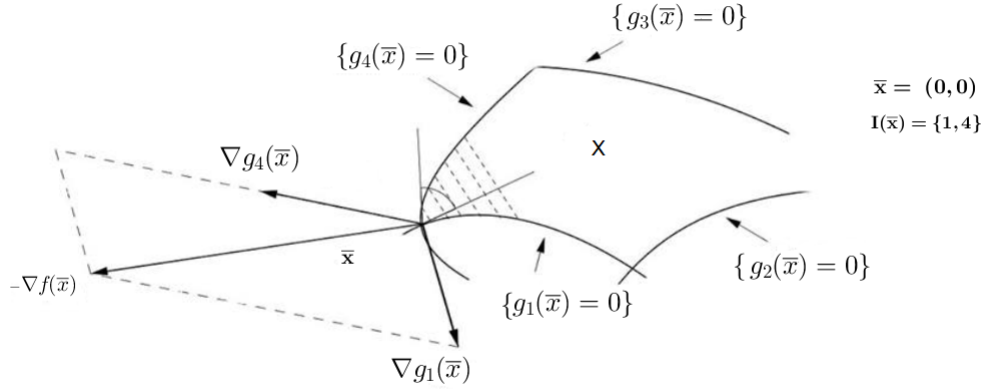
Figura 1.10: Ilustração Geométrica da condição de KKT

O próximo exemplo é dado em [2] como interpretação geométrica do Teorema 1.31.

Exemplo 1.33. Suponhamos que o conjunto viável X em $\mathbb{R}^2$ seja dado somente por restrições de desigualdade, ou seja

$$\begin{aligned} & \text{Minimizar } f(x) \\ & \text{Sujeito a } g_i(x) \leq 0, i = 1, \dots, 4. \end{aligned} \quad (P_2)$$

Na Figura 1.11 são traçadas as quatro restrições nos casos em que $g_i(x) = 0, i = 1, \dots, 4$ com $x \in \mathbb{R}^2$. Onde o conjunto de índices de restrições ativas em $\bar{x}$ está dado por $I(\bar{x}) = \{1, 4\}$.



Fonte: Adaptado de [24]

Figura 1.11: Interpretação geométrica da condição de KKT

Os gradientes $\nabla g_i(x)$ são ortogonais a suas respectivas funções $g_i(x)$ no ponto $\bar{x}$, onde os mesmos apontam na direção de máximo crescimento de $g_i(x)$ a partir de $\bar{x}$.

Se $\bar{x}$ é ponto de mínimo local do problema (P_2) , o Teorema 1.31 garante que existem multiplicadores $\mu_1, \mu_4 \geq 0$, tais que

$$-\nabla f(\bar{x}) = \mu_1 \nabla g_1(\bar{x}) + \mu_4 \nabla g_4(\bar{x})$$

isto é, $-\nabla f(\bar{x})$ pode ser escrito como combinação linear, com coeficientes positivos, dos gradientes das restrições ativas em $\bar{x}$.

As condições de qualificação são propriedades referentes a pontos de X , as quais dependem somente de X (e não da função objetivo) tais que a expressão

$$\text{Minimizador local} + \text{CQ} \Rightarrow \text{KKT}$$

sempre é verdadeira. Em outras palavras, sob qualquer condição de qualificação (CQ), todo minimizador local de (P) cumpre as condições de KKT. Observe que quanto mais fraca for a condição de qualificação, mais forte será a condição de otimalidade associada, deste modo, estamos interessados em condições de qualificação fracas.

Note que no Teorema 1.31, vimos que as condições KKT são necessárias para a otimalidade local sob uma condição de qualificação específica (Guignard). Entretanto, esta condição é de difícil averiguação. Por esta razão, consideram-se outras condições de qualificação, as quais, embora sejam menos gerais, são de mais fácil verificação. Para mais detalhes sobre este assunto, consultar [15].

Uma das condições de qualificação mais frequentes na literatura é a independência linear (LICQ) que afirma que os gradientes das restrições de igualdade e de desigualdade ativas são linearmente

independentes. Vejamos sua definição a seguir:

Definição 1.34 (Condição de Qualificação de Independência Linear). *Dizemos que a condição de qualificação de independência linear (LICQ) é satisfeita em $\bar{x}$ quando o conjunto formado pelos gradientes das restrições de igualdade e das restrições de desigualdade ativas é linearmente independente, isto é,*

$$\{\nabla h_j(\bar{x})\}_{j \in \{1, \dots, p\}} \cup \{\nabla g_i(\bar{x})\}_{i \in I(\bar{x})} \text{ é L.I..}$$

A condição de independência linear estabelece a existência e unicidade dos multiplicadores na solução do problema de (P) .

Definamos primeiro o conjunto de funções que tenham $\bar{x}$ como minimizador local do problema (P) , ou seja, $\mathcal{F} = \{f \in \mathcal{C}^1 : \bar{x} \text{ é ponto minimizador local de } (P)\}$. Além disso, para $f \in \mathcal{F}$ o conjunto de multiplicadores é dado por:

$$\Lambda(f) = \{(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}^m \times \mathbb{R}^p : (1.4) - (1.6) \text{ é satisfeito}\}.$$

Vejamos agora algumas propriedades do conjunto de multiplicadores, as quais podem ser encontradas com suas devidas demonstrações em [52]:

- (i) O conjunto $\Lambda(f)$ é fechado e convexo para toda $f \in \mathcal{F}$.
- (ii) O conjunto $\Lambda(f)$ é não vazio para toda $f \in \mathcal{F}$ se, e somente se, a condição de Guignard (1.3) é satisfeita.
- (iii) O conjunto de multiplicadores $\Lambda(f)$ é um conjunto unitário (*singleton*) para toda $f \in \mathcal{F}$ se, e somente se, LICQ é satisfeita.

Para um estudo detalhado sobre multiplicadores KKT, destacamos o trabalho de J. Gauvin [19] que provou que o conjunto de multiplicadores KKT é não vazio e limitado se a condição de qualificação de Mangasarian-Fromovitz (MFCQ) for válida, e ainda como vimos na Subseção 1.1.2, no trabalho de Wachsmuth [52] foi provado que a condição de qualificação de independência linear (LICQ) é uma condição necessária e suficiente para que o conjunto de multiplicadores seja um conjunto unitário (em inglês, *singleton*).

Agora, veremos critérios que garantam que as condições (1.4)-(1.6) também são suficientes para a otimalidade local. Tais resultados são obtidos para funções convexas e posteriormente generalizaremos a uma classe mais ampla de funções. Por simplicidade, consideramos o caso particular em que o problema (P) possui apenas restrições de desigualdade.

Teorema 1.35 (Teo. 1, [3]). *Considere que (P) possui apenas restrições de desigualdade. Seja $\bar{x} \in X$ e suponha que f e g_i sejam continuamente diferenciáveis e convexas. Se existe $\mu \in \mathbb{R}_+^m$ tal que satisfaz*

as condições

$$\nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) = 0, \quad (1.8)$$

$$\mu_i g_i(\bar{x}) = 0, \quad i = 1, \dots, m, \quad (1.9)$$

então $\bar{x}$ é uma solução ótima global do problema (P).

Uma noção de convexidade generalizada muito conhecida é a invexidade, a qual foi introduzida por Hanson em [22]. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é invexa se e somente se existe $\eta : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$f(x) - f(\bar{x}) \geq \nabla^T f(\bar{x}) \eta(x, \bar{x}), \quad \forall x, \bar{x} \in \mathbb{R}^n.$$

Assim sendo, toda função convexa diferenciável é invexa com $\eta(x, \bar{x}) = x - \bar{x}$.

Exemplo 1.36 (Hanson [22]). *Considere o problema a seguir:*

$$\begin{aligned} \text{Minimizar } f(x_1, x_2) &= x_1 - \text{sen } x_2 \\ \text{Sujeito a } g_1(x_1, x_2) &= \text{sen } x_1 - 4\text{sen } x_2 \leq 0, \\ g_2(x_1, x_2) &= 2\text{sen } x_1 + 7\text{sen } x_2 + x_1 - 6 \leq 0 \\ g_3(x_1, x_2) &= 2x_1 + 2x_2 - 3 \leq 0 \\ g_4(x_1, x_2) &= 4x_1^2 + 4x_2^2 - 9 \leq 0 \\ g_5(x_1, x_2) &= -\text{sen } x_1 \leq 0 \\ g_6(x_1, x_2) &= -\text{sen } x_2 \leq 0 \\ (x_1, x_2) &\in \mathbb{R}^2. \end{aligned} \quad (\text{P}_3)$$

A região viável descrita pelas restrições assim como algumas curvas de nível da função objetivo são exibidas na Figura 1.12.

Este problema claramente não é convexo. No entanto, ele é invexo com

$$\eta(x, y) = \left(\frac{\text{sen } x_1 - \text{sen } y_1}{\cos y_1}, \frac{\text{sen } x_2 - \text{sen } y_2}{\cos y_2} \right)$$

para todo $x = (x_1, x_2)$ e $y = (y_1, y_2)$ em X .

Temos que

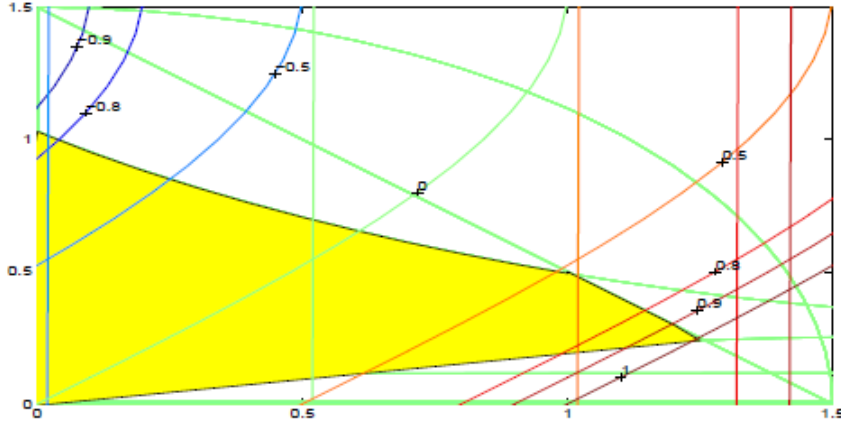
$$\nabla f(x_1, x_2) + \sum_{i=1}^6 g_i(x_1, x_2) = 0 \quad (1.10)$$

e é de fácil verificação que as condições de KKT (1.8)-(1.9), são satisfeitas por

$$\mu = [0, 1/7, 0, 0, 10/7, 0]^T \text{ e } \bar{x} = [0, \arcsen(6/7)]^T.$$

■

Observação 1.37. *Note que para aplicar o Teorema 1.35 não é necessário conhecer a função η . É suficiente saber que ela existe.*



Fonte: [41]

Figura 1.12: Região viável e curvas de nível da função objetivo

Uma função f diferenciável é invexa se e somente se todo ponto estacionário de f é minimizador global. Mais geralmente, se as funções forem todas invexas, ainda vale a suficiência das condições KKT (ver [22]). Entretanto, existem problemas de programação não linear cujos pontos KKT são ótimos mas não são invexos. Martin [38] introduziu uma classe de problemas, denominados como KT-invexos que é uma variação do conceito de invexidade para o problema (P) e é caracterizado pela seguinte propriedade fundamental: O problema (P) é KT-invexo se, e somente se, todo ponto que satisfaz as condições KKT é minimizador global de (P) .

Observe que se f é convexa e continuamente diferenciável, então f é invexa, com $\eta(x, \bar{x}) = x - \bar{x}$, porém a recíproca é falsa como pode ser visto no Exemplo 1.36. Note que se ξ é convexa e ϕ é invexa, com derivada inversível, então a composta $f = \xi \circ \phi$ é invexa. De fato, considere $\xi : \mathbb{R} \rightarrow \mathbb{R}$ e $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$. Tome $f = \xi \circ \phi$ e $x, \bar{x} \in \mathbb{R}^n$. Assim, temos que

$$\begin{aligned}
 f(x) - f(\bar{x}) &= \xi(\phi(x)) - \xi(\phi(\bar{x})) \\
 &\geq \xi'(\phi(\bar{x}))[\phi(x) - \phi(\bar{x})] \quad (f = \xi \circ \phi \Rightarrow f'(\bar{x}) = \xi'(\phi(\bar{x})) \cdot \phi'(\bar{x})) \\
 &= f'(\bar{x})[\phi'(\bar{x})]^{-1}[\phi(x) - \phi(\bar{x})] \\
 &= f'(\bar{x})\eta(x, \bar{x}).
 \end{aligned}$$

Desta desigualdade, segue que f é invexa, com $\eta(x, \bar{x}) = [\phi'(\bar{x})]^{-1}[\phi(x) - \phi(\bar{x})]$.

Portanto algumas funções invexas podem ser obtidas a partir de funções convexas por uma transformação adequada. Estas transformações fazem com que a propriedade de convexidade seja perdida, mas não a de invexidade.

Teorema 1.38 (Teo. 2.2, [22]). *Se $\bar{x} \in X$ um ponto KKT do problema (P) e suponha que f, g_i, h_j são funções invexas com respeito à mesma função η , então $\bar{x}$ é um minimizador global de (P) .*

Definição 1.39 (Martin, [38]). *O problema (P) é chamado KT-invexo se existe uma função $\eta : X \times X \rightarrow \mathbb{R}^n$ tal que*

$$\begin{aligned} f(x) - f(\bar{x}) &\geq \nabla^T f(\bar{x})\eta(x, \bar{x}) \\ -\nabla^T g_i(\bar{x})\eta(x, \bar{x}) &\geq 0, i \in I(\bar{x}) \end{aligned}$$

para quaisquer $x, \bar{x} \in X$.

A classe dos problemas KT-invexos é a maior classe de problemas para os quais as condições KKT são necessárias e também suficientes para a otimalidade. Esta propriedade pode ser demonstrada utilizando o Teorema de Alternativa de Motzkin [37]. Tal caracterização é bastante interessante, seja do ponto de vista teórico ou do ponto de vista das aplicações.

É imediato da definição anterior que todo problema invexo é KT-invexo. O seguinte exemplo apresentado por Martin [38] prova que a recíproca em geral é falsa.

Exemplo 1.40 (Exemplo 2.1, [38]). *Considere o problema*

$$\begin{aligned} &\text{Minimizar } f(x) \\ &\text{sujeito a } x \geq 0, x \in \mathbb{R}. \end{aligned} \tag{P_4}$$

onde suporemos que a função $f : \mathbb{R} \rightarrow \mathbb{R}$ é suave, com derivada positiva. Claramente, o único ponto KKT é o minimizador global $x = 0$. Se problema for invexo então f será convexa. De fato:

$$\begin{aligned} f(x) - f(\bar{x}) &\geq f'(\bar{x})\eta(x, \bar{x}), \\ -x + \bar{x} &\geq -\eta(x, \bar{x}), \forall x, \bar{x} \in \mathbb{R}. \end{aligned}$$

Como $f'(x) > 0$, temos que

$$f(x) - f(\bar{x}) \geq f'(\bar{x})(x - \bar{x}).$$

Entretanto, o problema é KT-invexo com $\eta(x, \bar{x}) = \frac{f(x) - f(\bar{x})}{f'(\bar{x})}$.

■

Teorema 1.41 (Teo. 2.1, [38]). *O problema (P) é KT-invexo se, e somente se, todo ponto KKT é minimizador global de (P) .*

1.1.3 Condições de Otimalidade de Segunda Ordem

Nesta seção, apresentamos condições necessárias e suficientes de otimalidade de segunda ordem (isto é, utilizando as informações das derivadas segundas) das funções do problema (P) . Assumiremos que estas são de classe $\mathcal{C}^2$.

Definamos primeiramente direções críticas em um ponto viável.

Definição 1.42 (Direções Críticas). Dizemos que uma direção $d \in \mathbb{R}^n$ é uma direção crítica para o problema (P) no ponto $\bar{x} \in X$ se

$$\begin{aligned}\nabla^T f(\bar{x})d &\leq 0, \\ \nabla^T g_i(\bar{x})d &\leq 0, \quad i \in I(\bar{x}), \\ \nabla^T h_j(\bar{x})d &= 0, \quad j \in J.\end{aligned}$$

Denotemos por $D_e(\bar{x})$ o conjunto de direções críticas no ponto $\bar{x}$ para o problema escalar (mono objetivo).

No primeiro resultado desta seção, que apresentamos a seguir, necessitamos da seguinte condição de qualificação.

Definição 1.43 (Condição de Qualificação Fraca de Segunda Ordem). O conjunto viável X satisfaz a condição de qualificação fraca de segunda ordem em $\bar{x}$, desde que, para qualquer $d \in D_e(\bar{x})$ não nula, se existe um vetor $a(d) \in \mathbb{R}^n$ tal que

$$\begin{aligned}d^T \nabla^2 g_i(\bar{x})d + \nabla^T g_i(\bar{x})a(d) &\leq 0, \quad i \in I(\bar{x}), \\ d^T \nabla^2 h_j(\bar{x})d + \nabla^T h_j(\bar{x})a(d) &= 0, \quad j \in J,\end{aligned}$$

então existe uma curva diferenciável $\alpha : [0, t_0] \rightarrow \mathbb{R}^n$ satisfazendo

$$\alpha(0) = \bar{x}, \quad \alpha'(0) = d, \quad \alpha''(0) = a(d) \quad e \quad \alpha(t) \in X, \quad \forall t \in [0, t_0].$$

Vejamos primeiramente uma condição necessária, e a seguir uma condição suficiente de segunda ordem.

Teorema 1.44 (Teo. 3.10, [1]). Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ duas vezes diferenciáveis em $\bar{x} \in \mathbb{R}^n$. Suponhamos que $\bar{x}$ seja uma solução local do problema (P) e que exista (μ, λ) satisfazendo as condições de KKT. Suponha ainda que a condição de qualificação de segunda ordem fraca é válida em $\bar{x}$. Então para toda d , que satisfaz

$$\begin{cases} \nabla^T g_i(\bar{x})d = 0, \quad i \in I(\bar{x}) \\ \nabla^T h_j(\bar{x})d = 0, \quad j \in J \end{cases}, \quad (1.11)$$

temos que

$$d^T \left[\nabla^2 f(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d \geq 0. \quad (1.12)$$

A seguir veremos uma condição suficiente de segunda ordem, em lugar de uma das condições de regularidade pode-se admitir a própria existência dos multiplicadores, o que é, uma hipótese mais fraca.

Teorema 1.45 (Teo. 4.3.7, [25]). *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$ duas vezes diferenciáveis em $\bar{x} \in \mathbb{R}^n$. Suponhamos que $\bar{x}$ é uma solução viável para o problema (P) e sejam $\mu \in \mathbb{R}_+^m$ e $\lambda \in \mathbb{R}^p$ tais que as condições de KKT sejam válidas. Nestas condições, se*

$$d^T \left[\nabla^2 f(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d > 0, \quad \forall d \in D_e(\bar{x}) \setminus \{0\}$$

então $\bar{x}$ é um minimizador local estrito² do problema (P).

Nesta seção discutimos a existência de soluções e condições necessárias e suficientes de otimalidade para problemas mono-objetivos. Na próxima seção, estudaremos resultados similares a esses, para o caso multiobjetivo.

1.2 Problema Multiobjetivo

Primeiramente, vamos considerar o seguinte problema de otimização multiobjetivo (ou vetorial):

$$\begin{aligned} &\text{Minimizar } F(x) = (f_1(x), \dots, f_r(x))^T \\ &\text{sujeito a } g_i(x) \leq 0, \quad i \in I, \\ &\quad h_j(x) = 0, \quad j \in J, \end{aligned} \tag{VOP}$$

onde $f_\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, $\ell \in L := \{1, \dots, r\}$, são as funções objetivo, $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i \in I := \{1, \dots, m\}$, e $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$, $j \in J := \{1, \dots, p\}$.

Note que o conjunto viável é o mesmo que foi dado no problema escalar

$$X := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \quad i \in I, \quad h_j(x) = 0, \quad j \in J\},$$

assim como o conjunto de índices ativos em um ponto $x \in \mathbb{R}^n$ relacionado às restrições de desigualdades $I(x) := \{i \in I \mid g_i(x) = 0\}$.

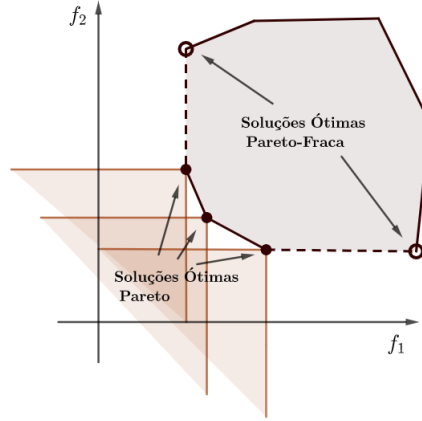
Porém, agora nosso objetivo é encontrar $\bar{x} \in X$ que satisfaça as restrições impostas sobre o problema e minimize as funções objetivo. A solução de um problema de otimização vetorial (VOP) é chamada eficiente, quando não é possível melhorar nenhum objetivo sem que haja uma deterioração em outro, e o conjunto de soluções é denominado conjunto de Pareto. Esta terminologia é devida a Pareto, um dos pioneiros no campo da Otimização Multiobjetivo, que abordou o tema em seu notável trabalho “*Cours d'Economie Politique*” [45].

Definição 1.46 (Solução de Pareto). *Um ponto $\bar{x} \in X$ é dito ser uma solução de Pareto (ou solução eficiente) para o problema (VOP) se não existe $x \in X$, $x \neq \bar{x}$ tal que $f_\ell(x) \leq f_\ell(\bar{x})$ para todo $\ell = 1, \dots, r$ com a desigualdade estrita para algum ℓ . Ou ainda, $\bar{x} \in X$ é considerado um ponto ótimo de Pareto se*

² $\bar{x}$ é um minimizador local estrito de f se existe $\delta > 0$ tal que $f(\bar{x}) < f(x)$ para todo $x \in B(\delta, \bar{x}) \cap X$.

não existir outro vetor $x \in X$ satisfazendo $F(x) - F(\bar{x}) \in -\mathbb{R}_+^r \setminus \{0\}$. Denotamos por P o conjunto de ótimos de Pareto.

Exemplos de pontos ótimos de Pareto podem ser observados na Figura 1.13.

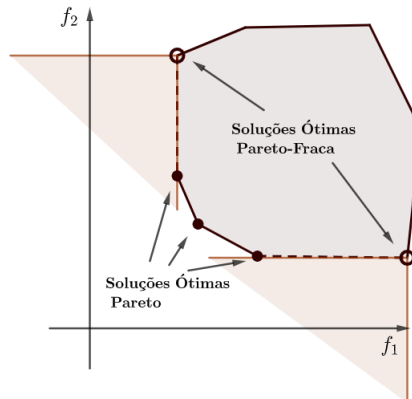


Fonte: O autor

Figura 1.13: Exemplos de pontos ótimos de Pareto.

Definição 1.47 (Solução de Pareto fraca). Um ponto $\bar{x} \in X$ é uma solução de Pareto fraca (ou fracamente eficiente) para o problema (VOP) se não existe $x \in X$, tal que $f_\ell(x) < f_\ell(\bar{x})$, para todo $\ell = 1, \dots, r$. Ou ainda, $\bar{x} \in X$ é dito ser um ponto de Pareto fraco se não existir outro vetor $x \in X$ tal que $F(x) - F(\bar{x}) \in -\text{int}(\mathbb{R}_+^r)$. Denotamos por P_f o conjunto de ótimos Pareto fraco.

Exemplos de pontos ótimos de Pareto Fraco podem ser observados na Figura 1.14.



Fonte: O autor

Figura 1.14: Exemplos de pontos ótimos de Pareto Fraco.

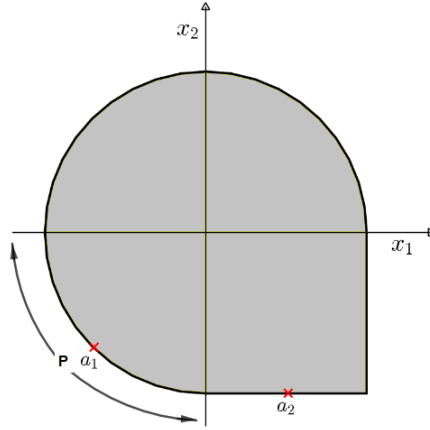
Observação 1.48. Para problemas de maximização multiobjetivo, os correspondentes conceitos de solução são obtidos revertendo-se os sentidos das desigualdades.

Das definições anteriores temos que eficiência (Pareto) implica eficiência fraca (Pareto fraca). A recíproca é falsa, como podemos ver no seguinte exemplo:

Exemplo 1.49 (Adaptado de Moulin e Soulié [42]).

$$\begin{aligned} & \text{Minimizar } F(x) = (f_1(x), f_2(x)) \\ & \text{sujeito a } x \in X \end{aligned} \tag{P_5}$$

onde $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ com $f_1(x) = x_1$ e $f_2(x) = x_2$ e $X = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\} \cup [0, 1] \times [0, -1] \subseteq \mathbb{R}^2$.



Fonte: Adaptação de [42]

Figura 1.15: Conjunto viável X .

- O ponto $a_1 = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ não pode melhorar estritamente f_1 , sem diminuir estritamente f_2 . O ponto a_1 é solução eficiente pertencente ao conjunto P , o qual é o conjunto das soluções eficientes, sendo P o arco compreendido entre $(-1, 0)$ e $(0, -1)$.
- Note que $a_2 = (\frac{1}{2}, -1)$ é solução fracamente eficiente de (P_5) e que o conjunto P_f definido por $P_f = \{P \cup (\lambda, -1) : 0 \leq \lambda \leq 1\}$ é o conjunto das soluções fracamente eficientes de (P_5) . Note que $a_2 \in P_f$ mas $a_2 \notin P$.

Definição 1.50. O conjunto de todos os valores das funções objetivo nas soluções de Pareto é chamada *fronteira de Pareto do Problema (VOP)*.

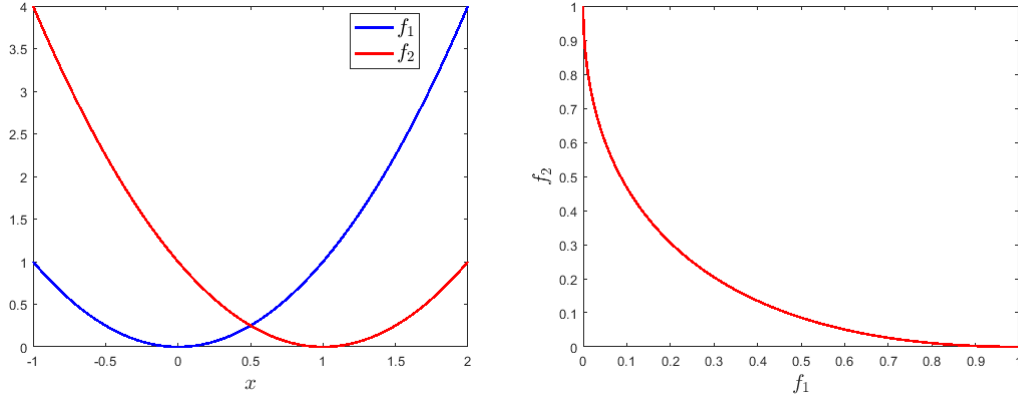
O próximo exemplo ilustra a Definição 1.50.

Exemplo 1.51. Considere o seguinte problema:

$$\begin{aligned} & \text{Minimizar } F(x) = (f_1(x), f_2(x)) \\ & \text{sujeito a } x \in \mathbb{R} \end{aligned} \quad (\text{P}_6)$$

onde $f_1(x) = x^2$ e $f_2(x) = (x - 1)^2$.

A Figura 1.16 ilustra o Exemplo 1.51. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , sendo fácil identificar que as soluções Pareto-ótimas estão no intervalo $[0, 1]$, já que fora deste intervalo ambas as f_1 e f_2 crescem. Note que neste intervalo acontece uma compensação entre as funções objetivos, isto é, enquanto uma função cresce a outra diminui. À direita é representada a fronteira de Pareto.



Fonte: O autor

Figura 1.16: Exemplo 1.51

1.2.1 Existência de Soluções

Nesta seção, discutiremos a existência de soluções para problemas de otimização multiobjetivo da forma (VOP). Para isso, veremos uma relação de comparação de vetores mais geral que a $\leq$ e $<$. Em problemas de otimização multiobjetivo, as atitudes de preferência do decisor desempenham um papel essencial que especifica o significado de otimização ou conveniência. Eles são frequentemente representados como relações binárias no espaço objetivo e são chamados de ordens de preferência.

Uma relação binária $\succ$ em $Y \subseteq \mathbb{R}^r$ é uma relação de dominação se é uma relação transitiva e não reflexiva, ou seja, dados $y^1, y^2, y^3 \in Y$, temos que $y^1 \succ y^2$ e $y^2 \succ y^3$ então $y^1 \succ y^3$ e para cada $y^1 \in Y$, $y^1 \not\succ y^1$. Uma relação de dominação é usada para expressar a preferência do decisor.

Neste contexto, utilizaremos a seguinte notação, dados $y^1, y^2 \in Y$, então:

- $y^1 \succ y^2$ ou $y^2 \prec y^1$ se y^1 é preferível a y^2 ;

- $y^1 \sim y^2$ se y^1 e y^2 forem indiferentes;
- $y^1 \succeq y^2$ se y^1 é preferível a y^2 ou se y^1 e y^2 são indiferentes.

Um elemento $\hat{y} \in Y$ é um **eficiente** de Y se não existe $y \in Y$ tal que $y \succ \hat{y}$. Denotamos

$$\varepsilon(Y, \succ) := \{\hat{y} \in Y \mid \nexists y \in Y \text{ tal que } y \succ \hat{y}\}$$

ao conjunto de elementos eficientes de Y .

Com esta nomenclatura, teríamos que $Y = F(X)$ e a relação de dominação em Y será definida por:

$$y^1 = F(x^1) \succ y^2 = F(x^2) \iff y^1 \leq y^2.$$

Ainda, se $\bar{y} = F(\bar{x})$ é um elemento eficiente de $Y = F(X)$ então $\bar{x}$ é solução eficiente do problema (VOP), e neste caso,

$$\varepsilon(Y, \succ) = \{y \in \mathbb{R}^r \mid y = F(\bar{x}), \bar{x} \text{ é solução eficiente de (VOP)}\}.$$

Assim, dada uma relação de dominação, podemos associar a ela uma multiaplicação, denominada estrutura de dominação. Sejam $Y \subseteq \mathbb{R}^r$, $y \in Y$ e $\succ$ uma relação de dominação em Y . A multiaplicação

$$\begin{aligned} D &: Y \rightrightarrows \mathbb{R}^r \\ y &\mapsto D(y) = \{d \in \mathbb{R}^r \mid y \prec y + d\} \cup \{0\} \end{aligned}$$

Note que dados $y^1, y^2 \in Y$ e $y^1 \neq y^2$, tem-se que $y^1 \succ y^2$ se, e somente se $y^2 \in y^1 + D(y^1)$. Deste modo, podemos redefinir eficiência em termos da estrutura de dominação. Assim,

$$\varepsilon(Y, D) := \{\hat{y} \in Y \mid \nexists y \in Y - \{\hat{y}\} \text{ tal que } \hat{y} \in y + D(y)\}.$$

Observe que $D = \mathbb{R}_+^r := \{y \in \mathbb{R}^r \mid y_r \geq 0\}$, este caso corresponde à chamada ordem de preferência Pareto e neste caso, cada $y \in \varepsilon(Y, D)$ é dito um elemento Pareto eficiente de Y . E ainda, se $D = \text{int}(\mathbb{R}_+^r) \cup \{0\}$ este caso corresponde à chamada ordem de preferência Pareto Fraca e neste caso, cada elemento $y \in \varepsilon(Y, D)$ é dito um elemento fracamente Pareto-eficiente de Y .

Um ponto $\hat{x} \in X$ é dito solução eficiente do problema (VOP) se $f(\hat{x}) \in \varepsilon(Y, D)$ ou, equivalentemente, se não existir $x \in X$ tal que $f(\hat{x}) \in f(x) + D(f(x)) \setminus \{0\}$.

Dizemos que a estrutura de dominação D é dita ser acíclica se ela não possui ciclos, isto é, para cada k , não existem $y^1, y^2, \dots, y^k \in Y$ tais que $y^1 \succ y^2 \succ \dots \succ y^k \succ y^1$.

Teorema 1.52 (Propo. 6.6, [24]). *Se a estrutura de dominação D é acíclica e se, para cada $y \in Y$ os conjuntos $D(y) \setminus \{0\}$ são abertos e se o conjunto Y é compacto e não vazio, então $\varepsilon(Y, D) \neq \emptyset$.*

Teorema 1.53 (Cor. 6.10, [24]). *Suponha que no problema (VOP) temos que a função objetivo F seja contínua e que o conjunto viável X seja compacto. Então o problema (VOP) possui soluções eficientes.*

1.2.2 Escalarização

Escalarização é um processo através do qual é possível transformar um problema vetorial (multiobjetivo) em um problema de otimização escalar (mono objetivo), de tal maneira que as soluções do problema multiobjetivo possam ser obtidas como soluções de um problema clássico de programação não-linear.

Existe uma grande variedade de técnicas para escalarizar um problema de otimização vetorial. Alguns exemplos podem ser vistos em [29]. Existem algumas técnicas que geram apenas uma solução ótima de Pareto e outras que tentam gerar todo o conjunto de soluções ótimas (fracas) de Pareto. Veremos duas técnicas diferentes de escalarização: a primeira técnica que veremos é a técnica da soma ponderada, esta técnica foi introduzido por Gass e Saaty [18] em 1955 e é provavelmente a técnica de escalarização mais utilizado, a segunda é a técnica de Chebyshev ponderado, introduzido por Bowman [4] em 1976.

Técnica da Soma Ponderada

Seja

$$W = \{w \in \mathbb{R}^r \mid w_\ell \geq 0, 1 \leq \ell \leq r \text{ e } \sum_{\ell=1}^r w_\ell = 1\}. \quad (1.13)$$

Para cada $w \in W$, consideremos o seguinte problema (ponderado)

$$\begin{aligned} &\text{Minimizar } \sum_{\ell=1}^r w_\ell f_\ell(x) = \Phi_w(x) \\ &\text{sujeito a } x \in X. \end{aligned} \quad (\text{WS})$$

Alguns resultados teóricos sobre a técnica da soma ponderada são apresentados a seguir. Tais resultados podem ser encontrados em [40].

Teorema 1.54 (Teo. 3.1.1, [40]). *Se existir $w \in W$ tal que $\bar{x} \in X$ é solução de (WS), então $\bar{x}$ é solução Pareto fraca de (VOP).*

Teorema 1.55 (Teo. 3.1.4, [40]). *Suponha que todas as funções de (VOP) sejam convexas. Se $\bar{x}$ é uma solução Pareto de (VOP) então existe $w \in W$ tal que $\bar{x}$ é solução de (WS).*

Teorema 1.56 (Teo. 3.1.6, [40]). *Se existe $w \in W$ tal que $w > 0$ e $\bar{x} \in X$ é solução de (WS), então $\bar{x}$ é solução Pareto de (VOP).*

Técnica de Chebyshev Ponderado

Antes de apresentar a técnica, precisamos da seguinte definição.

Definição 1.57. *Define-se o vetor objetivo ideal $f^* \in \mathbb{R}^r$, onde as componentes f_ℓ^* são obtidas minimizando cada uma das funções objetivo sujeitas individualmente às restrições, isto é, resolvendo*

$$f_\ell^* = \min_{x \in X} f_\ell(x) \quad (1.14)$$

para $\ell = 1, \dots, r$.

É óbvio que, se o mesmo $\bar{x}$ é a solução ótima para todos os problemas dados por (1.14), então $\bar{x}$ seria a solução do problema vetorial (VOP) e o conjunto de soluções de Pareto seria reduzido a ele. Isso não é possível em geral, pois pode haver algum conflito entre os objetivos. Mas as componentes do vetor objetivo ideal podem ser usadas como limitantes inferiores do conjunto de ótimos de Pareto para cada função objetivo.

Considere $w \in W$ como em (1.13). O problema de Chebyshev ponderado correspondente a (VOP) é dado por

$$\begin{aligned} &\text{Minimizar } \|f(x) - f^*\|_\infty^w = \Psi_w(x) \\ &\text{sujeito a } x \in X, \end{aligned} \quad (\text{WCS}_w)$$

onde $\|f(x) - f^*\|_\infty^w = \max_\ell \{w_\ell(f_\ell(x) - f_\ell^*)\}$.

Vejamos agora algumas relações entre o problema escalarizado Chebyshev ponderado (WCS_w) e o problema de otimização vetorial (VOP).

Teorema 1.58 (Teo. 3.4.3, [40]). *Para algum $w \in W$ o problema de Chebyshev ponderado (WCS_w) tem pelo menos uma solução ótima de Pareto para o problema (VOP).*

Teorema 1.59 (Cor. 3.4.4, [40]). *Se o problema de Chebyshev ponderado (WCS_w) tiver uma única solução para $w \in W$ fixado, essa deve ser uma solução ótima de Pareto para o problema (VOP).*

Como vimos na Seção anterior a convexidade do problema de otimização vetorial (VOP) é necessária para garantir que todas as soluções ótimas de Pareto possam ser encontradas pela técnica da soma ponderada. Por outro lado, o teorema a seguir mostra que todas as soluções ótimas de Pareto podem ser encontradas pela técnica de Chebyshev ponderado.

Teorema 1.60 (Teo. 3.4.5, [40]). *Seja $\bar{x} \in X$ uma solução ótima de Pareto para o problema (VOP). Então existe um vetor de ponderação $0 < w \in W$, tal que $\bar{x}$ é uma solução do problema (WCS_w).*

Como podemos observar a técnica de Chebyshev ponderado funciona para problemas convexos e não convexos, diferentemente da técnica da soma ponderada.

Teorema 1.61 (Teo. 3.6, [17]). *Se $w_\ell > 0$, para todo $\ell = 1, \dots, r$, então $\bar{x} \in X$ solução do problema Chebyshev ponderado (WCS_w) é solução ótima de Pareto para o problema (VOP).*

O teorema anterior mostra uma imensa vantagem desta técnica, pois podemos determinar todas as soluções Pareto eficientes para o problema (VOP), sejam elas na parte convexa e não convexa da fronteira. Entretanto, algumas soluções fracamente eficientes podem ser determinadas por esta técnica.

1.2.3 Condições Necessárias e Suficientes de Karush-Kuhn-Tucker

Na Subseção 1.1.2 estabelecemos condições de otimalidade para o problema de otimização escalar cujas funções são diferenciáveis. Nesta subseção temos como foco introduzir as noções de ponto vetorial de Fritz John e de Karush-Kuhn-Tucker para problemas de otimização vetorial (VOP). Inicialmente, condições necessárias de otimalidade (no sentido da eficiência fraca) para o problema (VOP). Como no caso escalar, veremos que sob hipóteses adequadas de convexidade generalizada, tais condições são, também suficientes.

O primeiro resultado é uma extensão do caso escalar, que se dá de modo natural, o qual indica as condições necessárias de Fritz John para otimalidade local de Pareto fraco.

Teorema 1.62 (Fritz John). *Uma condição necessária para $\bar{x}$ ser um ponto Pareto fraco é que existam vetores $\theta \in \mathbb{R}_+^r$, $\lambda \in \mathbb{R}_+^m$ e $\mu \in \mathbb{R}^p$ tais que*

$$\sum_{\ell \in L} \theta_\ell \nabla f_\ell(\bar{x}) + \sum_{i \in I} \lambda_i \nabla g_i(\bar{x}) + \sum_{j \in J} \mu_j \nabla h_j(\bar{x}) = 0, \quad (1.15)$$

$$\lambda_i g_i(\bar{x}) = 0, \quad i \in I, \quad (1.16)$$

$$(\theta, \lambda, \mu) \neq (0, 0, 0). \quad (1.17)$$

Os vetores (θ, λ, μ) que satisfazem o sistema (1.15) – (1.17) são conhecidos como multiplicadores de Fritz John. Além disso, as condições de Fritz John não garantem que o vetor $\theta \neq 0$. Quando temos a garantia de que os multiplicadores θ_ℓ , $\ell \in L$, são não todos nulos, chamamos o sistema (1.15)-(1.16) de condições de Karush-Kuhn-Tucker (KKT), com os respectivos multiplicadores de KKT.

Agora se adicionarmos uma condição de qualificação, podemos garantir que $\theta \neq 0$.

Teorema 1.63 (Teo. 3.3, [53]). *Seja $\bar{x}$ uma solução Pareto fraca para o problema (VOP) e a condição de qualificação de independência linear (LICQ) satisfeita em $\bar{x}$. Então existe (θ, λ, μ) tal que $\bar{x}$ satisfaz (1.15)-(1.16) com $\theta \neq 0$.*

Note que o teorema continua válido para condições de qualificação mais gerais, ver [36].

Vimos na Seção 1.1.2 que sob hipóteses de invexidade, todo ponto KKT é minimizador global. Não é difícil mostrar que o mesmo se aplica a problemas multiobjetivos. Isto é, se as funções forem invexas então todo ponto KKT é solução (global) fracamente eficiente. Novamente, é pertinente estudar

a classe de problemas multiobjetivos para os quais as condições KKT são necessárias e também suficientes para a otimalidade.

No trabalho de Osuna-Gómez et al. [44] é realizado um estudo das caracterizações das soluções para problemas de otimização multiobjetivo irrestrito e restrito. Nos deteremos ao estudo de problemas multiobjetivos com restrições. Por simplicidade, consideraremos o caso em que (VOP) possui apenas restrições de desigualdade.

Definamos agora uma condição de convexidade generalizada mais fraca para o problema de programação multiobjetivo. Esta condição é necessária e suficiente para que possamos garantir a igualdade entre o conjunto de pontos vetoriais Karush-Kuhn-Tucker e o conjunto de pontos fracamente eficientes.

Definição 1.64 (KT-pseudoinvexidade, [44]). *Considere o problema (VOP) apenas com restrições de desigualdade. O problema (VOP) é dito ser um problema vetorial KT-pseudoinvexo com respeito a η se para qualquer $x, \bar{x} \in X$ com $g(x) \leq 0$ e $g(\bar{x}) \leq 0$ existe $\eta(x, \bar{x}) \in \mathbb{R}^n$ tal que:*

$$\begin{aligned} f(x) < f(\bar{x}) &\Rightarrow \nabla f(\bar{x})\eta(x, \bar{x}) < 0, \\ -\nabla g_i(\bar{x})\eta(x, \bar{x}) &\geq 0, \quad i \in I(\bar{x}). \end{aligned}$$

Teorema 1.65 (Teo. 3.7, [44]). *Todo ponto viável $\bar{x}$ que satisfaz (1.15)-(1.17) é um ponto fracamente eficiente para (VOP) se, e somente se, o problema (VOP) é KT-pseudoinvexo.*

Como vimos na Seção 1.1.2, as condições de KKT são suficientes para a otimalidade quando o problema em questão é KT-invexo. Além disso, vimos o conceito de KT-pseudoinvexidade, o qual garante que se o problema (VOP) é KT-pseudoinvexo então todo ponto viável que satisfaz KKT é um ponto fracamente eficiente para (VOP) .

1.2.4 Condições de Otimalidade de Segunda Ordem

De modo similar ao caso escalar, existem pontos KKT que não são soluções globais fracamente eficientes de (VOP) e veremos que é possível estabelecer condições necessárias e suficientes de segunda ordem para a otimalidade global. Neste sentido, destacamos os trabalhos de Bigi e Castellani [7], Maciel, Santos e Sottosanto [36] e Wang [53]. Nesses trabalhos, foram obtidas condições de otimalidade necessárias e suficientes, usando informações sobre as derivadas segundas das funções do problema. Para isso devemos assumir que as funções envolvidas em (VOP) são de classe $\mathcal{C}^2$.

Uma primeira condição de qualificação para (VOP) , com base em curvas viáveis, é dada a seguir.

Definição 1.66 ((WSOCQ) Wang [53]). *Dizemos que a condição de qualificação fraca de segunda ordem*

(WSOCQ) é válida em um ponto viável $\bar{x} \in X$, desde que, para qualquer d , não nulo, que satisfaz

$$\begin{cases} \nabla^T h_j(\bar{x})d = 0, & \forall j \in J, \\ \nabla^T g_i(\bar{x})d = 0, & i \in I(\bar{x}), \end{cases}$$

se existe um vetor $w \in \mathbb{R}^n$ tal que

$$\begin{cases} d^T \nabla^2 h_j(\bar{x})d + \nabla^T h_j(\bar{x})w = 0, & \forall j \in J, \\ d^T \nabla^2 g_i(\bar{x})d + \nabla^T g_i(\bar{x})w \leq 0, & i \in I(\bar{x}), \end{cases}$$

então existe uma curva duas vezes diferenciável $\alpha : [0, t_0] \rightarrow \mathbb{R}^n$ satisfazendo

$$\alpha(0) = \bar{x}, \quad \alpha'(0) = d, \quad \alpha''(0) = w \quad e \quad \alpha(t) \in X, \quad \forall t \in [0, t_0].$$

O próximo teorema nos fornece uma condição necessária de segunda ordem para (VOP).

Teorema 1.67 (Teo. 3.2, Wang [53]). *Suponha que $\bar{x} \in X$ e (WSOCQ) é válida em $\bar{x}$, então uma condição necessária para $\bar{x}$ ser uma solução local Pareto fraca do problema (VOP) e para qualquer d , não nulo, que satisfaz o seguinte sistema*

$$\begin{cases} \nabla^T f_\ell(\bar{x})d \leq 0, & \ell \in L, \\ \nabla^T h_j(\bar{x})d = 0, & j \in J, \\ \nabla^T g_i(\bar{x})d = 0, & i \in I(\bar{x}), \end{cases}$$

existem $\theta \in \mathbb{R}_+^r$, $\theta \neq 0$, $\mu \in \mathbb{R}_+^m$ e $\lambda \in \mathbb{R}^p$ tais que Fritz John são válidas e

$$d^T \left[\sum_{\ell \in L} \theta_\ell \nabla^2 f_\ell(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d \geq 0.$$

A seguinte definição pode ser encontrada em [37].

Definição 1.68. *A condição de qualificação de primeira ordem de Kuhn-Tucker (KTCQ) é válida em x se e somente se, para qualquer vetor não nulo $d \in \mathcal{L}(x)$, existe $\tau > 0$ e uma curva $a : [0, \tau[\rightarrow X$ continuamente diferenciável em 0 tal que $a(0) = x$, $a'(0) = d$, e $g_i(a(t)) \leq 0$, para $i \in I$ e $h_j(a(t)) = 0$, para $j \in J$ e todo $t \in [0, \tau[$.*

A condição de qualificação de segunda ordem proposta por McCormick [39] para o caso escalar é baseado no seguinte conjunto

$$\bar{\mathcal{L}}(x) := \{d \in \mathbb{R}^n \mid \nabla^T g_i(x)d = 0, i \in I(\bar{x}), e \nabla^T h_j(x)d = 0, j \in J\}.$$

Para adaptar a condição de qualificação de segunda ordem de McCormick e dar uma condição similar para o caso vetorial, é considerado o seguinte conjunto

$$\hat{\mathcal{L}}(x) := \{d \in \mathbb{R}^n \mid \nabla^T g_i(x)d \leq 0, i \in I_0(\bar{x}), \nabla^T g_i(x)d = 0, i \in I_+(\bar{x}), e \nabla^T h_j(x)d = 0, j \in J\},$$

onde $I_0(\bar{x}) = \{i \in I(\bar{x}) \mid \mu_i = 0\}$ e $I_+(\bar{x}) = \{i \in I(\bar{x}) \mid \mu_i > 0\}$.

Embora aceitável para o caso escalar, a substituição de $\bar{\mathcal{L}}(\bar{x})$ por $\hat{\mathcal{L}}(\bar{x})$ produziria uma condição de qualificação de segunda ordem mais forte, mas sabemos que, quanto mais fraca a condição de qualificação, mais forte é a condição de otimalidade.

Definição 1.69. *A condição de qualificação de segunda ordem (SOCQ) é válida em x se e somente se para qualquer vetor não nulo $d \in \hat{\mathcal{L}}(x)$, existe $\tau > 0$ e uma curva duas vezes diferenciável $a : [0, \tau[\rightarrow X$, tal que $a(0) = x$, $a'(0) = d$, e $g_i(a(t)) = 0$ para todo $i \in I(x)$, $h_j(a(t)) = 0$, para todo $j \in J$, $t \in [0, \tau[$.*

O resultado a seguir fornece uma condição necessária, onde os multiplicadores não dependem da direção tomada, mas com um custo extra, pois devemos assumir duas condições de qualificação, diferente do caso da condição necessária dada por Wang [53], que depende da direção tomada.

Teorema 1.70 (Teo. 4.1, Maciel et al. [36]). *Seja $\bar{x} \in X$ e assuma que as condições de qualificação KTCQ e a SOCQ são válidas em $\bar{x}$. Se $\bar{x}$ é uma solução local de Pareto fraca de (VOP), então existem vetores $\theta \in \mathbb{R}_+^L$, $\mu \in \mathbb{R}_+^m$ e $\lambda \in \mathbb{R}^p$ tais que as condições de KKT são válidas e, para qualquer vetor $d \in \mathcal{L}(\bar{x})$ e que satisfaz*

$$\begin{cases} \theta_\ell \nabla^T f_\ell(\bar{x})d = 0, & \forall \ell \in L, \\ \mu_i \nabla^T g_i(\bar{x})d = 0, & \forall i \in I(\bar{x}), \end{cases}$$

segue também que

$$d^T \left[\sum_{\ell \in L} \theta_\ell \nabla^2 f_\ell(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d \geq 0.$$

A seguinte modificação da condição de qualificação de Mangasarian-Fromovitz foi introduzida por Kyparisis [33], para alcançar a unicidade dos multiplicadores, para problemas escalares.

Definição 1.71. *A condição de qualificação estrita de Mangasarian-Fromovitz (SMFCQ) é válida em um ponto estacionário $\bar{x}$ quando existem os multiplicadores de Fritz John tais que para os conjuntos de índices*

$$I_+(\bar{x}) := \{i \in I(\bar{x}) \mid \bar{\mu}_i > 0\}, \quad I_0(\bar{x}) := I(\bar{x}) \setminus I_+(\bar{x})$$

as seguintes condições são válidas:

- $\nabla g_i(\bar{x})$, $i \in I_+(\bar{x})$, $\nabla h_j(\bar{x})$, $j \in J$, são linearmente independentes,
- existe $d \in \mathbb{R}^n$ tal que

$$\begin{aligned} \nabla^T g_i(\bar{x})d &< 0, & i \in I_0(\bar{x}), \\ \nabla^T g_i(\bar{x})d &= 0, & i \in I_+(\bar{x}), \\ \nabla^T h_j(\bar{x})d &= 0, & j \in J. \end{aligned}$$

Com esta modificação de (MFCQ), Kyparisis [33] provou que o conjunto de multiplicadores KKT é um conjunto unitário se e somente se (SMFCQ) é válida.

Note que, no caso vetorial o multiplicador θ não é um número mas um vetor, Assim a definição de multiplicadores KKT deve ser escrito de maneira diferente. Seja

$$M_+(\bar{x}) := \{(\theta, \mu, \lambda) \mid (1.15) - (1.16) \text{ são válidas e } \|\theta\| = 1\}$$

o conjunto de multiplicadores KKT para o problema (VOP).

Para alcançar a unicidade dos multiplicadores sem definir um vetor unitário a priori, precisamos introduzir uma nova condição de regularidade, fortalecendo a (SMFCQ) da seguinte maneira:

Definição 1.72. A condição de regularidade estrita de Mangasarian-Fromovitz (SMFRC)_s é válida em um ponto estacionário $x \in X$ com respeito ao índice $s \in L$ quando existem multiplicadores de Fritz John, de modo que, para os conjuntos de índices $L_+ := \{\ell \in L \mid \bar{\theta}_\ell > 0\}$, $I_+(x) := \{i \in I(x) \mid \bar{\mu}_i > 0\}$, e $I_0(x) = I(x) \setminus I_+(x)$, as seguintes condições são válidas:

- $\nabla f_\ell(x), \ell \in L_+$ e $\ell \neq s$, $\nabla g_i(x), i \in I_+(x)$, e $\nabla h_j(x), j \in J$ são linearmente independentes;

- existe $d \in \mathbb{R}^n$ tal que

$$\begin{cases} \nabla^T f_\ell(x)d < 0, & \ell \notin L_+, \ell \neq s, \\ \nabla^T g_i(x)d < 0, & i \in I_0(x), \\ \nabla^T f_\ell(x)d = 0, & \ell \in L_+, \ell \neq s, \\ \nabla^T g_i(x)d = 0, & i \in I_+(x), \\ \nabla^T h_j(x)d = 0, & j \in J. \end{cases} \quad (1.18)$$

A condição acima é um requisito bastante forte. De fato, a solvabilidade do sistema (1.18) força algumas componentes fixas de todos os multiplicadores KKT a serem zero, como mostra o lema a seguir.

Lema 1.73 (Lema 1, [7]). Seja $(\bar{\theta}, \bar{\mu}, \bar{\lambda}) \in M_+(\bar{x})$ e suponha que (1.18) admita uma solução para algum $s \in L_+$. Se $(\theta', \mu', \lambda') \in M_+(\bar{x})$ então $\{\ell \in L \mid \theta'_\ell > 0\} \subseteq L_+$ e $\{i \in I \mid \mu'_i > 0\} \subseteq I_+(\bar{x})$.

No caso escalar, $L_+ = L$ consiste apenas no índice correspondente à função objetivo e, consequentemente, (SMFRC) reduz a (SMFCQ).

Observação 1.74. Quando (SMFRC)_s é válida para todo $s \in L_+$, podemos simplesmente dizer que a condição de regularidade estrita de Mangasarian-Fromovitz (SMFRC) é válida, sem qualquer possível mal-entendido.

Definição 1.75 (Direção Crítica). Dizemos que $d \in \mathbb{R}^n$ é uma direção crítica para o problema (VOP) no ponto $\bar{x} \in X$ se

$$\begin{aligned}\nabla^T f_\ell(\bar{x})d &\leq 0, & \ell \in L, \\ \nabla^T g_j(\bar{x})d &\leq 0, & j \in I(\bar{x}), \\ \nabla^T h_i(\bar{x})d &= 0, & i \in I.\end{aligned}$$

Denote por $D_m(\bar{x})$ o conjunto de direções críticas em $\bar{x}$.

Teorema 1.76 (Teo. 4, Bigi e Castellani [7]). Seja $\bar{x} \in X$ um ponto local de Pareto fraco para o problema (VOP) e (θ, μ, λ) tal que (1.15)-(1.16) são válidas. Se (SMFRC) é válida em $\bar{x}$, então existe $\gamma > 0$ tal que $\gamma(\theta, \mu, \lambda) \in M_+(\bar{x})$ e

$$\begin{cases} \theta_\ell \nabla^T f_\ell(\bar{x})d = 0, & \forall \ell \in L, \\ \mu_j \nabla^T g_j(\bar{x})d = 0, & \forall j \in I(\bar{x}), \end{cases} \quad (1.19)$$

$$d^T \left[\sum_{\ell \in L} \theta_\ell \nabla^2 f_\ell(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d \geq 0 \quad (1.20)$$

são válidas para toda direção $d \in D_m(\bar{x})$.

Agora, vejamos condições suficientes de segunda ordem para a eficiência fraca.

Teorema 1.77 (Teo. 3.5, Wang [53]). Se para cada $d \in D_m(\bar{x}) \setminus \{0\}$, existem $\theta \in \mathbb{R}_+^r$, $\mu \in \mathbb{R}_+^m$ e $\lambda \in \mathbb{R}^p$ tais que as condições de KKT são válidas e

$$d^T \left[\sum_{\ell \in L} \theta_\ell \nabla^2 f_\ell(\bar{x}) + \sum_{i \in I} \mu_i \nabla^2 g_i(\bar{x}) + \sum_{j \in J} \lambda_j \nabla^2 h_j(\bar{x}) \right] d > 0, \quad (1.21)$$

então $\bar{x}$ é uma solução local de Pareto do problema (VOP).

Note que este resultado é semelhante ao dado na seção anterior para o caso mono objetivo.

Neste capítulo tratamos de conceitos básicos de soluções, existência de soluções, condições necessárias e suficientes de Karush-Kuhn-Tucker (KKT) e condições de segunda ordem para os problemas de programação não linear, tanto para problemas mono objetivo quanto para problemas multiobjetivos. No próximo capítulo discutiremos condições de otimalidade utilizando a teoria de medida e integração. O estudo será apresentado para problemas mono objetivo e em seguida os resultados serão estendidos para problemas multiobjetivos, como uma contribuição deste trabalho.

Capítulo 2

CARACTERIZAÇÃO DE OTIMALIDADE VIA INTEGRAÇÃO

Existem inúmeros estudos sobre como caracterizar otimalidade, no entanto, quase todas as tentativas são de natureza local e mesmo assim tendem a exigir vários graus de diferenciabilidade, a menos que sejam impostas algumas hipóteses de convexidade ou convexidade generalizada, como vimos no Capítulo 1.

A atenção dada à busca de condições necessárias e suficientes de otimalidade sem a exigência da diferenciabilidade ou, ainda, sem a exigência de convexidade - no caso das condições suficientes - é de fundamental importância.

O cálculo integral, sendo útil em conjuntos compactos e não vazios, possui um caráter mais global que o cálculo diferencial, o qual possui caráter mais local. Portanto, as condições de otimalidade baseadas em caracterizações integrais podem ser utilizadas em uma variedade maior de problemas. A abordagem que veremos neste capítulo difere das tradicionais que vimos no Capítulo 1 as quais são baseadas em derivadas, esta apela para o uso da teoria da medida e integração.

Primeiramente veremos tais caracterizações para problemas mono objetivo e na seção subsequente iremos estendê-las para problemas de otimização multiobjetivo, representando a principal contribuição desta dissertação.

2.1 Problema Mono Objetivo

Nesta seção vamos considerar o problema

$$\begin{array}{ll} \text{Minimizar } f(x) & \\ x \in X & \end{array} \quad (\text{P})$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e X é um subconjunto não vazio de $\mathbb{R}^n$. Note que o conjunto viável não é necessariamente formado por restrições de igualdade e desigualdade como no Capítulo 1.

Serão considerados os casos onde f é contínua e também o caso em que f pode ser descontínua.

2.1.1 Caso em que a função objetivo é contínua

Os primeiros resultados desta subseção são apresentados por Falk em [16], para problemas de maximização. Logo, para fins de uniformização de escrita, para o problema dado em (P), precisamos reescrever os resultados para problemas de minimização. Antes de começar a discussão sobre caracterização de otimalidade via integral, precisaremos dos seguintes lemas.

Lema 2.1. *Seja $\xi : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função contínua, com $\xi(x) \leq 0$ para todo $x \in X$. Suponha que X seja um conjunto compacto e $\text{int}(X) \neq \emptyset$. Então, para todo $\alpha \in [0, 1]$, todo $x \in X$ e todos $r, s \in [0, \infty)$, temos*

$$[-\xi(x)]^{\alpha r + (1-\alpha)s} \leq \alpha [-\xi(x)]^r + (1-\alpha)[- \xi(x)]^s. \quad (2.1)$$

Demonstração. Para cada $x \in X$ dado, considere a função $\Xi(r) = [-\xi(x)]^r$. Para mostrar a desigualdade (2.1), basta mostrar que Ξ é uma função convexa.

Se $\xi(x) = 0$, o resultado é imediato. Agora, suponha que $\xi(x) \neq 0$. Aplicando o logaritmo em $\Xi(r) = [-\xi(x)]^r$ obtemos:

$$\ln(\Xi(r)) = \ln([-\xi(x)]^r) = r \ln(-\xi(x)). \quad (2.2)$$

Derivando a igualdade (2.2) em relação à variável r temos:

$$\begin{aligned} \frac{d}{dr}[\ln(\Xi(r))] &= \frac{d}{dr} r \ln(-\xi(x)) \\ \frac{1}{\Xi(r)} \Xi'(r) &= \ln(-\xi(x)) \end{aligned}$$

ou seja,

$$\Xi'(r) = \ln(-\xi(x)) \Xi(r). \quad (2.3)$$

Agora, derivando a igualdade (2.3) novamente em relação à r teremos:

$$\Xi''(r) = \Xi(r) [\ln(-\xi(x))]^2.$$

Substituindo a expressão de Ξ e usando o fato que $\xi(x) \leq 0$, temos

$$\Xi''(r) = [-\xi(x)]^r [\ln(-\xi(x))]^2 \geq 0.$$

Logo a função Ξ é convexa. ■

Obtemos quase imediatamente do Lema 2.1 o seguinte resultado.

Lema 2.2 (Lema 1, [16]). *Sob as hipóteses do lema anterior, temos que a função*

$$\Upsilon(t) = \int_X [-\xi(x)]^t dx \quad (2.4)$$

é convexa sobre $[0, \infty)$.

Demonstração. Segue do Lema 2.1 que

$$[-\xi(x)]^{\alpha r + (1-\alpha)s} \leq \alpha[-\xi(x)]^r + (1-\alpha)[- \xi(x)]^s,$$

para todo $\alpha \in [0, 1]$, todo $x \in X$ e todo $r, s \in [0, \infty)$. Então

$$\begin{aligned} \Upsilon(\alpha r + (1-\alpha)s) &= \int_X [-\xi(x)]^{\alpha r + (1-\alpha)s} dx \\ &\leq \alpha \int_X [-\xi(x)]^r dx + (1-\alpha) \int_X [-\xi(x)]^s dx \\ &= \alpha \Upsilon(r) + (1-\alpha) \Upsilon(s). \end{aligned}$$

■

A seguir temos o nosso primeiro teorema de representação integral de soluções ótimas globais.

Teorema 2.3 (Teo. 1, [16]). *Sejam $X \subset \mathbb{R}^n$ um compacto com interior não vazio e $f : X \rightarrow (-\infty, 0]$ uma função contínua. Seja ainda, $\bar{x} \in X$ tal que $f(\bar{x}) = -1$. Então $\Upsilon(t) = \int_X [-f(x)]^t dx \rightarrow \infty$ quando $t \rightarrow \infty$ se, e somente se, $\bar{x}$ não é solução global de (P) .*

Demonstração. Primeiro, suponhamos que $\bar{x} \in X$ não é solução global de (P) . Então existe $x_0 \in X$ tal que $f(x_0) < -1$.

Como f é contínua, segue que existe uma vizinhança V de x_0 tal que para $\tau > 0$ suficientemente pequeno $f(x) < -1 - \tau$, para todo $x \in X \cap V$. Logo,

$$\Upsilon(t) = \int_X [-f(x)]^t dx = \int_{X \setminus V} [-f(x)]^t dx + \int_{X \cap V} [-f(x)]^t dx. \quad (2.5)$$

Porém $-f$ é não negativa, donde obtemos,

$$\Upsilon(t) \geq \int_{X \cap V} [-f(x)]^t dx \geq (1 + \tau)^t \mathcal{V}(X \cap V). \quad (2.6)$$

Como X é um conjunto compacto, o conjunto $X \cap V$ tem volume positivo ($\mathcal{V}(X \cap V) > 0$). Além disso, temos que $\tau > 0$, donde segue da desigualdade (2.6) que $\Upsilon(t) \rightarrow \infty$ quando $t \rightarrow \infty$.

Para mostrar a recíproca, suponha, por contradição, que $\bar{x}$ é solução global de (P) . Neste caso, para todo $x \in X$, temos $f(x) \geq f(\bar{x}) = -1$ e, portanto,

$$\Upsilon(t) = \int_X [-f(x)]^t dx \leq \int_X dx = \mathcal{V}(X), \quad (2.7)$$

e como X é compacto, segue que $\mathcal{V}(X)$ é finito. Assim, a função $\Upsilon(t)$ é limitada, donde segue que $\Upsilon(t) \nrightarrow \infty$. ■

Observação 2.4. Na demonstração do Teorema 2.3, observamos que:

- Se $f(\bar{x}) = 0$, haveria nada a demonstrar;

- A hipótese $f \leq 0$ não é muito restritiva, pois a função $x \mapsto -\exp(-f(x))$ é não positiva e seus minimizadores coincidem com os minimizadores de f ;
- Se tivéssemos $f(\bar{x}) < -1$, bastaria aplicar o Teorema 2.3 à função $\xi(x) = -\frac{f(x)}{f(\bar{x})}$.

É consequência quase imediata do Teorema 2.3 o seguinte resultado:

Corolário 2.5. *Sob as hipóteses do Teorema 2.3, considere a sequência $\{\Upsilon_n\}_{n \in \mathbb{N}}$, $\Upsilon_n = \int_X [-f(x)]^n dx$. Então:*

- (i) $\bar{x} \in X$ é uma solução ótima global de (P) se, e somente se $\{\Upsilon_n\}_{n \in \mathbb{N}}$ converge.
- (ii) Se para algum n , temos $\Upsilon_{n+1} > \Upsilon_n$, então $\bar{x}$ não é solução ótima global do problema (P) .

Demonstração. (i) Imediata do Teorema 2.3.

- (ii) Suponha, por absurdo, que $\bar{x}$ é solução ótima de (P) . Logo, $f(x) \geq f(\bar{x}) = -1$ para todo $x \in X$. Além disso temos que $-1 \leq f(x) \leq 0$, ou ainda, $0 \leq -f(x) \leq 1$. Portanto, temos para todo $n \in \mathbb{N}$ que

$$\int_X ([-f(x)]^{n+1} - [-f(x)]^n) dx \leq 0. \quad (2.8)$$

O que é um absurdo, pois por hipótese temos que existe n tal que $\Upsilon_{n+1} > \Upsilon_n$. Segue que $\bar{x}$ não é solução global para o problema (P) . ■

A convexidade da função Υ definida em (2.4) nos fornece um método de verificação que pode ser útil na eliminação de candidatos $\bar{x}$ que não são minimizadores globais. Pois, se $\Upsilon(t) \rightarrow \infty$ quando $t \rightarrow \infty$, a função Υ deve ser monótona crescente antes de algum ponto t_0 e monótona decrescente no intervalo $[1, t_0]$. Portanto, se é possível obter dois pontos $1 \leq t_1 < t_2$ tais que $\Upsilon(t_2) > \Upsilon(t_1)$, segue-se imediatamente que $\bar{x}$ não é um minimizador global para o problema (P) , como vimos no resultado do Corolário 2.5.

Exemplo 2.6. *Considere a função $f : X \rightarrow \mathbb{R}$ dada por $f(x) = \frac{x^3}{3} - \frac{3x^2}{2} + 2x$ onde $X = [0, 3]$.*

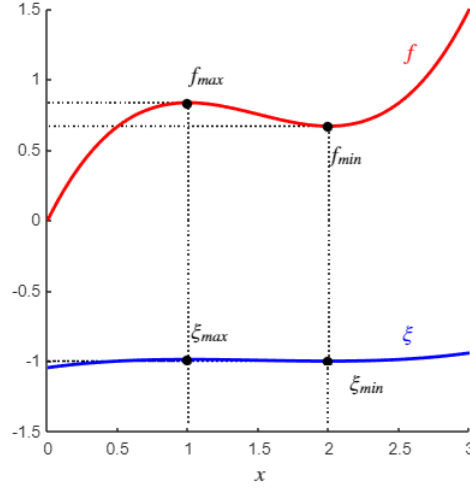
Vejamos primeiramente os pontos críticos de f , ou seja, os zeros de $f'(x) = x^2 - 3x + 2$. Onde, temos $\bar{x} = 1$ e $\hat{x} = 2$.

Note que, $f''(\hat{x}) > 0$. Assim, $\hat{x}$ é ao menos minimizador local. Vejamos se $\hat{x}$ é um minimizador global.

Uma cota superior é obtida da seguinte observação, $f(x) \leq \frac{x^3}{3} + 2x$ para todo $x \in [0, 3]$, consequentemente $f(x) \leq 15$, para todo $x \in [0, 3]$. Em $\hat{x} = 2$, temos que $f(2) - 15 = \frac{2}{3} - 15 = -\frac{43}{3}$. Assim, consideramos a função

$$\xi(x) = \frac{3}{43} \left(\frac{x^3}{3} - \frac{3x^2}{2} + 2x - 15 \right),$$

no qual é não positiva em $[0, 3]$ e assume valor -1 em $\hat{x} = 2$. Além disso, f e ξ possuem os mesmos pontos extremos locais e globais.



Fonte: O autor

Figura 2.1: Gráfico das funções f e ξ

Assim, a função

$$\Upsilon(n) = \int_{[0,3]} [-\xi(x)]^n dx \quad (2.9)$$

é uma função convexa não monótona. Então

$$\Upsilon(20) = \int_0^3 \left[\frac{3}{43} \left(\frac{x^3}{3} - \frac{3x^2}{2} + 2x - 15 \right) \right]^{20} dx \approx 2.8226,$$

e

$$\Upsilon(21) = \int_0^3 \left[-\frac{3}{43} \left(\frac{x^3}{3} - \frac{3x^2}{2} + 2x - 15 \right) \right]^{21} dx \approx 2.8230.$$

Como $\Upsilon_{21} > \Upsilon_{20}$, temos que $\hat{x}$ não é um minimizador global.

■

Outras caracterizações de otimalidade global por integração podem ser obtidas através dos conjuntos de nível da função f . Para cada $c \in \mathbb{R}$, define-se o conjunto de nível da função f por

$$H_c = \{x \in \mathbb{R}^n \mid f(x) \leq c\}.$$

No que se segue, assumiremos as seguintes hipóteses:

A1. f é contínua.

A2. Existe um $b \in \mathbb{R}$ tal que H_b é não vazio e compacto.

Inicialmente, consideraremos o problema irrestrito

$$\begin{aligned} &\text{Minimizar } f(x) \\ &x \in \mathbb{R}^n \end{aligned} \tag{P'}$$

Sob as hipóteses A1 e A2 o problema (P') tem solução.

Doravante, consideraremos a tripla $(\mathbb{R}^n, \Omega, \mu)$, onde Ω é uma σ -álgebra de subconjuntos do espaço topológico $\mathbb{R}^n$ e μ é a medida de Lebesgue. As integrais aqui consideradas são integrais de Lebesgue.

Proposição 2.7. *Seja $\bar{c} \in \mathbb{R}$. Sob as hipóteses A1 e A2, temos que, se $\mu(H_{\bar{c}}) = 0$ então $\bar{c}$ é o valor ótimo de f em $\mathbb{R}^n$ e $H_{\bar{c}}$ é o conjunto dos minimizadores globais de (P') .*

Demonstração. Suponha, por absurdo, que $\bar{c}$ não é valor mínimo de (P') . Seja $\hat{c}$ o valor mínimo de f . Seja $\zeta > 0$ tal que $\bar{c} - \hat{c} = 2\zeta > 0$. Sendo f contínua, o conjunto $E = \{x \in \mathbb{R}^n \mid f(x) < \bar{c} - \zeta\}$ é aberto. Note que E é não vazio. (Com efeito: $E = \emptyset$ se e somente se $\bar{c} - \zeta \leq \hat{c}$. Mas, por construção: $\bar{c} - \zeta = \zeta + \hat{c} > \hat{c}$. Logo, E é não vazio.) Sendo $\hat{c} < \bar{c}$, tem-se $H_{\hat{c}} \subset H_{\bar{c}}$. Além disso, $H_{\hat{c}} \neq \emptyset$ e $E \subset H_{\bar{c}}$. Logo, $0 < \mu(E) \leq \mu(H_{\bar{c}})$, o que contradiz a hipótese. Portanto, $\bar{c}$ é o valor ótimo de (P') . ■

Consequentemente, temos que se $\bar{c}$ é o valor ótimo de (P') , então, $\mu(H_c) > 0$, para cada $c > \bar{c}$.

Em Estatística, uma medida de tendência central é um valor central ou valor típico para uma distribuição de probabilidade. Existem diferentes tipos de medidas de tendência central, e cada uma delas tem suas vantagens e desvantagens, que só vão depender dos dados e dos fins desejados. Os tipos mais comuns, ou até mesmo, mais utilizados são: a média aritmética (ou, simplesmente, média ou valor médio), a mediana, a moda, a média geométrica e a média quadrática.

Estamos interessados apenas no primeiro tipo de medida de tendência central, a média. No decorrer do texto nos limitaremos a chamar de média ou, simplesmente, valor médio. Vejamos a seguir a definição de valor médio.

Definição 2.8 (Def. 1.1, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$ e considere $c > \bar{c}$. Definimos*

$$M(f, c) = \frac{1}{\mu(H_c)} \int_{H_c} f(x) d\mu \tag{2.10}$$

como sendo o valor médio¹ da função f sobre seu conjunto de nível $H_c = \{x \in \mathbb{R}^n \mid f(x) \leq c\}$.

Temos que $\mu(H_c) > 0$ para $c > \bar{c}$. Consequentemente, o valor médio está bem definido devido à continuidade da função f .

A seguir, vejamos algumas propriedades do valor médio.

¹A média, ou valor médio, representa o centro de dados ou centro de distribuição de dados.

Proposição 2.9 (Propo. 1.0, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Para $c > \bar{c}$, temos que $M(f, c) \leq c$.*

Demonstração. Por definição $f(x) \leq c$ para $x \in H_c$ para o qual

$$\begin{aligned} M(f, c) &= \frac{1}{\mu(H_c)} \int_{H_c} f(x) d\mu \\ &\leq \frac{1}{\mu(H_c)} \int_{H_c} c d\mu \\ &= \frac{1}{\mu(H_c)} c \mu(H_c) = c. \end{aligned}$$

■

A próxima proposição garante que a função definida em (2.10) é não decrescente em relação à segunda variável.

Proposição 2.10 (Prop. 1.1, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Se $c_2 \geq c_1 > \bar{c}$, então $M(f, c_2) \geq M(f, c_1)$.*

Demonstração. Aplicando a Proposição 2.9 e conhecendo as propriedades de integração, temos

$$\begin{aligned} M(f, c_2) &= \frac{1}{\mu(H_{c_2})} \int_{H_{c_2}} f(x) d\mu \\ &= \frac{1}{\mu(H_{c_2})} \left[\int_{H_{c_1}} f(x) d\mu + \int_{H_{c_2} \setminus H_{c_1}} f(x) d\mu \right] \\ &\geq \frac{1}{\mu(H_{c_2})} \int_{H_{c_1}} f(x) d\mu + \frac{\mu(H_{c_2}) - \mu(H_{c_1})}{\mu(H_{c_2})} c_1 \\ &\geq \frac{1}{\mu(H_{c_2})} \int_{H_{c_1}} f(x) d\mu + \frac{\mu(H_{c_2}) - \mu(H_{c_1})}{\mu(H_{c_2})} \frac{1}{\mu(H_{c_1})} \int_{H_{c_1}} f(x) d\mu \\ &= \frac{\mu(H_{c_1}) + \mu(H_{c_2}) - \mu(H_{c_1})}{\mu(H_{c_1})\mu(H_{c_2})} \int_{H_{c_1}} f(x) d\mu \\ &= \frac{1}{\mu(H_{c_1})} \int_{H_{c_1}} f(x) d\mu \\ &= M(f, c_1). \end{aligned}$$

■

O lema a seguir estabelece que se uma sequência $\{c_k\}$ decrescente é convergente então a sequência de conjunto de nível $\{H_{c_k}\}$ também é.

Lema 2.11 (Lema 1.2, [23]). *Considere $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Suponha que $\{c_k\}_{k \in \mathbb{N}}$ seja uma sequência decrescente que converge para $c \geq \bar{c}$ quando $k \rightarrow \infty$. Então,*

$$H_c = \bigcap_{k=1}^{\infty} H_{c_k} = \lim_{k \rightarrow \infty} H_{c_k} \quad (2.11)$$

e

$$\lim_{k \rightarrow \infty} \mu(H_{c_k}) = \mu(H_c) \quad (2.12)$$

Demonstração. Pela definição de conjunto de nível e pelo fato que a sequência $\{c_k\}$ é decrescente convergente a c , temos $H_c \subset H_{c_k} \subset H_{c_{k-1}}$ para todo $k \in \mathbb{N}$. Consequentemente, $\lim_{k \rightarrow \infty} H_{c_k} = \bigcap_{k=1}^{\infty} H_{c_k}$.

Se $x \in \bigcap_{k=1}^{\infty} H_{c_k}$, então $f(x) \leq c_k$, para todo $k \in \mathbb{N}$. Portanto, $f(x) \leq c$, isto é, $x \in H_c$. A conclusão segue imediatamente da continuidade da medida μ . ■

A seguir temos uma proposição a qual nos fornece de certa forma a continuidade do valor médio de uma função sobre o conjunto de nível.

Proposição 2.12 (Propo. 1.3, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Suponha que $\{c_k\}_{k \in \mathbb{N}}$ é uma sequência decrescente que converge para $c > \bar{c}$. Então,*

$$M(f, c) = \lim_{c_k \downarrow c} M(f, c_k). \quad (2.13)$$

Demonstração. De acordo com a Proposição 2.9, a sequência $\{M(f, c_k)\}_{k \in \mathbb{N}}$ é decrescente e

$$M(f, c_k) \geq M(f, c)$$

para todo $k \in \mathbb{N}$, para o qual o limite dado em (2.13) existe. Além disso,

$$\begin{aligned} 0 &\leq \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu - \frac{1}{\mu(H_c)} \int_{H_c} f(x) d\mu \\ &\leq \left| \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu - \frac{1}{\mu(H_c)} \int_{H_{c_k}} f(x) d\mu \right| + \left| \frac{1}{\mu(H_c)} \int_{H_{c_k}} f(x) d\mu - \frac{1}{\mu(H_c)} \int_{H_c} f(x) d\mu \right| \\ &\leq \left| \left[\frac{1}{\mu(H_{c_k})} - \frac{1}{\mu(H_c)} \right] \int_{H_{c_k}} f(x) d\mu \right| + \frac{1}{\mu(H_c)} \left| \int_{H_{c_k} \setminus H_c} f(x) d\mu \right| \end{aligned}$$

Os dois últimos termos tendem a zero quando c_k tende para c por causa da continuidade da medida μ e a continuidade absoluta da integral de uma função mensurável limitada. ■

Vejamos a seguir algumas propriedades do valor médio.

Proposição 2.13 (Propo. 1.4, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. O valor médio de uma função sobre seu conjunto de nível tem as seguintes propriedades:*

1. $M(\lambda, c) = \lambda$ para toda constante $\lambda \geq c$.
2. $M(\lambda f, \lambda c) = \lambda M(f, c)$ para toda constante $\lambda > 0$ e $c > \bar{c}$.
3. $M(f + \lambda, c + \lambda) = M(f, c) + \lambda$ para toda constante $\lambda > 0$ e $c > \bar{c}$.

Demonstração. 1. Imediato da definição da função M dada em (2.10).

2. Como $\{x \in \mathbb{R}^n \mid \lambda f(x) \leq \lambda c\} = \{x \in \mathbb{R}^n \mid f(x) \leq c\} = H_c$ para $\lambda > 0$, temos

$$M(\lambda f, \lambda c) = \frac{1}{\mu(H_c)} \int_{H_c} \lambda f(x) d\mu = \lambda M(f, c).$$

3. Temos $\{x \in \mathbb{R}^n \mid f(x) + \lambda \leq c + \lambda\} = \{x \in \mathbb{R}^n \mid f(x) \leq c\} = H_c$ para cada constante $\lambda \in \mathbb{R}$.

Consequentemente,

$$M(f + \lambda, c + \lambda) = \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda] d\mu = M(f, c) + \lambda.$$

■

Vimos na Definição 2.8 o conceito de valor médio para $c > \bar{c}$, onde $\bar{c}$ é o valor mínimo de f em $\mathbb{R}^n$, mas a pergunta natural que surge é: “o que ocorre quando $c = \bar{c}$?”. Quando isso ocorre pode acontecer de a medida $\mu(H_c)$ ser nula, assim a Definição 2.8 não teria sentido. A próxima definição contorna este problema.

Definição 2.14 (Def 1.3, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$, considere $c \geq \bar{c}$ e seja $\{c_k\}_{k \in \mathbb{N}}$ uma sequência decrescente cujo limite é c , isto é, $c_k \downarrow c$. Definimos então o valor médio $M(f, c)$ como:*

$$M(f, c) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu. \quad (2.14)$$

O limite dado em (2.14) está bem definido, pois como vimos anteriormente $\{M(f, c_k)\}_{k \in \mathbb{N}}$ é uma sequência decrescente e limitada.

Note ainda que o limite (2.14) não depende da sequência decrescente tomada. Suponha que tomemos outra sequência decrescente $\{b_k\}_{k \in \mathbb{N}}$, que também tende a c quando $k \rightarrow \infty$. Tomando a união das duas sequências $\{c_k\}_{k \in \mathbb{N}}$ e $\{b_k\}_{k \in \mathbb{N}}$, e reordenando-as, obtemos uma nova sequência decrescente $\{d_k\}_{k \in \mathbb{N}}$ que também tende a c . Portanto, temos uma nova sequência decrescente limitada $\left\{ \frac{1}{\mu(H_{d_k})} \int_{H_{d_k}} f(x) d\mu \right\}_{k \in \mathbb{N}}$ cujo limite existe. Assim, como duas subsequências da mesma sequência, os seguintes limites existem e são iguais:

$$\lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{b_k})} \int_{H_{b_k}} f(x) d\mu. \quad (2.15)$$

Pela Proposição 2.12, é claro que a Definição 2.14 estende a Definição 2.8 para o caso de $c \geq \bar{c}$. Além disso, temos que as Proposições 2.9, 2.10, 2.12 e 2.13 continuam sendo válidas para $c \geq \bar{c}$.

Temos todas as ferramentas necessárias para provar a caracterização do valor médio para a otimalidade global.

Teorema 2.15 (Condições do Valor Médio). *Para o problema (P') , as seguintes afirmações são equivalentes:*

- (a) um ponto $\bar{x} \in \mathbb{R}^n$ é um minimizador global com $\bar{c} = f(\bar{x})$ como o valor de mínimo global correspondente;
- (b) $M(f, c) \geq \bar{c}$ para $c > \bar{c}$;
- (c) $M(f, \bar{c}) = \bar{c}$.

Demonstração. Suponha por absurdo que $\bar{c}$ não seja o valor mínimo global de f e seja $\hat{c}$ o valor mínimo de f . Então, $\bar{c} - \hat{c} = 2\zeta > 0$. Temos ainda que $\mu(H_{\hat{c}+\zeta}) > 0$ e $\mu(H_{\bar{c}}) > 0$.

$$\begin{aligned}
M(f, \bar{c}) &= \frac{1}{\mu(H_{\bar{c}})} \int_{H_{\bar{c}}} f(x) d\mu \\
&= \frac{1}{\mu(H_{\bar{c}})} \left[\int_{H_{\bar{c}} \setminus H_{\hat{c}+\zeta}} f(x) d\mu + \int_{H_{\hat{c}+\zeta}} f(x) d\mu \right] \\
&\leq \frac{\bar{c}}{\mu(H_{\bar{c}})} [\mu(H_{\bar{c}}) - \mu(H_{\hat{c}+\zeta})] + \frac{\hat{c} + \zeta}{\mu(H_{\bar{c}})} \mu(H_{\hat{c}+\zeta}) \\
&= \bar{c} - \beta,
\end{aligned}$$

onde $\beta = \zeta \frac{\mu(H_{\hat{c}+\zeta})}{\mu(H_{\bar{c}})} > 0$, ou seja, $M(f, \bar{c}) < \bar{c}$. Isto estabelece a suficiência de (b) ou (c).

Para demonstrar a necessidade de (b) e (c), suponha que $\bar{c}$ é o valor mínimo global de f em $\mathbb{R}^n$. Então $f(x) \geq \bar{c}$ para todo $x \in \mathbb{R}^n$. Assim, para $c > \bar{c}$, temos

$$M(f, c) = \frac{1}{\mu(H_c)} \int_{H_c} f(x) d\mu \geq \frac{1}{\mu(H_c)} \int_{H_c} \bar{c} d\mu = \bar{c}. \quad (2.16)$$

Agora, tome uma sequência decrescente $\{c_k\}_{k \in \mathbb{N}}$ tal que $\lim_{k \rightarrow \infty} c_k = \bar{c}$. De (b) temos,

$$\lim_{k \rightarrow \infty} M(f, c_k) \geq \bar{c}, \quad (2.17)$$

isto é, $M(f, \bar{c}) \geq \bar{c}$. Mas $M(f, c) \leq c$ para $c \geq \bar{c}$, de modo que $M(f, \bar{c}) \leq \bar{c}$. Portanto,

$$M(f, \bar{c}) = \bar{c}.$$

■

Vejamos a seguir um exemplo que ilustra bem o Teorema 2.15.

Exemplo 2.16. Fixado $a > 0$, considere o seguinte problema

$$\begin{aligned}
&\text{Minimizar } f(x) = |x|^a \\
&\text{Sujeito a } x \in \mathbb{R}.
\end{aligned} \quad (P_5)$$

Para $c > 0$, $H_c = \{x \in \mathbb{R} \mid |x|^a \leq c\} = [-c^{1/a}, c^{1/a}]$,

$$M(f, c) = \frac{1}{2c^{1/a}} \int_{-c^{1/a}}^{c^{1/a}} |x|^a dx = \frac{1}{c^{1/a}} \int_0^{c^{1/a}} x^a dx = \frac{1}{1+a} c.$$

Seja $\bar{c}$ o valor mínimo global de f . Então

$$M(f, \bar{c}) = \frac{1}{1+a} \bar{c} = \bar{c}.$$

Assim, $\bar{c} = 0$ e $H_{\bar{c}} = \{0\}$.

Observação 2.17. Note que, um ponto $\bar{x}$ é minimizador de f se, e somente se é minimizador global de $f + a$ para todo $a \in \mathbb{R}$ e da função af , para $a > 0$.

De fato, seja $\bar{c} = f(\bar{x})$. De acordo com a Proposição 2.13,

$$M(f + a, \bar{c} + a) = M(f, \bar{c}) + a = \bar{c} + a,$$

$$M(af, a\bar{c}) = aM(f, \bar{c}) = a\bar{c}.$$

Assim, $\bar{c} + a$ e $a\bar{c}$ são valores de mínimo global e $\bar{x}$ é minimizador global para $f + a$ e af , respectivamente.

Como foi visto, embora o valor médio seja uma medida importante, ele sozinho não fornece toda a informação relevante. Veremos agora a definição de variância de uma função sobre o seu conjunto de nível, a qual representa a medida de quanto os dados estão dispersos em relação a seu valor médio. Ou seja, quanto maior a variância, maior será a dispersão nos dados.

Definição 2.18 (Def. 2.1, [23]). Considere $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$, e $c > \bar{c}$. Definimos

$$V(f, c) = \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c)]^2 d\mu \quad (2.18)$$

e

$$V_1(f, c) = \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - c]^2 d\mu \quad (2.19)$$

como sendo a variância e a variância modificada, respectivamente, da função f sobre o conjunto de nível H_c .

É claro que a variância (2.18) e a variância modificada (2.19) estão bem definidas, pois $\mu(H_c) > 0$ para $c > \bar{c}$ e f é contínua. A seguir veremos algumas propriedades da variância.

Proposição 2.19 (Propo. 2.1, [23]). Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Temos então as seguintes propriedades:

$$(a) \text{ Para } c > \bar{c}, \text{ temos } V(f, c) = \frac{1}{\mu(H_c)} \int_{H_c} [f(x)]^2 d\mu - [M(f, c)]^2;$$

$$(b) \ V(f, c) \geq 0 \text{ para } c > \bar{c};$$

$$(c) \ V(\lambda, c) = 0 \text{ para toda constante } \lambda \geq c;$$

$$(d) \ V(\lambda f, \lambda c) = \lambda^2 V(f, c) \text{ para toda constante } \lambda > 0 \text{ e } c > \bar{c};$$

$$(e) \ V(f + \lambda, c + \lambda) = V(f, c) \text{ para } c > \bar{c}.$$

Demonstração. (a)

$$\begin{aligned}
V(f, c) &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c)]^2 d\mu \\
&= \frac{1}{\mu(H_c)} \int_{H_c} \{[f(x)]^2 - 2M(f, c)f(x) + [M(f, c)]^2\} d\mu \\
&= \frac{1}{\mu(H_c)} \int_{H_c} [f(x)]^2 d\mu - \frac{2M(f, c)}{\mu(H_c)} \int_{H_c} f(x) d\mu + [M(f, c)]^2 \\
&= \frac{1}{\mu(H_c)} \int_{H_c} [f(x)]^2 d\mu - [M(f, c)]^2.
\end{aligned}$$

(b)-(c) são imediatas da definição da função V definida em (2.18).

(d) Como $\{x \in \mathbb{R}^n \mid \lambda f(x) \leq \lambda c\} = \{x \in \mathbb{R}^n \mid f(x) \leq c\} = H_c$ para $\lambda > 0$ e $c > \bar{c}$, temos

$$\begin{aligned}
V(\lambda f, \lambda c) &= \frac{1}{\mu(H_c)} \int_{H_c} [\lambda f(x) - M(\lambda f, \lambda c)]^2 d\mu \\
&= \frac{1}{\mu(H_c)} \int_{H_c} [\lambda f(x) - \lambda M(f, c)]^2 d\mu \\
&= \frac{\lambda^2}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c)]^2 d\mu \\
&= \lambda^2 V(f, c).
\end{aligned}$$

(e) Temos $\{x \in \mathbb{R}^n \mid f(x) + \lambda \leq c + \lambda\} = \{x \in \mathbb{R}^n \mid f(x) \leq c\} = H_c$ para cada constante $\lambda \in \mathbb{R}$.

Assim,

$$\begin{aligned}
V(f + \lambda, c + \lambda) &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda - M(f + \lambda, c + \lambda)]^2 d\mu \\
&= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda - M(f, c) - \lambda]^2 d\mu \\
&= V(f, c).
\end{aligned}$$

■

Como a variância $V(f, c)$ geralmente não é monótona em c , o seguinte lema se faz necessário.

Lema 2.20 (Lema 2.2, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$, e que $c > \bar{c}$, logo $\frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu$ é não decrescente em c para $c > \bar{c}$ se $f + \lambda \geq 0$.*

Demonstração. Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$ e suponha que $c_1 \geq c_2 \geq \bar{c}$. Então

$$\begin{aligned}
\frac{1}{\mu(H_{c_1})} \int_{H_{c_1}} [f(x) + \lambda]^2 d\mu &= \frac{1}{\mu(H_{c_1})} \int_{H_{c_1} \setminus H_{c_2}} [f(x) + \lambda]^2 d\mu + \frac{1}{\mu(H_{c_1})} \int_{H_{c_2}} [f(x) + \lambda]^2 d\mu \\
&\geq (c_2 + \lambda)^2 \frac{\mu(H_{c_1}) - \mu(H_{c_2})}{\mu(H_{c_1})} + \frac{1}{\mu(H_{c_1})} \int_{H_{c_2}} [f(x) + \lambda]^2 d\mu \\
&\geq \frac{\mu(H_{c_1}) - \mu(H_{c_2})}{\mu(H_{c_1})} \frac{1}{\mu(H_{c_2})} \int_{H_{c_2}} [f(x) + \lambda]^2 d\mu + \frac{1}{\mu(H_{c_1})} \int_{H_{c_2}} [f(x) + \lambda]^2 d\mu \\
&= \frac{1}{\mu(H_{c_2})} \int_{H_{c_2}} [f(x) + \lambda]^2 d\mu.
\end{aligned}$$

Assim, $\frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu$ é não decrescente em c para $c > \bar{c}$. ■

Agora, supondo que $\{c_k\}_{k \in \mathbb{N}}$ seja uma sequência decrescente que converge para c quando $k \rightarrow \infty$, temos que $\left\{ \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu \right\}$ é também uma sequência decrescente.

Proposição 2.21 (Propo. 2.3, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Se $\{c_k\}_{k \in \mathbb{N}}$ é uma sequência decrescente que converge para $c > \bar{c}$, isto é, $c_k \downarrow c$, então*

$$V(f, c) = \lim_{k \rightarrow \infty} V(f, c_k). \quad (2.20)$$

Demonstração. Pelo Lema 2.20 a sequência $\left\{ \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu \right\}$ é decrescente e limitada inferiormente por $\frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu$, o limite

$$\lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu$$

existe. Além disso,

$$\begin{aligned} 0 &\leq \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu - \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu \\ &\leq \left| \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu - \frac{1}{\mu(H_c)} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu \right| \\ &\quad + \left| \frac{1}{\mu(H_c)} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu - \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu \right| \\ &\leq \left| \left[\frac{1}{\mu(H_{c_k})} - \frac{1}{\mu(H_c)} \right] \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu \right| + \frac{1}{\mu(H_c)} \left| \int_{H_{c_k} \setminus H_c} [f(x) + \lambda]^2 d\mu \right|. \end{aligned}$$

A continuidade da medida μ e a continuidade absoluta da integral da função contínua $[f(x) + \lambda]^2$ implica que cada parcela do lado direito da desigualdade tende a zero. Ora de acordo com a Proposição 2.13, temos

$$[M(f + \lambda, c + \lambda)]^2 = [M(f, c) + \lambda]^2 = \lim_{k \rightarrow \infty} [M(f, c_k) + \lambda]^2 = \lim_{k \rightarrow \infty} [M(f(x) + \lambda, c_k + \lambda)]^2.$$

Assim, como $H_{c_k} = \{x \in \mathbb{R}^n \mid f(x) + \lambda \leq c_k + \lambda\}$, obtemos

$$\begin{aligned} \lim_{k \rightarrow \infty} V(f, c_k) &= \lim_{k \rightarrow \infty} V(f + \lambda, c_k + \lambda) \\ &= \lim_{k \rightarrow \infty} \left[\frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu - [M(f + \lambda, c_k + \lambda)]^2 \right] \\ &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) + \lambda]^2 d\mu - [M(f + \lambda, c + \lambda)]^2 \\ &= V(f + \lambda, c + \lambda) = V(f, c). \end{aligned}$$

■

A seguinte proposição é útil para obter propriedades da variância modificada a partir da Proposição 2.19.

Proposição 2.22 (Propo. 2.6, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. Para $c > \bar{c}$, temos*

$$V_1(f, c) = V(f, c) + [M(f, c) - c]^2. \quad (2.21)$$

Demonstração.

$$\begin{aligned} V_1(f, c) &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - c]^2 d\mu \\ &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c) + M(f, c) - c]^2 d\mu \\ &= \frac{1}{\mu(H_c)} \int_{H_c} [[f(x) - M(f, c)]^2 + 2[M(f, c) - c][f(x) - M(f, c)] + [M(f, c) - c]^2] d\mu \\ &= \frac{1}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c)]^2 d\mu + \frac{2[M(f, c) - c]}{\mu(H_c)} \int_{H_c} [f(x) - M(f, c)] d\mu + [M(f, c) - c]^2 \\ &= V(f, c) + [M(f, c) - c]^2. \end{aligned}$$

■

Vejamos a seguir algumas propriedades da variância modificada.

Proposição 2.23. *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$. A variância modificada tem as seguintes propriedades:*

- (i) $V_1(f, c) \geq V(f, c)$ para $c > \bar{c}$;
- (ii) Se $\{c_k\}_{k \in \mathbb{N}}$ é uma sequência decrescente que converge para $c > \bar{c}$, então $\lim_{c_k \downarrow c} V_1(f, c_k) = V_1(f, c)$.

Note que, como no caso do valor médio, a variância e a variância modificada foram definidas somente para valores de $c > \bar{c}$, logo pode-se definir a variância $V(f, c)$ e a variância modificada $V_1(f, c)$ de f sobre o seu conjunto de nível para todos $c \geq \bar{c}$ utilizando limites, como veremos na seguinte definição.

Definição 2.24 (Def 2.3, [23]). *Seja $\bar{c}$ o valor mínimo de f em $\mathbb{R}^n$ e $c \geq \bar{c}$. Considere $\{c_k\}_{k \in \mathbb{N}}$ uma sequência decrescente que converge para c quando $k \rightarrow \infty$. Define-se*

$$V(f, c) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) - M(f, c)]^2 d\mu \quad (2.22)$$

e

$$V_1(f, c) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) - c]^2 d\mu \quad (2.23)$$

a variância e variância modificada, respectivamente, de f sobre o seu conjunto de nível H_c .

Com todas as ferramentas em mãos pode-se provar a seguinte caracterização da variância e da variância modificada para a otimalidade global.

Teorema 2.25. Para o problema (P') , as seguintes afirmações são equivalentes:

- (a) Um ponto $\bar{x} \in \mathbb{R}^n$ é um minimizador global com $\bar{c} = f(\bar{x})$ como o valor mínimo global de f correspondente;
- (b) $V(f, \bar{c}) = 0$;
- (c) $V_1(f, \bar{c}) = 0$.

Demonstração. (b) $\Rightarrow$ (a) Suponhamos, por absurdo, que $\bar{x}$ não é um minimizador global, isto é, $\bar{c} = f(\bar{x})$ não é o valor de mínimo global de f . Temos que $\mu(H_{\bar{c}}) > 0$. Temos, por hipótese, que

$$V(f, \bar{c}) = \frac{1}{\mu(H_{\bar{c}})} \int_{H_{\bar{c}}} [f(x) - M(f, \bar{c})]^2 d\mu = 0.$$

Consideremos o conjunto $A := \{x \in H_{\bar{c}} \mid f(x) \neq M(f, \bar{c})\}$ onde $\mu(A) = 0$. Então $f(x) = M(f, \bar{c})$ para quase todo ponto $x \in H_{\bar{c}}$. Mas $\bar{x} \in H_{\bar{c}} = \{x \in \mathbb{R}^n \mid f(x) \leq \bar{c}\}$. Portanto, $\bar{c} = f(\bar{x}) = M(f, \bar{c})$, isto é, $\bar{x}$ é um minimizador global de f . Isto é uma contradição.

(a) $\Rightarrow$ (b) Suponhamos, que $\bar{x}$ é um minimizador global, e por absurdo, que $V(f, \bar{c}) = 2\zeta > 0$ com $\bar{c} = f(\bar{x})$. Seja $\{c_k\}_{k \in \mathbb{N}}$ uma sequência decrescente a qual tende a $\bar{c}$ quando $k \rightarrow \infty$. Portanto, existe um inteiro positivo N tal que

$$V(f, c_k) > \zeta \text{ para } k > N.$$

Isto significa que

$$\begin{aligned} \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda]^2 d\mu &= \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} ([f(x)]^2 + 2\lambda f(x) + \lambda^2) d\mu \\ &= \frac{1}{\mu(H_{c_k})} \left[\int_{H_{c_k}} [f(x)]^2 d\mu + 2\lambda \int_{H_{c_k}} f(x) d\mu + \lambda^2 \mu(H_{c_k}) \right] \\ &= \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x)]^2 d\mu + 2\lambda M(f, c_k) + \lambda^2 \\ &= V(f, c_k) + M(f, c_k)^2 + 2\lambda M(f, c_k) + \lambda^2 \\ &= V(f, c_k) + [M(f, c_k) + \lambda]^2 \\ &> \frac{V(f, c_k)}{2} + [M(f, c_k) + \lambda]^2 \\ &> \frac{\zeta}{2} + \left[\frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) + \lambda] d\mu \right]^2, \quad \text{para } k > N, \end{aligned}$$

onde λ é um número real tal que $f(x) + \lambda \geq 0$. Como $\bar{c} \leq f(x) \leq c_k$ para $x \in H_{c_k}$, temos

$$(c_k + \lambda)^2 > (c_k + \lambda)^2 + \frac{\zeta}{2}.$$

Tomando $k \rightarrow \infty$, temos

$$(\bar{c} + \lambda)^2 \geq (\bar{c} + \lambda)^2 + \zeta,$$

o que claramente é um absurdo.

- Se $V_1(f, \bar{c}) = 0$, então $V(f, \bar{c}) = 0$, como $0 \leq V(f, \bar{c}) \leq V_1(f, \bar{c})$. Assim, $\bar{x}$ é um minimizador global e $\bar{c}$ é o valor de mínimo global de f .
- Se $\bar{x}$ é um minimizador global de f , então $M(f, \bar{c}) = \bar{c} = f(\bar{x})$. Isto significa que $V_1(f, \bar{c}) = V(f, \bar{c})$. Portanto, $V_1(f, \bar{c}) = V(f, \bar{c}) = 0$.

■

Exemplo 2.26. Voltemos ao Exemplo 2.16. Tíñhamos que $f(x) = |x|^a$, fixado $a > 0$ e que $H_c = [-c^{1/a}, c^{1/a}]$.

$$\begin{aligned}
 V_1(f, c) &= \frac{1}{2c^{1/a}} \int_{-c^{1/a}}^{c^{1/a}} [|x|^a - c]^2 dx \\
 &= \frac{1}{c^{1/a}} \left[\int_0^{c^{1/a}} x^{2a} - 2c \int_0^{c^{1/a}} x^a dx \right] + c^2 \\
 &= \frac{c^2}{2a+1} - \frac{2c^2}{a+1} + c^2 = \frac{2a^2}{(1+a)(1+2a)} c^2.
 \end{aligned}$$

Seja $\bar{c}$ o valor mínimo global de f . Então

$$V_1(f, \bar{c}) = \frac{2a^2}{(1+a)(1+2a)} \bar{c}^2 = 0.$$

Assim, $\bar{c} = 0$ e $H_{\bar{c}} = \{0\}$.

■

Consideremos agora o problema (P) , com $X \subset \mathbb{R}^n$. Seja, pois,

$$\begin{aligned}
 &\text{Minimizar } f(x) \\
 &x \in X.
 \end{aligned} \tag{P}$$

É possível estender os resultados vistos anteriormente para o caso de problemas com restrições. Para este fim, necessitamos do conceito de conjunto robusto.

Definição 2.27. Um conjunto D é dito ser robusto se e somente se

$$cl(D) = cl(int D). \tag{2.24}$$

Definição 2.28. Um ponto $x \in cl(D)$ é dito ser um ponto robusto para o conjunto D se e somente se para cada vizinhança

$$V \cap int D \neq \emptyset. \tag{2.25}$$

Um conjunto aberto é robusto. O conceito de robustez é de certa forma uma generalização do conceito de conjuntos abertos. Um conjunto fechado pode não ser robusto.

Observação 2.29. Observe que o conceito de robustez está profundamente relacionado à topologia dada. Um subconjunto de números inteiros não é robusto em $\mathbb{R}$, mas é robusto em relação à topologia discreta no conjunto de números inteiros.

O conceito de pontos robustos é definido apenas em $cl(D)$ pois se $x \notin cl(D)$, existe uma vizinhança V de x tal que $V \cap cl(D) = \emptyset$. Assim, os pontos que não estão contidos no $cl(D)$ são sempre não robustos.

Teorema 2.30 (Teo. 2.1, [55]). *Um conjunto D é robusto se, e somente se, cada ponto de D é robusto para D .*

Teorema 2.31 (Teo. 2.2, [55]). *Suponha que x é robusto para D e $x \in int\ G$, então x é robusto para $D \cap int\ G$ e também para $D \cap G$.*

Teorema 2.32 (Teo. 2.3, [55]). *A união de conjuntos robustos é robusto.*

Teorema 2.33 (Teo. 2.4, [55]). *A interseção de um conjunto robusto e um conjunto aberto é robusto.*

As seguintes afirmações são válidas e podem ser provadas usando os teoremas acima.

1. Se D é robusto então $cl(D)$ também é robusto.
2. Um ponto x é robusto para D se e somente se $x \in cl(int\ D)$.
3. Se D é robusto e F é um conjunto fechado, então $D \setminus F$ é robusto.

Suponha que o conjunto $X \subset \mathbb{R}^n$ é robusto. Considere o problema de encontrar o valor mínimo global de f sobre X , ou seja

$$\begin{aligned} &\text{Minimizar } f(x) \\ &x \in X. \end{aligned} \tag{P*}$$

Podemos construir uma medida adequada, derivada do espaço $(\mathbb{R}^n, \Omega, \mu)$. A família de conjuntos $\Omega_X = \{X \cap B | B \in \Omega\}$ é uma σ -álgebra. A medida μ_X é definida por

$$\mu_X(A) = \mu(A \cap X) \text{ para } A \in \Omega, \tag{2.26}$$

portanto temos que (X, Ω_X, μ_X) é um espaço mensurável.

Definição 2.34. *Suponha que $X \subset \mathbb{R}^n$ é um conjunto robusto. O espaço mensurável (X, Ω_X, μ_X) é chamado um espaço de medida de rejeição.*

Vejamos que as versões do valor médio, variância e variância modificada de uma função sobre seus conjuntos de nível para um espaço de medida de rejeição não diferem muito das vistas anteriormente para problemas irrestritos.

Definição 2.35. *Seja $\bar{c}$ o valor mínimo de f em X . Suponha que $\{c_k\}$ é uma sequência decrescente que converge para $c \geq \bar{c}$. Os limites*

$$M(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} f(x) d\mu, \quad (2.27)$$

$$V(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} [f(x) - M(f, c, X)]^2 d\mu, \quad (2.28)$$

$$V_1(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} [f(x) - c]^2 d\mu, \quad (2.29)$$

são chamados respectivamente de valor médio de rejeição, variância de rejeição e variância modificada de rejeição de f sobre $H_c \cap X$.

Teorema 2.36. *Para o problema (P^*) , as seguintes afirmações são equivalentes:*

- (a) $\bar{x} \in X$ é um minimizador global e $\bar{c} = f(\bar{x})$ é o valor mínimo global correspondente;
- (b) $M(f, c, X) \geq \bar{c}$ para $c > \bar{c}$;
- (c) $M(f, \bar{c}, X) = \bar{c}$;
- (d) $V(f, \bar{c}, X) = 0$;
- (e) $V_1(f, \bar{c}, X) = 0$.

Exemplo 2.37. *Considere o problema (P^*) onde $f(x) = x$ e o conjunto robusto $X = [1, 2]$. Para qualquer c , o conjunto de nível $H_c = \{x \mid x \leq c\} = (-\infty, c]$, de modo que*

$$M(f, c, X) = \frac{1}{\mu(H_c \cap X)} \int_{H_c \cap X} f(x) d\mu = \frac{1}{c-1} \int_1^c x dx = \frac{1}{2(c-1)}(c^2 - 1) = \frac{c+1}{2} \quad (2.30)$$

para $c > 1$. Aplicando a condição do valor médio, temos que

$$M(f, c, X) = \bar{c} = \frac{\bar{c} + 1}{2}. \quad (2.31)$$

Assim, $\bar{c} = 1$ e $\bar{H} = \{1\}$.

■

2.1.2 Caso em que a função objetivo é descontínua

Os resultados da subseção anterior podem ser generalizados a uma certa classe de problemas cuja função objetivo é descontínua. Para isto, necessitamos do conceito de função relativamente robusta.

Definição 2.38. *Dizemos que a função f é relativamente robusta com respeito a X em $x_0 \in cl(X)$ se $x_0 \in F_c = \{x \mid f(x) < c\}$ implica x_0 é robusto em $F_c \cap X$.*

Consideraremos as seguintes hipóteses para o problema (P):

A3. $X \subset \mathbb{R}^n$ é compacto e não vazio;

A4. $f : X \rightarrow \mathbb{R}$ é semicontínua inferiormente;

A5. f é relativamente robusta a X em um minimizador global.

Lema 2.39. *Assuma que as hipóteses A3, A4 e A5 sejam válidas. Se $H_{\bar{c}} \cap X$ é não vazio e $\mu(H_{\bar{c}} \cap X) = 0$, então $\bar{c}$ é o valor mínimo de f em X .*

Demonstração. Suponha, por contradição, que $\bar{c}$ não é o valor de mínimo global e $\hat{c} < \bar{c}$ é. Seja $2\zeta = \bar{c} - \hat{c} > 0$. Existe um minimizador global $\hat{x}$ tal que $\hat{c} = f(\hat{x})$ e f é relativamente robusta a X em $\hat{x}$. Temos $\hat{x} \in F_{\bar{c}-\zeta}$. Assim, considerando V uma vizinhança de $\hat{x}$, temos que $V \cap \text{int}(X \cap F_{\bar{c}-\zeta})$ é não vazia. Portanto $\text{int}(X \cap F_{\bar{c}-\zeta})$ é não vazio e

$$\text{int}(X \cap F_{\bar{c}-\zeta}) \subset X \cap H_{\bar{c}},$$

o qual implica que

$$\mu(X \cap H_{\bar{c}}) \geq \mu(X \cap F_{\bar{c}-\zeta}) > 0,$$

o que é uma contradição. ■

Definição 2.40. *Seja $\bar{c}$ o valor mínimo de f em X , $c > \bar{c}$ e suponha que as hipóteses A3, A4 e A5 sejam válidas. Definimos o valor médio, variância e variância modificada, respectivamente, como segue*

$$M(f, c, X) = \frac{1}{\mu(X \cap H_c)} \int_{X \cap H_c} f(x) d\mu, \quad (2.32)$$

$$V(f, c, X) = \frac{1}{\mu(X \cap H_c)} \int_{X \cap H_c} [f(x) - M(f, c, X)]^2 d\mu, \quad (2.33)$$

$$V_1(f, c, X) = \frac{1}{\mu(X \cap H_c)} \int_{X \cap H_c} [f(x) - c]^2 d\mu. \quad (2.34)$$

Como a função f é mensurável, $X \cap H_c$ é compacto e $\mu(X \cap H_c) > 0$, então as integrais dadas na definição anterior estão bem definidas. Além disso, quando $c = \bar{c}$, $\mu(X \cap H_c)$ pode ser igual a zero. A Definição 2.40 pode ser estendida por um processo de limite para o caso em que $c \geq \bar{c}$.

Definição 2.41. *Sob as suposições da Definição 2.40, define-se valor médio, variância e variância modificada, respectivamente, para $c \geq \bar{c}$ como segue:*

$$M(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(X \cap H_{c_k})} \int_{X \cap H_{c_k}} f(x) d\mu, \quad (2.35)$$

$$V(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(X \cap H_{c_k})} \int_{X \cap H_{c_k}} [f(x) - M(f, c, X)]^2 d\mu, \quad (2.36)$$

$$V_1(f, c, X) = \lim_{k \rightarrow \infty} \frac{1}{\mu(X \cap H_{c_k})} \int_{X \cap H_{c_k}} [f(x) - c]^2 d\mu. \quad (2.37)$$

Os limites existem e são independentes das escolhas de $\{c_k\}$. Os conceitos estendidos estão bem definidos e consistentes com os da Definição 2.40. Com esses conceitos, a otimalidade global é caracterizada da seguinte forma.

Teorema 2.42. *Sob as hipóteses A3, A4 e A5, as seguintes afirmações são equivalentes:*

- $\bar{x} \in X$ é um minimizador global de (P) e $\bar{c} = f(\bar{x})$ é o mínimo global correspondente;
- $M(f, \bar{c}, X) = \bar{c}$;
- $V(f, \bar{c}, X) = 0$;
- $V_1(f, \bar{c}, X) = 0$.

Vejam os a seguir, que ao utilizarmos as técnicas de escalarização, os quais foram apresentados no Capítulo 1, conseguimos estender os resultados dados nesta seção para problemas multiobjetivos.

2.2 Problema Multiobjetivo

Consideremos primeiramente o problema de otimização multiobjetivo dado por:

$$\begin{aligned} &\text{Minimizar } F(x) = (f_1(x), \dots, f_r(x)) \\ &x \in X, \end{aligned} \quad \text{ldots} \quad (P_v)$$

onde $X \subset \mathbb{R}^n$ é um conjunto não vazio e $f_\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, para $\ell = 1, \dots, r$, funções de valores reais.

Assim como no caso mono objetivo, temos dois casos a serem tratados, o caso em que F é contínua e o caso em que F pode não ser contínua.

2.2.1 Caso em que as funções objetivo são contínuas

Consideremos que F seja contínua, ou seja, suas componentes f_ℓ , $\ell = 1, \dots, r$ em (P_v) sejam contínuas, e o conjunto X é compacto.

Note que, o resultado dado no Teorema 2.3 não pode ser aplicado diretamente nas funções obtidas através das técnicas de escalarização, pois as funções $\Phi_w = \sum_{\ell=1}^r w_\ell f_\ell(x)$ e $\Psi_w = \|f(x) - f^*\|_w$ dadas nos problemas (WS) e (WCS_w) , respectivamente, não tem como contradomínio o intervalo $(-\infty, 0]$. Porém, sabemos que ambas as funções Φ_w e Ψ_w são funções contínuas sobre o compacto X . Sendo assim, pelo Teorema de Weierstrass, temos que $\Phi_w(x) \leq M_1$, assim como $\Psi_w(x) \leq M_2$, onde M_1 e M_2 são constantes a serem determinadas.

Observe que, $\Phi_w - M_1 \leq 0$, assim como $\Psi_w - M_2 \leq 0$. Portanto, pelo Teorema 2.3 obtemos os seguintes resultados.

Teorema 2.43. *Sejam X um conjunto compacto com interior não vazio, $\tilde{\Phi}_w : X \rightarrow (-\infty, 0]$ uma função contínua dada por $\tilde{\Phi}_w(x) := \Phi_w(x) - M_1$ e $\bar{x} \in X$ tal que $\tilde{\Phi}_w(\bar{x}) \neq 0$. Então*

$$\Upsilon_w(t) = \int_X \left[\frac{\tilde{\Phi}_w(x)}{\tilde{\Phi}_w(\bar{x})} \right]^t d\mu \rightarrow \infty \quad (2.38)$$

quando $t \rightarrow \infty$ se, e somente se, $\bar{x}$ não é solução global de (WS) .

Teorema 2.44. *Sejam X um conjunto compacto com interior não vazio, $\tilde{\Psi}_w(x) := \Psi_w(x) - M_2$ e $\bar{x} \in X$ tal que $\tilde{\Psi}_w(\bar{x}) \neq 0$. Então*

$$\Upsilon_w(t) = \int_X \left[\frac{\tilde{\Psi}_w(x)}{\tilde{\Psi}_w(\bar{x})} \right]^t d\mu \rightarrow \infty \quad (2.39)$$

quando $t \rightarrow \infty$ se, e somente se, $\bar{x}$ não é solução global de (WCS_w) .

Note, que as integrais dadas nestes teoremas estão bem definidas, pois ambas as funções $\frac{\tilde{\Phi}_w(x)}{\tilde{\Phi}_w(\bar{x})}$ e $\frac{\tilde{\Psi}_w(x)}{\tilde{\Psi}_w(\bar{x})}$ são contínuas e não negativas. Ainda, os mesmos pontos que minimizam Φ_w minimizam $\tilde{\Phi}_w$, assim como, os mesmos pontos que minimizam Ψ_w minimizam $\tilde{\Psi}_w$.

Agora, considerando o conjunto $W^* = \{w \in \mathbb{R}^r \mid w_\ell > 0, \forall \ell = 1, \dots, r, \sum_{\ell=1}^r w_\ell = 1\}$, e os Teoremas 1.56, 1.60 e 1.61 obtemos os seguintes resultados.

Teorema 2.45. *Seja X um compacto com interior não vazio e $\bar{x} \in X$ tal que $\tilde{\Phi}_w(\bar{x}) \neq 0$. Se para algum $w \in W^*$ (respectivamente, $w \in W$), temos que a sequência de integrais dada por $\int_X \left[\frac{\tilde{\Phi}_w(x)}{\tilde{\Phi}_w(\bar{x})} \right]^t d\mu$ converge então $\bar{x}$ é solução ótima de Pareto (respectivamente Pareto fraca) para o problema (P_v) .*

Demonstração. Segue imediatamente do Teorema 2.43 que $\bar{x} \in X$ é solução global se, e somente se $\int_X \left[\frac{\tilde{\Phi}_w(x)}{\tilde{\Phi}_w(\bar{x})} \right]^t d\mu$ converge. Além disso, pelo Teorema 1.56, se $\bar{x}$ é solução de (WS) para algum $w \in W^*$ então $\bar{x}$ é solução ótima de Pareto para (P_v) . Assim como, pelo Teorema 1.54, se $\bar{x}$ é solução de (WS) para algum $w \in W$ então $\bar{x}$ é solução ótima de Pareto fraca para (P_v) . ■

Não temos a garantia que a recíproca de Teorema 2.45 seja válida sem a suposição de convexidade ou convexidade generalizada.

Teorema 2.46. *Seja X um compacto com interior não vazio e $\bar{x} \in X$ tal que $\tilde{\Psi}_w(\bar{x}) \neq 0$. Temos que $\bar{x}$ é solução ótima de Pareto para (P_v) se, e somente se, existe $w \in W^*$ tal que a sequência de integrais dada por $\int_X \left[\frac{\tilde{\Psi}_w(x)}{\tilde{\Psi}_w(\bar{x})} \right]^t d\mu$ converge.*

Demonstração. Pelo Teorema 2.44, $\bar{x}$ é solução global de (WCS_w) se, e somente se $\int_X \left[\frac{\tilde{\Psi}_w(x)}{\tilde{\Psi}_w(\bar{x})} \right]^t d\mu$ converge.

($\Rightarrow$) Segue imediatamente dos Teoremas 2.44 e 1.56, se $\bar{x} \in X$ é solução ótima de Pareto para (P_v) , então existe $w \in W^*$ tal que é solução do problema (WCS_w) .

($\Leftarrow$) Segue imediatamente dos Teoremas 2.44 e 1.60, se $\bar{x} \in X$ é solução do problema (WCS_w) , para algum $w \in W^*$, então $\bar{x}$ é solução ótima de Pareto para (P_v) . ■

Note que, com simples estimativas, podemos encontrar quem são as constantes M_1 e M_2 , cujas existências são garantidas pelo Teorema de Weierstrass.

Para o caso do problema ponderado (WS) temos:

$$\begin{aligned}\Phi_w(x) &= \sum_{\ell=1}^r w_\ell f_\ell(x) \\ &\leq r \max_{\ell} w_\ell f_\ell(x) = M_1.\end{aligned}$$

E para o caso do problema Chebyshev ponderado (WCS_w) , denotando $\widetilde{M} = \min_{\ell} \{w_\ell f_\ell^*\}$ e $\widetilde{N}_\ell = \max_{x \in X} w_\ell f_\ell(x)$, $\ell \in \{1, \dots, r\}$, temos

$$\begin{aligned}\Psi_w(x) &= \max_{\ell} [w_\ell (f_\ell(x) - f_\ell^*)] \\ &= \max_{\ell} [w_\ell f_\ell(x) - w_\ell f_\ell^*] \\ &\leq \max_{\ell} [w_\ell f_\ell(x) - \widetilde{M}] \\ &\leq \max_{\ell} \widetilde{N}_\ell - \widetilde{M} = M_2.\end{aligned}$$

Utilizando a técnica de escalarização de Chebyshev, obtemos caracterizações de otimalidade para o problema multiobjetivo (P_v) utilizando o valor médio, a variância e a variância modificada. Assumimos que as funções f_ℓ , $\ell = 1, \dots, r$, são contínuas e o conjunto X é compacto. Para cada $w \in \mathbb{R}^r$, $w > 0$ definem-se

$$\begin{aligned}f_\ell^* &= \min_{x \in X} \{f_\ell(x)\} \\ \Psi_w(x) &= \max_{\ell=1, \dots, r} \{w_\ell (f_\ell(x) - f_\ell^*)\}.\end{aligned}$$

Vimos que, se $\bar{x} \in X$ é solução eficiente se e somente se existe $w > 0$ tal que $\bar{x}$ é minimizador global de Ψ_w em X . Além disso é claro que Ψ_w é contínua.

Destas observações e dos Teoremas 2.8 e 2.18 obtemos as seguintes proposições:

Proposição 2.47 (Condições do Valor Médio). *Suponha que as funções f_ℓ , $\ell = 1, \dots, r$, são contínuas e o conjunto X é compacto. Seja $\bar{x} \in X$. As seguintes afirmações são equivalentes:*

- (i) $\bar{x}$ é solução Pareto eficiente de (P_v) ;
- (ii) Existe $w > 0$ tal que $\bar{x}$ minimiza Ψ_w em X e $\bar{c} = \Psi_w(\bar{x})$;
- (iii) Existe $w > 0$ tal que $M(\Psi_w, c, X) \geq \bar{c}$ para todo $c > \bar{c} = \Psi_w(\bar{x})$;

(iv) Existe $w > 0$ tal que $M(\Psi_w, \bar{c}, X) = \bar{c} = \Psi_w(\bar{x})$;

onde

$$M(\Psi_w, c, X) = \frac{1}{\mu(H_c \cap X)} \int_{H_c \cap X} \Psi_w(x) d\mu$$

é o valor médio de Ψ_w sobre o seu conjunto de nível $H_c = \{x \in X \mid \Psi_w(x) \leq c\}$.

Proposição 2.48 (Condições da Variância). *Suponha que as funções f_ℓ , $\ell = 1, \dots, r$, são contínuas e o conjunto X é compacto. Seja $\bar{x} \in X$. As seguintes afirmações são equivalentes:*

(i) $\bar{x}$ é solução Pareto eficiente de (P_v) ;

(ii) Existe $w > 0$ tal que $\bar{x}$ minimiza Ψ_w em X e $\bar{c} = \Psi_w(\bar{x})$;

(iii) Existe $w > 0$ tal que $V(\Psi_w, \bar{c}, X) = 0$, para $\bar{c} = \Psi_w(\bar{x})$;

(iv) Existe $w > 0$ tal que $V_1(\Psi_w, \bar{c}, X) = 0$, para $\bar{c} = \Psi_w(\bar{x})$.

onde

$$\begin{aligned} V(\Psi_w, c, X) &= \frac{1}{\mu(H_c \cap X)} \int_{H_c \cap X} [\Psi_w(x) - M(\Psi_w, c, X)]^2 d\mu, \\ V_1(\Psi_w, c, X) &= \frac{1}{\mu(H_c \cap X)} \int_{H_c \cap X} [\Psi_w(x) - c]^2 d\mu, \end{aligned}$$

são a variância e a variância modificada, respectivamente, de Ψ_w sobre o seu conjunto de nível H_c .

2.2.2 Caso em que as funções objetivo podem ser descontínuas

Nesta seção consideramos o caso em que F não é necessariamente contínua. Observamos que os resultados das Proposições 2.47 e 2.48 se estendem a uma certa classe de problemas não contínuos.

Consideremos as seguintes hipóteses:

A3. X é compacto e não vazio;

A4. Existe $w > 0$ tal que a função Ψ_w é semicontínua inferiormente e relativamente robusta a X em um minimizador de Ψ_w .

Do Teorema 2.42 obtemos os seguintes resultados:

Proposição 2.49. *Seja $\bar{x} \in X$. Sob as hipóteses A3 e A4, as seguintes afirmações são equivalentes:*

(i) $\bar{x}$ é solução Pareto eficiente de (P_v) ;

(ii) $\bar{x}$ minimiza Ψ_w em X e $\bar{c} = \Psi_w(\bar{x})$;

(iii) $M(\Psi_w, c, X) \geq \bar{c}$, para todo $c > \bar{c} = \Psi_w(\bar{x})$;

$$(vi) \quad M(\Psi_w, \bar{c}, X) = \bar{c} = \Psi_w(\bar{x}).$$

Proposição 2.50. *Seja $\bar{x} \in X$. Sob as hipóteses A3 e A4, as seguintes afirmações são equivalentes:*

- (i) $\bar{x}$ é solução Pareto eficiente de (P_v) ;
- (ii) $\bar{x}$ minimiza Ψ_w em X e $\bar{c} = \Psi_w(\bar{x})$;
- (iii) $V(\Psi_w, \bar{c}, X) = 0$, para $\bar{c} = \Psi_w(\bar{x})$;
- (vi) $V_1(\Psi_w, \bar{c}, X) = 0$, para $\bar{c} = \Psi_w(\bar{x})$.

Note que, no caso da escalarização da soma ponderada não conseguimos a robustez da função Φ_w , pois a soma de funções robustas pode não ser robusta, como mostra o seguinte exemplo.

Exemplo 2.51. *Considere as seguintes funções:*

$$\xi_1(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \quad e \quad \xi_2(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

onde ξ_1 e ξ_2 são funções robustas. Mas a soma delas

$$\xi(x) = \xi_1(x) + \xi_2(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

é não robusta em $x = 0$.

Com os resultados vistos até o momento, no Capítulo 3 apresentamos algoritmos para o caso mono objetivo e ainda propomos um algoritmo para problemas multiobjetivos. Resultados numéricos são apresentados.

Capítulo 3

ALGORITMOS E TESTES NUMÉRICOS

Neste capítulo veremos o método de Monte Carlo o qual é um método estatístico que utiliza sequências de números aleatórios para fazer simulações numéricas. Tal método será aplicado nas implementações dos algoritmos para encontrar minimizadores globais de uma função sobre um conjunto, para problemas mono objetivo. Ainda, neste capítulo propomos um algoritmo para busca de soluções de Pareto para problemas multiobjetivo e comprovamos sua eficácia em uma coleção de problemas teste.

3.1 Método de Monte Carlo

Este método está baseado em conceitos e princípios da mecânica estatística e pode ser utilizado de diferentes maneiras: como uma técnica para calcular integrais, para modelar processos estocásticos, calcular propriedades de estado e para simular sistemas de partículas interagentes.

Além de diversas aplicações na Física, Química e Medicina, atualmente o método de Monte Carlo é aplicado em diversas áreas, como em: Finanças, na modelagem e simulação de um mercado de opção, [50, 5]; Engenharia, na gestão de portfólio de uma empresa de seguros ou na análise de um problema de estoque, [43, 46]; Biologia, na análise de sistemas de tratamento de câncer ou para definir estratégias de otimização e paralelização de uma infecção pelo HIV, [34, 12].

Nós estamos particularmente interessados na aplicação do método de Monte Carlo para o cálculo de integrais unidimensionais e multidimensionais. Nossas principais referências no tema são os livros [27, 48].

3.1.1 Caso unidimensional

Vejamos uma integração simples para fixar melhor o método. Considere o problema de calcular a integral de uma função f de uma única variável no intervalo $[a, b]$, ou seja,

$$I = \int_a^b f(x) dx. \quad (3.1)$$

Considere a mudança de variável da forma:

$$z = \frac{x - a}{b - a} \quad (3.2)$$

que transforma o intervalo $[a, b]$ em $[0, 1]$. Equivalentemente,

$$x = a + (b - a)z \quad (3.3)$$

e a integral dada em (3.1) passa a ser escrita como:

$$I = \int_0^1 f(a + (b - a)z)(b - a)dz. \quad (3.4)$$

Denotando $g(z) = (b - a)f(a + (b - a)z)$, temos

$$I = \int_0^1 g(z)dz. \quad (3.5)$$

Discretizando o intervalo $[0, 1]$ em N pontos da forma $z_1, z_2, \dots, z_N$ uniformemente distribuídos temos que se esses N pontos são variáveis aleatórias uniformemente distribuídas em $[0, 1]$, e que $g(z_1), g(z_2), \dots, g(z_N)$ são variáveis aleatórias independentes. Assim, pode-se aproximar I gerando uma grande quantidade de números aleatórios z_i e tomar como aproximação o valor médio de $g(z_i)$, ou seja

$$I = \int_0^1 g(z)dz \approx \frac{1}{N} \sum_{i=1}^N g(z_i). \quad (3.6)$$

A precisão da estimativa I é o erro padrão da média $\sigma_P(I)$, calculado na forma:

$$\sigma_P(I) = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (g(z_i) - I)^2}. \quad (3.7)$$

A seguir temos o algoritmo para o cálculo de integrais de funções unidimensionais.

Algoritmo 3.1. Método de Monte Carlo - Unidimensional

Passo 0: Dado: $N \in \mathbb{N}$ pontos da discretização.

Passo 1: Discretize o intervalo $[0, 1]$ uniformemente.

Passo 2: Faça a mudança de variáveis

$$x = a + (b - a)z.$$

Passo 3: Calcule o valor aproximado da integral dado por:

$$I = \int_0^1 g(z)dz \approx \frac{1}{N} \sum_{i=1}^N g(z_i).$$

Passo 4: Calcule o erro cometido

$$\sigma_P(I) = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (g(z_i) - I)^2}.$$

Caso a discretização do intervalo $[0, 1]$ seja de forma randômica deve-se rodar o algoritmo m vezes, cada vez com o mesmo número N de pontos. Para cada experimento obtemos o valor da média I e calculamos o desvio padrão da média. Como a sequência de números aleatórios é diferente, I varia para cada experimento. Definimos então I_j como o valor I da integral obtido na j -ésima rodada do Algoritmo 3.1, e $\sigma_P(I)_j$ sendo o erro cometido na j -ésima rodada do Algoritmo 3.1. Calculamos então:

- A média dos valores de I_j , que será o valor aproximado da integral, ou seja

$$Integral = \frac{1}{m} \sum_{j=1}^m I_j. \quad (3.8)$$

- O erro médio

$$Erro = \frac{1}{m} \sum_{j=1}^m \sigma_P(I)_j. \quad (3.9)$$

- O desvio padrão

$$\sigma_m = \sqrt{\frac{1}{m} \sum_{j=1}^m (I_j - Integral)^2}. \quad (3.10)$$

Exemplo 3.1. Considere o problema de calcular a integral (3.1) em que $f(x) = \sqrt{x}$ e $[a, b] = [0, 1]$. O valor analítico desta integral é $2/3$. Vamos calculá-la usando o método de Monte Carlo descrito no Algoritmo 3.1 com três tipos de discretizações do intervalo $[0, 1]$ em $N > 1$ pontos:

D1) Aleatórios uniformemente distribuídos, obtidos pelo comando `rand(N, 1)` do *Matlab*.

D2) Aleatórios uniformemente distribuídos, mas obrigando que $z_1 = 0$ e $z_N = 1$.

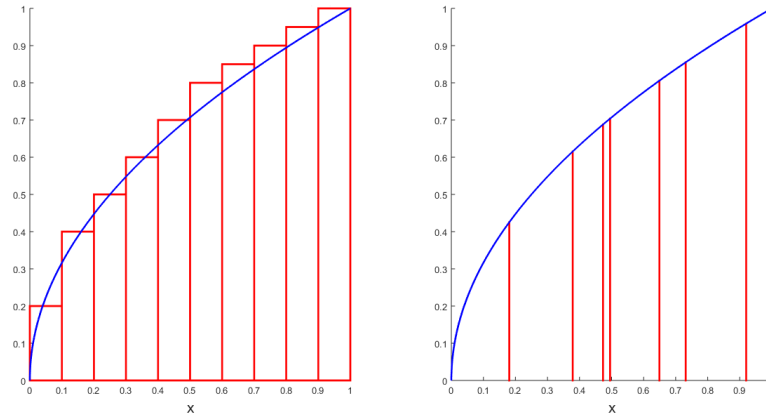
D3) Uniformemente distribuídos com intervalo $\frac{1}{N-1}$, ou seja:

$$z_i = \frac{i-1}{N-1}, \quad \forall i = 1, 2, \dots, N.$$

Podemos comparar o método de integração Monte Carlo dado pelo Algoritmo 3.1 com o método de integração por retângulos. No primeiro caso temos que os valores de x_i são escolhidos de forma aleatória a partir de uma distribuição uniforme em $[a, b]$. Já no segundo caso temos que os pontos x_i são tomados espaçados com intervalos escolhidos de mesma largura $\frac{b-a}{N}$, ou seja, escolhidos de forma regular.

Na Figura 3.1 temos a comparação entre o método de integração por retângulos e por Monte Carlo. A imagem da esquerda representa o método de integração por retângulos, enquanto à direita representada o método de integração de Monte Carlo.

Rodamos o Algoritmo 3.1 m vezes utilizando as discretizações D1, D2 e D3 para diferentes valores de N . A discretização D3 é rodada uma única vez visto que não é aleatória, enquanto as discretizações D1 e D2 foram rodadas $m = 100$ para cada valor de N considerado.



Fonte: O autor

Figura 3.1: Comparação entre o método de integração por retângulos e por Monte Carlo

A Tabela 3.1 mostra a média da integral (3.8), do erro cometido (3.9) e do desvio padrão (3.10) pelo método de Monte Carlo para diferentes valores de N usados nas discretizações do intervalo $[0, 1]$.

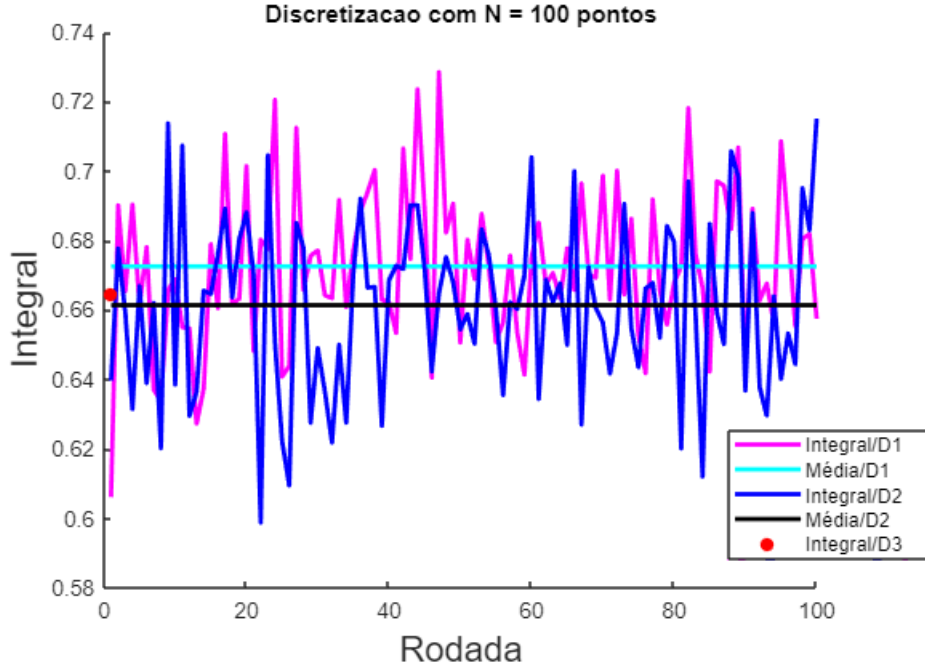
Discretização	N	Integral	Erro	σ_m
D1	100	0.6725	1.5218×10^{-16}	0.0222
D2		0.6614	1.2450×10^{-16}	0.0247
D3		0.6648	1.9080×10^{-16}	0
D1	1000	0.6672	4.7363×10^{-16}	0.0070
D2		0.6672	5.0625×10^{-16}	0.0072
D3		0.6665	5.8516×10^{-16}	0
D1	10000	0.6670	1.2354×10^{-15}	0.0022
D2		0.6666	1.4791×10^{-15}	0.0023
D3		0.6666	2.0920×10^{-15}	0
D1	100000	0.6667	5.0842×10^{-16}	7.0636×10^{-4}
D2		0.6667	5.3445×10^{-16}	7.3964×10^{-4}
D3		0.6667	6.3210×10^{-16}	0

Tabela 3.1: Resultados obtidos com o Algoritmo 3.1 para $m = 100$ rodadas.

Observe que quanto maior for o número N de pontos uniformemente distribuídos sobre o intervalo de integração, menor será a estimativa para o erro assim como o desvio padrão.

Nas Figuras 3.2, 3.3, 3.4 e 3.5 podemos observar a média e o valor da integral utilizando

as discretizações D1, D2 e D3 para diferentes valores de N , sendo eles 100, 1000, 10000 e 100000 respectivamente.



Fonte: O autor

Figura 3.2: Comparação entre a média e o valor da integral para $N=100$.

3.1.2 Caso multidimensional

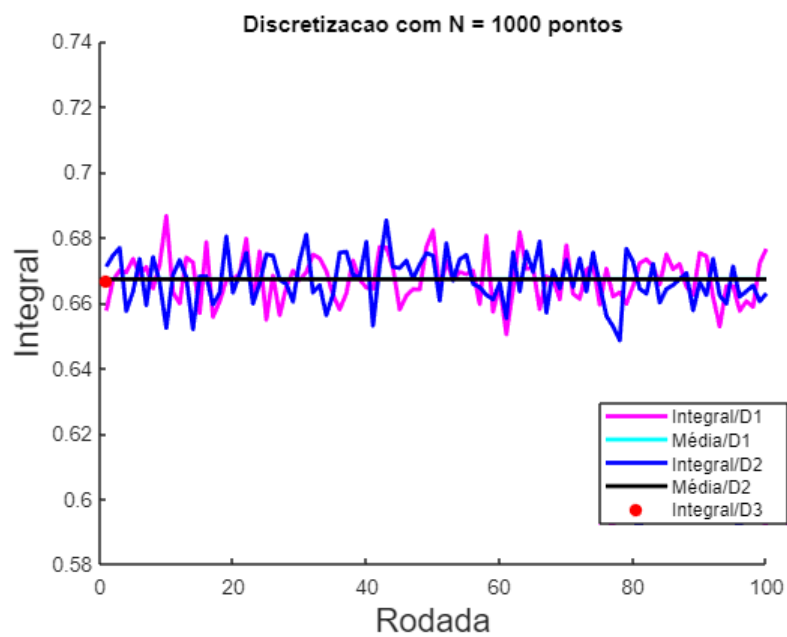
Vimos anteriormente como calcular a integral de uma função de uma única variável definida num intervalo $[a, b]$. Generalizaremos agora para funções com mais de uma variável. Consideremos então f uma função de n variáveis, onde queremos calcular:

$$I = \int_{a_n}^{b_n} \int_{a_{n-1}}^{b_{n-1}} \dots \int_{a_1}^{b_1} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n. \quad (3.11)$$

Note que, se o intervalo de integração de cada integral não for o intervalo $[0, 1]$, podemos aplicar a substituição de variáveis dada em (3.2) em cada integral e obter:

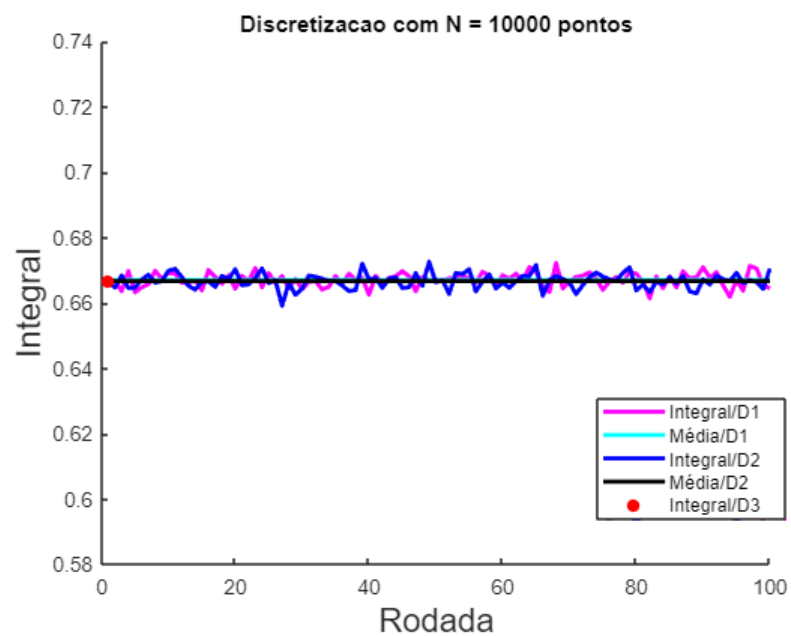
$$I = \int_0^1 \int_0^1 \dots \int_0^1 g(z_1, z_2, \dots, z_n) dz_1 dz_2 \dots dz_n. \quad (3.12)$$

Para calcular I é preciso gerar n conjuntos independentes, um para cada variável. Cada um



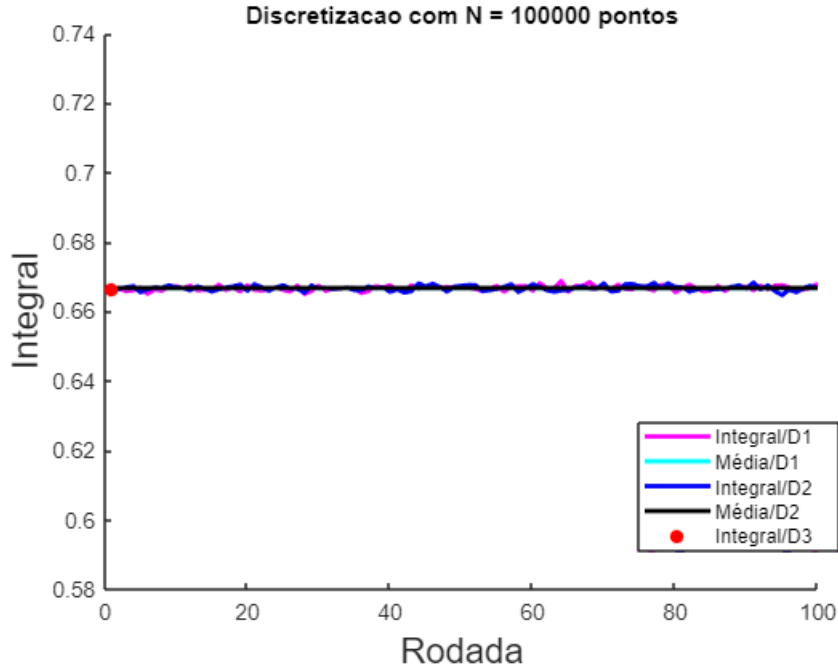
Fonte: O autor

Figura 3.3: Comparação entre a média e o valor da integral para $N=1000$.



Fonte: O autor

Figura 3.4: Comparação entre a média e o valor da integral para $N=10000$.



Fonte: O autor

Figura 3.5: Comparação entre a média e o valor da integral para $N=100000$.

desses conjuntos consistindo de N pontos uniformemente distribuídos sobre o intervalo $[0, 1]$:

$$\begin{array}{c} z_1^1, z_2^1, \dots, z_n^1 \\ z_1^2, z_2^2, \dots, z_n^2 \\ \vdots \\ z_1^N, z_2^N, \dots, z_n^N. \end{array}$$

Assim, as variáveis $g(z_1^i, z_2^i, \dots, z_n^i)$, $i = 1, 2, \dots, N$ são todas uniformemente distribuídas, pode-se aproximar I gerando uma grande quantidade de conjuntos de números aleatórios $z_1^i, z_2^i, \dots, z_n^i$ e tomar como aproximação o valor médio de $g(z_1^i, z_2^i, \dots, z_n^i)$, ou seja

$$I \approx \frac{1}{N} \sum_{i=1}^N g(z_1^i, z_2^i, \dots, z_n^i). \quad (3.13)$$

A precisão da estimativa I é o erro padrão da média $\sigma_P(I)$, calculado:

$$\sigma_P(I) = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (g(z_1^i, z_2^i, \dots, z_n^i) - I)^2}. \quad (3.14)$$

Por simplicidade consideremos que para todas as variáveis a discretização tem o mesmo número de pontos.

Algoritmo 3.2. Método de Monte Carlo - Multidimensional

Passo 0: Dado: $N \in \mathbb{N}$ o número de pontos da discretização.

Passo 1: Discretize o cuboide $[0, 1]^n$ uniformemente.

Passo 2: Faça a mudança de variáveis

$$x_i = a_i + (b_i - a_i)z_i, \quad \forall i = 1, \dots, n.$$

Passo 3: Assim a integral é dada por:

$$I = \int_0^1 \int_0^1 \dots \int_0^1 g(z_1, z_2, \dots, z_n) dz_1 dz_2 \dots dz_n \approx \frac{1}{N} \sum_{i=1}^N g(z_1^i, z_2^i, \dots, z_n^i).$$

Passo 4: Calcule o erro cometido

$$\sigma_P(I) = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (g(z_1^i, z_2^i, \dots, z_n^i) - I)^2}.$$

Semelhante ao caso unidimensional, caso as discretizações do cuboide $[0, 1]^n$ sejam de forma randômica deve-se rodar o algoritmo m vezes e calcular a média das rodadas, o erro médio e o desvio padrão.

Exemplo 3.2. Considere o problema de calcular a integral (3.11) em que $f(x_1, x_2) = \frac{x_1}{x_2} + \frac{x_2}{x_1}$ e $[a_1, b_1] \times [a_2, b_2] = [1, 2] \times [1, 4]$. O valor analítico desta integral é $21/2 \cdot \ln(2)$. Vamos calculá-la usando o método de Monte Carlo descrito no Algoritmo 3.2 com os três tipos de discretizações do cuboide $[0, 1]^2$, em $N > 1$ pontos, análogas às utilizadas no Exemplo 3.1.

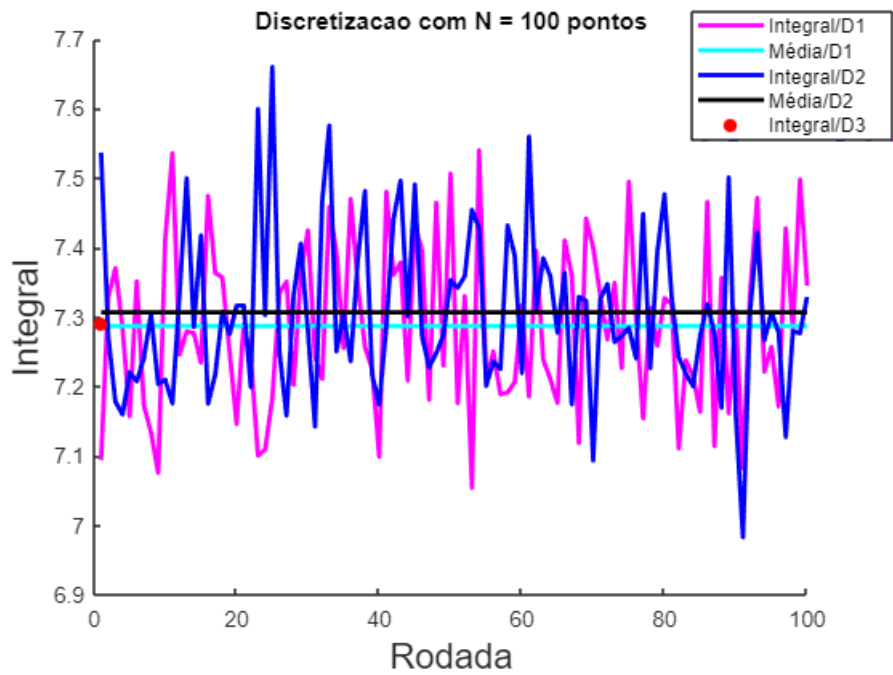
Rodamos o Algoritmo 3.2 m vezes utilizando as discretizações D1, D2 e D3 para diferentes valores de N . A discretização D3 é rodada uma única vez visto que não é aleatória, enquanto as discretizações D1 e D2 foram rodadas $m = 100$ para cada valor de N considerado.

A Tabela 3.2 mostra a média da integral, do erro cometido e do desvio padrão pelo método de Monte Carlo para diferentes valores de N usados nas discretizações do intervalo $[0, 1]$ obtidos através do Algoritmo 3.2.

Nas Figuras 3.6, 3.7, 3.8 e 3.9 podemos observar a média e o valor da integral utilizando as discretizações D1, D2 e D3 para diferentes valores de N , sendo eles 100, 1000, 2500 e 5000 respectivamente.

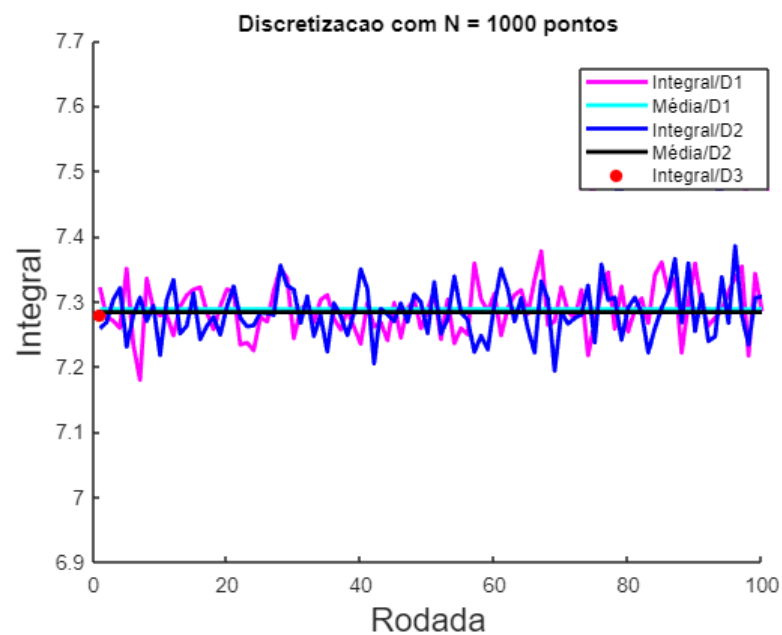
Discretização	N	Integral	Erro	σ_m
D1	100	7.2870	0.0488	0.1181
D2		7.3066	0.0489	0.1161
D3		7.2898	0.0488	0
D1	1000	7.2883	0.0049	0.0376
D2		7.2835	0.0049	0.0375
D3		7.2792	0.0049	0
D1	2500	7.2781	0.0019	0.0243
D2		7.2842	0.0020	0.0248
D3		7.2785	0.0019	0
D1	5000	7.2787	9.7456×10^{-4}	0.0181
D2		7.2793	9.7465×10^{-4}	0.0172
D3		7.2783	9.7451×10^{-4}	0

Tabela 3.2: Resultados obtidos com o Algoritmo 3.2 para $m = 100$ rodadas.



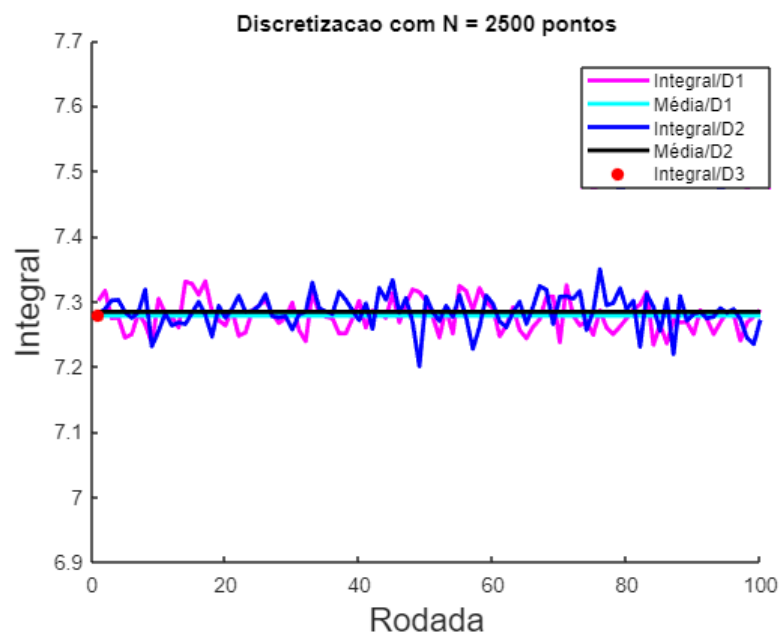
Fonte: O autor

Figura 3.6: Comparação entre a média e o valor da integral para $N=100$.



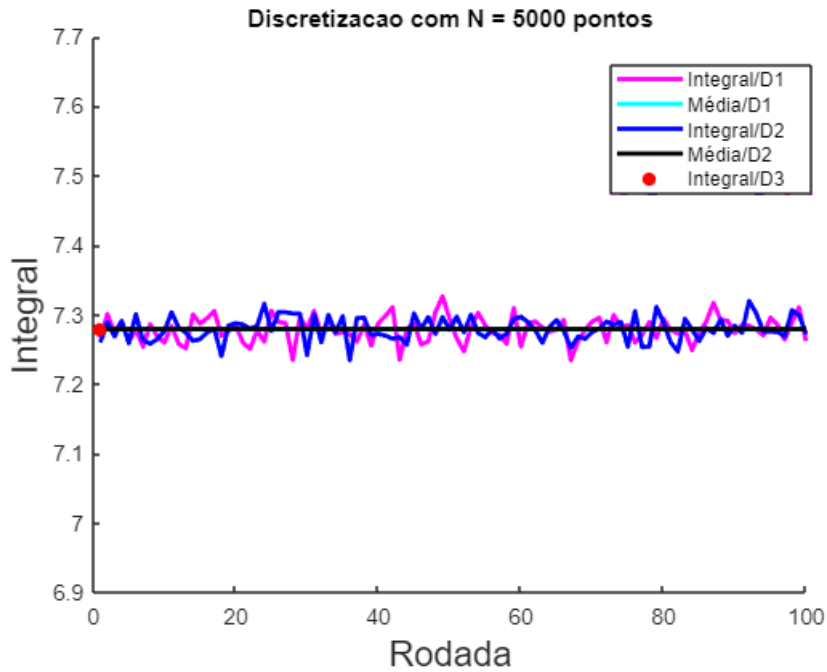
Fonte: O autor

Figura 3.7: Comparação entre a média e o valor da integral para N=1000.



Fonte: O autor

Figura 3.8: Comparação entre a média e o valor da integral para N=2500.



Fonte: O autor

Figura 3.9: Comparação entre a média e o valor da integral para $N=5000$.

3.2 Problema Mono objetivo

Nesta seção abordaremos os casos em que o problema é irrestrito e que o problema é restrito.

3.2.1 Problema irrestrito

Inicialmente, consideramos o problema irrestrito

$$\begin{aligned} &\text{Minimizar } f(x) \\ &x \in \mathbb{R}^n. \end{aligned} \tag{P'}$$

Veremos o método do valor médio do conjunto de nível, este será o primeiro algoritmo a ser tratado para encontrar minimizadores globais, para isso suponhamos que f é uma função contínua e que existe um número real c tal que o conjunto de nível

$$H_c = \{x \in \mathbb{R}^n \mid f(x) \leq c\}, \tag{3.15}$$

é não vazio e compacto.

Algoritmo

Tome um número real c_0 tal que o conjunto de nível H_{c_0} é não vazio. Por exemplo, pode-se tomar $c_0 \geq \bar{c} = \min_{x \in \mathbb{R}^n} f(x)$ ou $x_0 \in \mathbb{R}^n$ e definir $c_0 = f(x_0)$. Caso a medida de $\mu(H_{c_0})$ for igual a zero, então c_0 é o valor mínimo global e H_{c_0} é o conjunto de minimizadores globais de f . Caso contrário, calcule o valor médio $M(f, c_0)$ de f em seu conjunto de nível H_{c_0} , e defina

$$c_1 = M(f, c_0) = \frac{1}{\mu(H_{c_0})} \int_{H_{c_0}} f(x) d\mu.$$

É claro que

$$c_0 \geq c_1 \geq \bar{c} = \min_{x \in \mathbb{R}^n} f(x).$$

Assim, gera-se uma sequência não crescente do valor médio de f definida, para $k \geq 0$ por

$$c_{k+1} = M(f, c_k) = \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu$$

em que $H_{c_{k+1}} = \{x \in \mathbb{R}^n \mid f(x) \leq c_{k+1}\}$.

Dessa maneira,

$$c_0 \geq c_1 \geq \dots \geq c_k \geq c_{k+1} \geq \dots \geq \bar{c}$$

e pela definição de conjuntos de níveis

$$H_{c_0} \supset H_{c_1} \supset \dots \supset H_{c_k} \supset H_{c_{k+1}} \supset \dots \supset H_{\bar{c}}.$$

Desta pequena discussão sobre como obter tais sequências e aplicando as condições de otimalidade global obtidas nos Teoremas 2.15 e 2.25, temos o seguinte algoritmo.

Algoritmo 3.3. Método do Valor Médio dos conjuntos de nível

Passo 0 - Considere um ponto $x_0 \in \mathbb{R}^n$ e $\varepsilon > 0$. Seja

$$c_0 = f(x_0) \quad e \quad H_{c_0} = \{x \in \mathbb{R}^n \mid f(x) \leq c_0\}.$$

Defina $k := 0$.

Passo 1 - Calcule o valor médio

$$c_{k+1} = M(f, c_k) = \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} f(x) d\mu$$

e seja

$$H_{c_{k+1}} = \{x \in \mathbb{R}^n \mid f(x) \leq c_{k+1}\}.$$

Passo 2 - Calcule a variância modificada

$$VF = V_1(f, c_k) = \frac{1}{\mu(H_{c_k})} \int_{H_{c_k}} [f(x) - c_k]^2 d\mu.$$

Passo 3 - Se $VF \geq \varepsilon$, seja $k := k + 1$ e volte para o passo 1, caso contrário vá para o passo 4.

Passo 4 - Defina $\bar{c} = c_{k+1}$ e $\bar{H} = H_{c_{k+1}}$. PARE.

O número $\bar{c}$ e o conjunto $\bar{H}$ aproximam o valor mínimo global e o conjunto de minimizadores globais, respectivamente. O passo 3 aplica a condição de otimalidade global 2.25 da variância modificada $V_1(f, \bar{c}) = 0$ em uma forma de aproximação, pois se considerar $\varepsilon = 0$ no Algoritmo 3.3, o processo iterativo não termina. Obtemos duas sequências $\{c_k\}_{k \in \mathbb{N}}$ e $\{H_{c_k}\}_{k \in \mathbb{N}}$ as quais são ambas limitadas inferiormente. Portanto, são convergentes. Seja

$$\bar{c} = \lim_{k \rightarrow \infty} c_k \quad \text{e} \quad \bar{H} = \lim_{k \rightarrow \infty} H_{c_k}.$$

Assim, temos o seguinte teorema de convergência.

Teorema 3.3 (Teo. 1.1, [23]). *Considere as sequências $\{c_k\}$ e $\{H_{c_k}\}$ geradas pelo Algoritmo 3.3. O limite $\bar{c} = \lim_{k \rightarrow \infty} c_k$ é o valor mínimo global e o limite $\bar{H} = \lim_{k \rightarrow \infty} H_{c_k}$ é o conjunto de minimizadores globais de f .*

Demonstração. Pela construção do algoritmo, temos

$$c_{k+1} = M(f, c_k), \quad k = 0, 1, \dots$$

já que $\{c_{k+1}\}_{k \in \mathbb{N}}$ e $\{c_k\}_{k \in \mathbb{N}}$ são sequências não crescentes e tem o mesmo limite $\bar{c}$.

Temos $\lim_{k \rightarrow \infty} M(f, c_k) = M(f, \bar{c})$. Agora, aplicando o Teorema 2.15, concluímos que $\bar{c}$ é o valor mínimo global de f em $\mathbb{R}^n$. Além disso, temos também que

$$\lim_{k \rightarrow \infty} H_{c_k} = H_{\bar{c}} = \bar{H} = \{x \in \mathbb{R}^n \mid f(x) = \bar{c}\}.$$

■

3.2.2 Problema restrito

Consideremos agora o problema restrito:

$$\begin{aligned} &\text{Minimizar } f(x) \\ &x \in X \end{aligned} \tag{P}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ e X é um subconjunto não vazio de $\mathbb{R}^n$.

Um algoritmo de rejeição para encontrar o valor mínimo global e o conjunto de minimizadores globais de f sobre o conjunto de restrições X é descrito no Algoritmo 3.4.

Algoritmo 3.4. Algoritmo de rejeição

Passo 0 - Considere um ponto $x_0 \in X$ e $\varepsilon > 0$. Seja

$$c_0 = f(x_0) \quad \text{e} \quad H_{c_0} = \{x \in X \mid f(x) \leq c_0\}.$$

Defina $k := 0$.

Passo 1 - Calcule o valor médio

$$c_{k+1} = M(f, c_k, X) = \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} f(x) d\mu$$

e seja

$$H_{c_{k+1}} = \{x \in X \mid f(x) \leq c_{k+1}\}.$$

Passo 2 - Calcule a variância modificada

$$VF = V_1(f, c_k, X) = \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} [f(x) - c_k]^2 d\mu.$$

Passo 3 - Se $VF \geq \varepsilon$, seja $k := k + 1$ e volte para o passo 1, caso contrário vai para o passo 4.

Passo 4 - Defina $\hat{c} = c_{k+1}$ e $\hat{H} = H_{c_{k+1}}$. PARE.

O número $\hat{c}$ é uma aproximação do valor mínimo global e o conjunto $\hat{H}$ é a aproximação correspondente do conjunto de minimizadores globais de f em X . Novamente, no passo 3, a condição de otimalidade global $V_1(f, c, X) = 0$ é aplicada em uma forma de aproximação, pois de definirmos $\varepsilon = 0$ neste algoritmo, então o procedimento iterativo não será encerrado e obteremos duas sequências decrescentes, $\{c_k\}$ e $\{H_{c_k}\}$, as quais são limitadas inferiormente. Portanto, são convergentes. Seja $\bar{c} = \lim_{k \rightarrow \infty} c_k$ e $\bar{H} = \lim_{k \rightarrow \infty} (H_{c_k} \cap X)$.

Utilizando a condição de otimalidade do valor médio de rejeição, pode ser provado facilmente o seguinte teorema de convergência.

Teorema 3.4 (Teo. 2.1, [23]). *O limite $\bar{c} = \lim_{k \rightarrow \infty} c_k$ é o valor mínimo global e $\bar{H} = \lim_{k \rightarrow \infty} (H_{c_k} \cap X)$ é o conjunto de minimizadores globais de f em X .*

Demonstração. Pela construção do algoritmo, temos

$$c_{k+1} = M(f, c_k, X), \quad k = 0, 1, \dots$$

já que $\{c_{k+1}\}_{k \in \mathbb{N}}$ e $\{c_k\}_{k \in \mathbb{N}}$ são sequências decrescentes e tem o mesmo limite $\bar{c}$.

Temos $\lim_{k \rightarrow \infty} M(f, c_k, X) = M(f, \bar{c}, X)$. Agora, aplicando o Teorema 2.36, concluimos que $\bar{c}$ é o valor mínimo global de f em X . Além disso, temos também que

$$\lim_{k \rightarrow \infty} H_{c_k} \cap X = H_{\bar{c}} \cap X = \bar{H} = \{x \in X \mid f(x) = \bar{c}\}.$$

■

A seguir, veremos como utilizar o método de Monte Carlo nas implementações das condições de otimalidade e dos algoritmos teóricos vistos na Subseção 3.2.1.

Começamos com um modelo simples de implementação. Suponhamos que $X = D$ seja um cuboide em $\mathbb{R}^n$ dado por $D = \{x \in \mathbb{R}^n \mid a \leq x \leq b\}$ e f é uma função com um único minimizador global

$\bar{x} \in D$. Em outras palavras, para uma sequência decrescente $\{c_k\}$ a qual converge para o valor mínimo $\bar{c}$, a distância de pontos do conjunto de nível satisfaz

$$\rho_k = \rho(H_{c_k}) = \max_{x, y \in H_{c_k}} \|x - y\| \rightarrow 0 \text{ quando } k \rightarrow \infty,$$

onde $H_{c_k} = \{x \in \mathbb{R}^n \mid f(x) \leq c_k\}$. Temos

$$\bar{c} = \min_{x \in D} f(x) = \min_{x \in H_{c_k} \cap D} f(x) = \min_{x \in D_k} f(x)$$

onde $D_k := \{x \in \mathbb{R}^n \mid a_k \leq x \leq b_k\}$ é o menor cuboide o qual contém o conjunto de nível $H_{c_k} \cap D$.

Lema 3.5. *Considere a sequência $\{c_k\}$ decrescente e $D_k := \{x \in \mathbb{R}^n \mid a_k \leq x \leq b_k\}$, onde*

$$\begin{aligned} a_k^i &= \min\{x^i \mid (x^1, x^2, \dots, x^i, \dots, x^n) \in H_{c_k}\}, \\ b_k^i &= \max\{x^i \mid (x^1, x^2, \dots, x^i, \dots, x^n) \in H_{c_k}\}, \end{aligned}$$

como sendo menor cuboide o qual contém o conjunto de nível $H_{c_k} \cap D$. Então $\{\bar{x}\} = \bigcap_{k=1}^{\infty} D_k$, onde $\bar{x}$ é o único minimizador global de f em D_k .

Demonstração. É claro que $\bar{x} \in \bigcap_{k=1}^{\infty} [H_{c_k} \cap D] \subset \bigcap_{k=1}^{\infty} D_k$. Suponha que $\{\bar{x}\} \neq \bigcap_{k=1}^{\infty} D_k$, isto é, existe $x_1 \in \bigcap_{k=1}^{\infty} D_k$ tal que $x_1 \neq \bar{x}$ e $f(x_1) = d > \bar{c}$, onde $\bar{c}$ é o valor mínimo de f . Seja D_0 o menor cuboide em $\mathbb{R}^n$ contendo $\bar{x}$ e x_1 .

Então $D_0 \subset D_k$, $k = 1, 2, \dots$ com $b_0 - a_0 > 0$. Agora, existe $x_{\alpha_i}, x_{\beta_i} \in H_{c_k}$ tal que $a_k = x_{\alpha_i}$ e $b_k = x_{\beta_i}$. Como

$$\begin{aligned} \|x_{\alpha_i} - x_{\beta_i}\| &= \left[(x_1^{\alpha_i} - x_1^{\beta_i})^2 + (x_2^{\alpha_i} - x_2^{\beta_i})^2 + \dots + (x_n^{\alpha_i} - x_n^{\beta_i})^2 \right]^{1/2} \\ &\geq \left[(x_i^{\alpha_i} - x_i^{\beta_i})^2 \right]^{1/2} = b_i^k - a_i^k, \end{aligned}$$

temos, para $i = 1, \dots, n$,

$$b_k^i - a_k^i \leq \|x_{\alpha_i} - x_{\beta_i}\| \leq \rho(H_{c_k}) \rightarrow 0, \text{ quando } k \rightarrow \infty,$$

o que é uma contradição, pois $b_0 - a_0 > 0$. ■

A seguir temos um algoritmo para este modelo simples, semelhante ao dado na Subseção 3.2.1.

Algoritmo 3.5. Modelo Simples

Passo 0 - Considere $c_0 > \bar{c} = \min f(x)$ e $\varepsilon > 0$. Seja $D_0 = D$ um cuboide inicial.

Defina $k := 0$.

Passo 1 - Calcule o valor médio

$$c_{k+1} = M(f, c_k, D_k) = \frac{1}{\mu(D_k)} \int_{D_k} f(x) d\mu.$$

Seja D_{k+1} o menor cuboide contendo o conjunto de nível

$$H_{c_{k+1}} = \{x \in D_k \mid f(x) \leq c_{k+1}\}.$$

Passo 2 - Calcule a variância modificada

$$VF = V_1(f, c_k, D_k) = \frac{1}{\mu(D_k)} \int_{D_k} (f(x) - c_k)^2 d\mu.$$

Passo 3 - Se $VF \geq \varepsilon$, seja $k := k + 1$ e volte para o passo 1, caso contrário vá para o passo 4.

Passo 4 - Defina $\bar{c} = c_{k+1}$ e $\{\bar{x}\} \subset D_{k+1}$. PARE.

Note que ao invés de calcular as integrais sobre H_{c_k} em cada iteração calculamos sobre D_k , o que é claramente mais fácil. Também, calcular $M(f, c_k, D_k)$ e $V_1(f, c_k, D_k)$ ao invés de $M(f, c_k)$ e $V_1(f, c_k)$ nos dá maior precisão pois $\rho(D_k) \rightarrow 0$ quando $k \rightarrow \infty$.

Discutiremos a seguir alguns detalhes do Algoritmo 3.5 com a utilização do método de Monte Carlo, visto na Seção 3.1, para o cálculo das integrais e aproximação do conjunto de nível.

1. Aproximando H_{c_0} e $M(f, c_0, D)$

Considere N o número de pontos aleatórios da discretização do intervalo $[0, 1]$. Seja $\xi = (\xi_1^j, \xi_2^j, \dots, \xi_n^j)$, para $j = 1, \dots, N$, um conjunto de números aleatórios uniformemente distribuído sobre $[0, 1]^n$. Portanto, cada ξ_i^j para $i = 1, \dots, n$ e $j = 1, \dots, N$, é um número aleatório no intervalo $[0, 1]$. Considere a mudança de variáveis

$$x = a + (b - a)\xi.$$

Então $x = (x_1, \dots, x_n)$ é uniformemente distribuído sobre D .

Consideramos os t menores valores da função f avaliada nos pontos aleatórios e os ordenamos em ordem decrescente, denotando por

$$FV[1] \geq FV[2] \geq \dots \geq FV[t].$$

Considere o conjunto $\mathcal{W}$ dos argumentos desses valores funcionais, ou seja, o conjunto dos pontos da amostra correspondentes aos t menores valores da função.

O conjunto $\mathcal{W}$ é chamado de conjunto de aceitação o qual pode ser considerado como uma aproximação para o conjunto de nível H_{c_0} , onde $c_0 = FV[1]$ é o maior valor de $\{FV[\ell]\}$, $\ell \in \{1, \dots, t\}$. O inteiro positivo t é chamado de índice estatístico. É claro que $f(x) \leq c_0$ para todo $x \in \mathcal{W}$. Também, o valor médio de f sobre o conjunto de nível H_{c_0} pode ser aproximado pelo valor médio de $\{FV[\ell]\}$:

$$c_1 = M(f, c_0, D) \approx \frac{FV[1] + FV[2] + \dots + FV[t]}{t}.$$

2. Geração do cuboide D_1

O cuboide de dimensão n :

$$D_1 = \{x \in \mathbb{R}^n \mid a_1 \leq x \leq b_1\}$$

pode ser gerado estatisticamente pelo seguinte procedimento. Suponha que as amostras aleatórias no conjunto de aceitação $\mathcal{W}$ são $\tau_1, \tau_2, \dots, \tau_t \in \mathbb{R}^n$. Considere, para $i = 1, \dots, n$,

$$\sigma_{min}^i = \min\{\tau_1^i, \tau_2^i, \dots, \tau_n^i\} \quad \text{e} \quad \sigma_{max}^i = \max\{\tau_1^i, \tau_2^i, \dots, \tau_n^i\},$$

onde $\tau_j = (\tau_1^j, \tau_2^j, \dots, \tau_n^j)$, para todo $j = 1, 2, \dots, t$. Define-se então

$$a_1 = \sigma_{min} - \frac{\sigma_{max} - \sigma_{min}}{t - 1} \quad \text{e} \quad b_1 = \sigma_{max} + \frac{\sigma_{max} - \sigma_{min}}{t - 1}.$$

3. Continuando o processo iterativo

As amostras agora são tomadas no novo domínio D_1 . Considere $\xi = (\xi_1^j, \xi_2^j, \dots, \xi_n^j)$, para $j = 1, \dots, N$, um conjunto de números aleatórios uniformemente distribuído sobre $[0, 1]^n$. Tome um ponto da amostra aleatório $x = (x_1, x_2, \dots, x_n)$ em D_1 , da forma

$$x = a_1 + (b_1 - a_1)\xi.$$

Avalie $f(x)$. Se $f(x) \geq FV[1]$, então abandone este ponto da amostra, caso contrário, atualize os conjuntos $\{FV[\ell]\}$ e $\mathcal{W}$ tal que o novo conjunto $\{FV[\ell]\}$ seja composto pelos t melhores valores da função obtidos até o momento. O conjunto de aceitação $\mathcal{W}$ é atualizado de acordo. Repetindo este processo até que $FV[1] \leq c_1$, obtemos novos $\{FV[\ell]\}$ e $\mathcal{W}$.

Continuando este processo temos uma sequência decrescente de valores médios $\{c_k\}$ e uma sequência de cuboides $\{D_k\}$ de dimensão n .

4. Solução iterativa

Em cada iteração, o menor valor $FV[t]$ no conjunto $\{FV[\ell]\}$ e o ponto correspondente em $\mathcal{W}$ podem ser considerados como uma solução iterativa para o problema de minimização sobre D_k .

5. Critério de convergência

A variância modificada VF de $\{FV[\ell]\}$, a qual é dada por

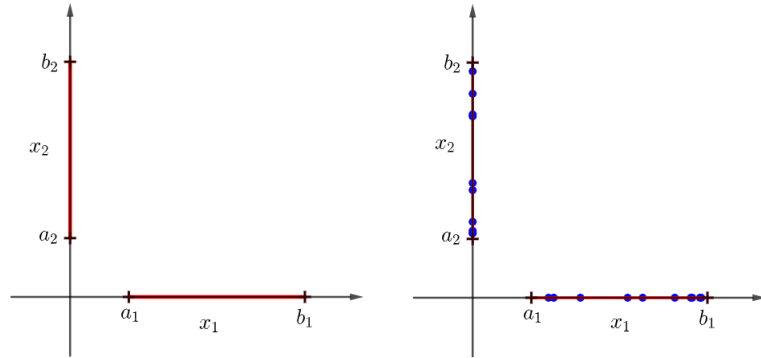
$$VF = \frac{1}{t - 1} \sum_{\ell=1}^t (FV[\ell] - FV[1])^2,$$

pode ser considerada uma aproximação de $V_1(f, c_k, D_k)$ em cada iteração. Se VF é menor que a precisão ε dada, então o processo iterativo termina, e a iteração atual no passo 4 do Algoritmo 3.5 pode servir como uma estimativa do valor mínimo global e do minimizador global.

A seguir temos um exemplo simples de um cuboide no $\mathbb{R}^2$ para ilustrar os detalhes discutidos anteriormente.

Consideremos primeiramente o cuboide $D = \{x \in \mathbb{R}^2 \mid a \leq x \leq b\}$, ou seja, temos que as variáveis são limitadas inferiormente e superiormente, $a_1 \leq x_1 \leq b_1$ e $a_2 \leq x_2 \leq b_2$. Cada variável é discretizada em N pontos aleatórios uniformemente distribuídos sobre o cuboide. Por simplicidade consideramos a mesma quantidade de pontos na discretização de cada variável.

Na Figura 3.10 à esquerda temos representado o cuboide no $\mathbb{R}^2$ enquanto à direita temos a representação da discretização de cada variável em N pontos.

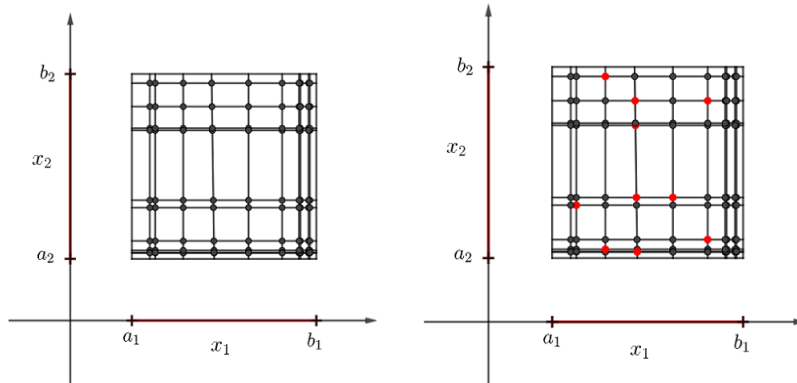


Fonte: O autor

Figura 3.10: Cuboide inicial e discretização.

Obtemos a malha formada pelos pontos da discretização. Avaliamos a função objetivo em cada ponto desta malha e ordenamos os seus valores. Seleccionamos os t melhores valores da função avaliada, ou seja, consideramos uma amostra menor contendo os t menores valores da função avaliada.

Na Figura 3.11 à esquerda temos representado a malha formada pelos N^2 pontos enquanto à direita temos a representação da amostra dos t pontos.

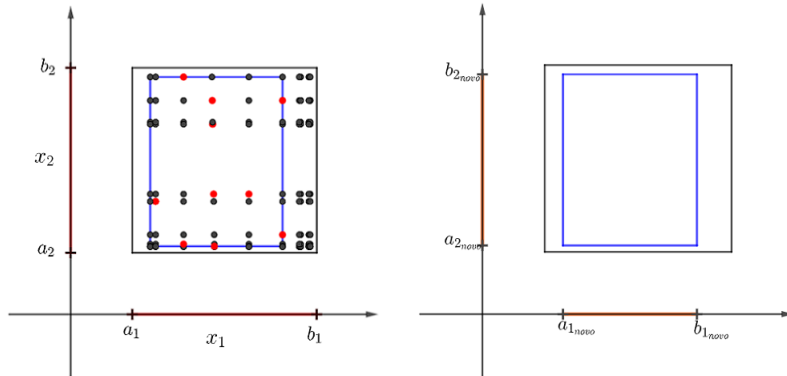


Fonte: O autor

Figura 3.11: Malha e amostra.

Obtemos o novo cuboide, que contém os t pontos da amostra correspondentes aos t menores valores da função avaliada. Repetimos o processo iterativamente.

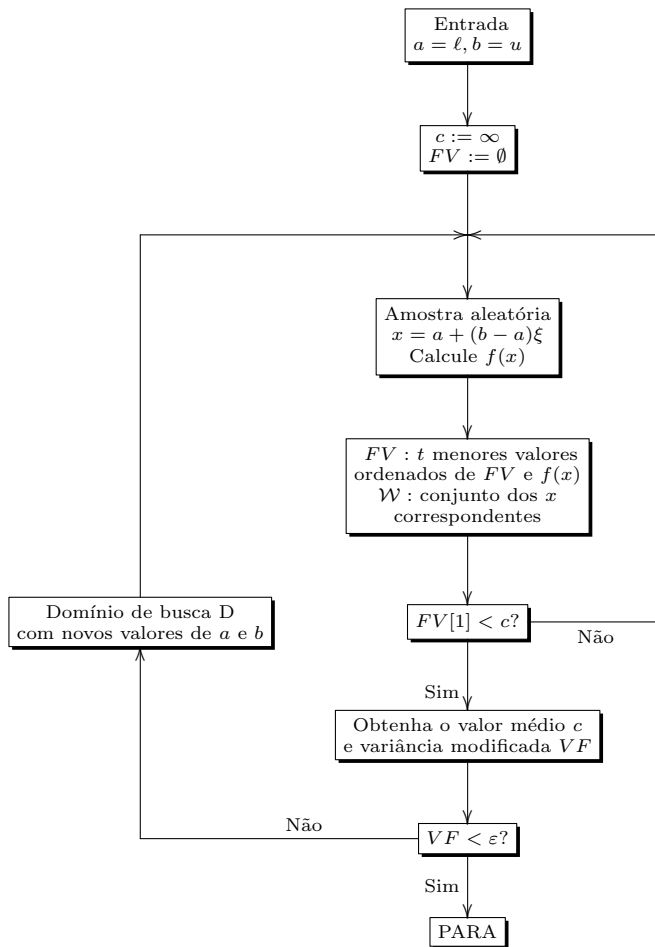
Na Figura 3.12 à esquerda temos representado o menor cuboide que contém os t melhores pontos enquanto à direita temos a representação do novo cuboide.



Fonte: O autor

Figura 3.12: Cuboide que contém os t menores valores de f .

Fluxograma do Algoritmo 3.5



3.2.3 Testes Numéricos

Nesta subseção apresentamos experimentos numéricos, os quais foram realizados em um notebook Samsung Intel Celeron, CPU 1007U @ 1.50Ghz com 4GB RAM, 64 bits, usando Matlab 2020a e foram considerados 19 problemas os quais podem ser encontrados em [23, 35, 56].

Consideremos valor inicial de $c_0 = 10^{12}$, $\varepsilon = 10^{-20}$ e o cuboide $D_0 = D$ fornecido em cada problema. O Algoritmo 3.5 foi rodado 10 vezes para cada um dos 19 problemas. A média dos resultados bem como o desvio padrão referente às 10 rodadas foram calculadas. A escolha do número N de discretizações e o número t de menores valores da função foi variada entre as instâncias dos problemas.

Tais resultados serão descritos em forma de tabelas, onde cada coluna será correspondente a uma informação do problema. A primeira coluna corresponde ao número da rodada, a segunda ao minimizador $\bar{x}$ encontrado, a terceira coluna ao valor de $f(\bar{x})$, ou seja, ao valor mínimo correspondente a $\bar{x}$, a quarta ao número de iterações necessárias na rodada e, por fim, a quinta coluna ao valor da variância modificada. Começamos com um problema “simples” para certificar-se que o algoritmo está funcionando corretamente.

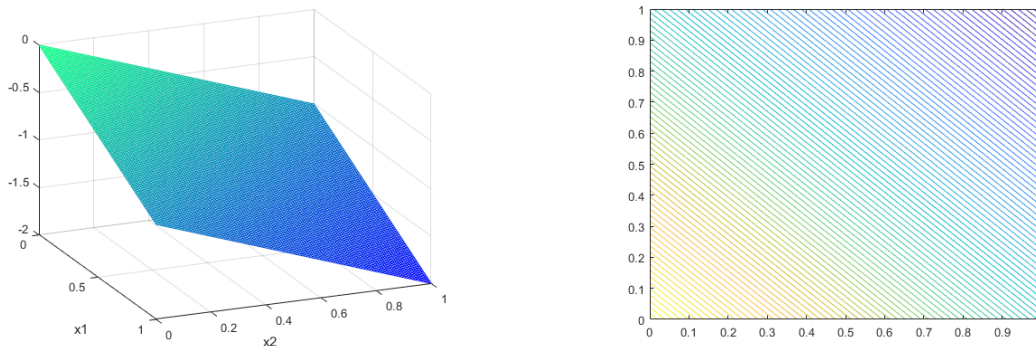
Problema 3.6. Dado $n \geq 2$ considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por:

$$f(x) = -\sum_{i=1}^n x_i,$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid 0 \leq x \leq 1\}$.

Sabemos que a solução deste problema é $\bar{x} = (1, \dots, 1)$ com valor ótimo correspondente igual a $f(\bar{x}) = -n$.

A Figura 3.13 ilustra o Problema 3.6 para $n = 2$. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.13: Problema 3.6

Nas Tabelas 3.3, 3.4, 3.5 e 3.6 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com diferentes valores de n . Para tais resultados foi considerado n igual a 2, 3, 4 e 5, respectivamente.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(9.9804, 9.9964) \cdot e - 01$	-1.9977	9	$8.5864e - 23$
2	$(9.9976, 9.9972) \cdot e - 01$	-1.9995	9	$3.8428e - 22$
3	$(9.9858, 9.9993) \cdot e - 01$	-1.9985	9	$2.7382e - 21$
4	$(9.9951, 9.9887) \cdot e - 01$	-1.9984	8	$7.3046e - 21$
5	$(9.9982, 9.9972) \cdot e - 01$	-1.9995	9	$5.0620e - 23$
6	$(9.9968, 9.9992) \cdot e - 01$	-1.9996	9	$7.4671e - 22$
7	$(9.9999, 9.9824) \cdot e - 01$	-1.9982	8	$2.7654e - 22$
8	$(9.9892, 9.9995) \cdot e - 01$	-1.9989	9	$1.9612e - 22$
9	$(9.9694, 9.9957) \cdot e - 01$	-1.9965	9	$5.3161e - 23$
10	$(9.9983, 9.9965) \cdot e - 01$	-1.9995	9	$4.6737e - 22$
Média	$(0.9991, 0.9995)$	-1.9986	8.8	$1.2303e - 21$

Tabela 3.3: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 2$, $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(9.5306, 9.6911, 9.8675) \cdot e - 01$	-2.9089	13	$5.0200e - 22$
2	$(9.9620, 9.9892, 9.9476) \cdot e - 01$	-2.9899	9	$6.7936e - 22$
3	$(9.8986, 9.9067, 9.9862) \cdot e - 01$	-2.9792	9	$2.3019e - 22$
4	$(9.9850, 9.8393, 9.9415) \cdot e - 01$	-2.9766	10	$4.3750e - 22$
5	$(9.7643, 9.9813, 9.8825) \cdot e - 01$	-2.9628	13	$1.8589e - 21$
6	$(9.8304, 10.003, 9.4757) \cdot e - 01$	-2.9309	10	$5.2112e - 22$
7	$(9.9328, 9.9714, 9.9110) \cdot e - 01$	-2.9815	10	$7.9051e - 22$
8	$(9.9204, 9.9509, 9.9533) \cdot e - 01$	-2.9825	11	$7.9971e - 21$
9	$(9.9896, 9.9728, 9.9554) \cdot e - 01$	-2.9918	9	$5.0790e - 21$
10	$(9.8759, 9.6789, 9.9611) \cdot e - 01$	-2.9516	10	$5.2270e - 22$
Média	$(0.9869, 0.9898, 0.9888)$	-2.9656	10.4	$1.8618e - 21$

Tabela 3.4: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 3$, $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0274.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(9.9458, 9.7592, 9.8593, 9.8712) \cdot e - 01$	-3.9436	7	$3.1347e - 22$
2	$(9.7820, 9.6580, 9.9858, 9.9954) \cdot e - 01$	-3.9421	10	$1.2696e - 21$
3	$(9.9378, 9.8020, 10.002, 9.7753) \cdot e - 01$	-3.9517	9	$8.1682e - 23$
4	$(9.9221, 9.8797, 9.9950, 10.001) \cdot e - 01$	-3.9797	8	$1.4076e - 21$
5	$(9.9663, 9.9143, 9.9488, 9.9182) \cdot e - 01$	-3.9748	8	$3.2716e - 22$
6	$(9.9986, 9.9593, 10.003, 9.8279) \cdot e - 01$	-3.9789	8	$2.4319e - 22$
7	$(9.7493, 9.7588, 9.9357, 9.9992) \cdot e - 01$	-3.9443	7	$1.7550e - 22$
8	$(9.9693, 9.9038, 9.9644, 9.9084) \cdot e - 01$	-3.9746	7	$7.7356e - 21$
9	$(9.9111, 9.8844, 9.8195, 9.9995) \cdot e - 01$	-3.9615	9	$1.0355e - 22$
10	$(9.9929, 9.8956, 9.9799, 9.9822) \cdot e - 01$	-3.9851	8	$2.5453e - 23$
Média	$(0.9918, 0.9842, 0.9949, 0.9928)$	-3.9636	8.1	$1.1683e - 21$

Tabela 3.5: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 4$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0169.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(9.5070, 9.9170, 9.8827, 9.9983, 9.9118) \cdot e - 01$	-4.9217	11	$1.1750e - 21$
2	$(9.9038, 9.9334, 9.9499, 9.6175, 9.8857) \cdot e - 01$	-4.9290	10	$3.1341e - 22$
3	$(9.5493, 9.9727, 9.9390, 9.9732, 9.7137) \cdot e - 01$	-4.9148	10	$7.0055e - 22$
4	$(9.9119, 9.8987, 9.6014, 9.8796, 9.8392) \cdot e - 01$	-4.9131	11	$9.5431e - 23$
5	$(9.9708, 9.9262, 9.8349, 9.6657, 9.8895) \cdot e - 01$	-4.9287	11	$2.6240e - 21$
6	$(9.9491, 9.7029, 9.8127, 9.8834, 9.6907) \cdot e - 01$	-4.9039	11	$4.1510e - 22$
7	$(9.9873, 9.8206, 9.9372, 9.6674, 9.9173) \cdot e - 01$	-4.9330	9	$2.5292e - 21$
8	$(9.9642, 9.7447, 9.9407, 9.9275, 9.9533) \cdot e - 01$	-4.9530	7	$4.5456e - 21$
9	$(9.6156, 9.9898, 9.8828, 9.9632, 9.9408) \cdot e - 01$	-4.9392	10	$2.4538e - 21$
10	$(9.6415, 10.001, 9.6957, 9.9608, 9.9031) \cdot e - 01$	-4.9202	10	$2.2929e - 22$
Média	$(0.9800, 0.9891, 0.9848, 0.9854, 0.9865)$	-4.9257	10	$1.5081e - 21$

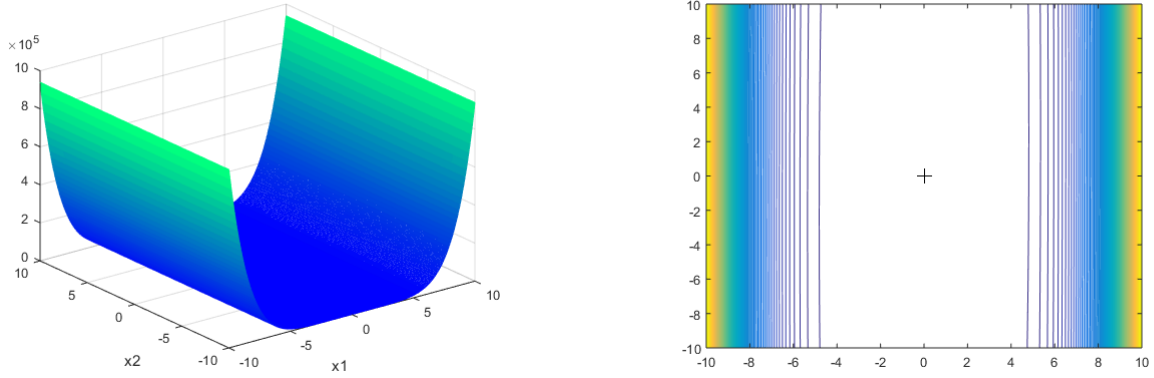
Tabela 3.6: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.6 com $n = 5$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0141.

Problema 3.7 (Problema A.2, [56]). *Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

$$f(x_1, x_2) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1),$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid -10 \leq x_1 \leq 10; -10 \leq x_2 \leq 10\}$.

A Figura 3.14 ilustra o Problema 3.7. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.14: Problema 3.7

Na Tabela 3.7 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.7.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-3.8245e-08, -8.6517e-08)$	$2.3229e-12$	4	$5.3956e-24$
2	$(-7.4874e-08, -2.1146e-07)$	$2.9427e-12$	4	$6.3078e-24$
3	$(-1.2665e-08, -5.4797e-08)$	$1.2264e-12$	4	$1.6920e-24$
4	$(7.3315e-08, 6.0238e-08)$	$2.3863e-12$	4	$5.1601e-24$
5	$(-3.1832e-08, -8.3902e-08)$	$1.9175e-12$	4	$5.9011e-24$
6	$(-9.3128e-08, -7.2461e-08)$	$3.8996e-12$	4	$1.2561e-23$
7	$(-1.7670e-09, -1.1044e-08)$	$1.6387e-12$	4	$4.2473e-24$
8	$(4.5640e-08, 2.8234e-08)$	$8.3487e-13$	4	$1.1891e-24$
9	$(-4.6835e-08, 1.6147e-07)$	$3.3974e-12$	4	$1.0681e-23$
10	$(1.1810e-08, 1.0787e-08)$	$8.5895e-13$	4	$8.9897e-25$
Média	$(-0.1686, -0.2595) \cdot e-07$	$2.1425e-12$	4	$5.4034e-24$

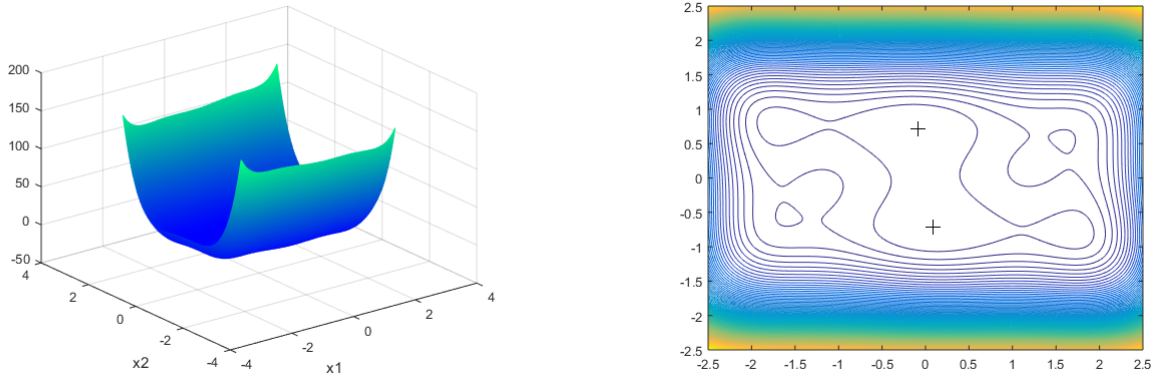
Tabela 3.7: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.7 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.0486e-12$.

Problema 3.8 (Problema A.3, [56]). Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4,$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid -2.5 \leq x_1 \leq 2.5; -2.5 \leq x_2 \leq 2.5\}$.

A Figura 3.15 ilustra o Problema 3.8. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.15: Problema 3.8

Na Tabela 3.8 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.8.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-9.0227e-02, 7.1565e-01)$	$-1.0316e+00$	5	$6.0243e-21$
2	$(8.9842e-02, -7.1266e-01)$	$-1.0316e+00$	5	$9.7392e-25$
3	$(8.9843e-02, -7.1266e-01)$	$-1.0316e+00$	4	$4.4470e-21$
4	$(8.9842e-02, -7.1266e-01)$	$-1.0316e+00$	8	$2.1530e-24$
5	$(-8.9843e-02, 7.1266e-01)$	$-1.0316e+00$	5	$1.5369e-23$
6	$(-9.0121e-02, 7.1484e-01)$	$-1.0316e+00$	5	$5.2813e-24$
7	$(9.2247e-02, -7.1280e-01)$	$-1.0316e+00$	5	$8.9794e-23$
8	$(8.9798e-02, -7.1231e-01)$	$-1.0316e+00$	14	$3.6688e-24$
9	$(8.9841e-02, -7.1266e-01)$	$-1.0316e+00$	5	$1.4212e-22$
10	$(8.9842e-02, -7.1266e-01)$	$-1.0316e+00$	10	$1.1927e-24$
Média		-1.0316	6.6	$1.0732e-21$

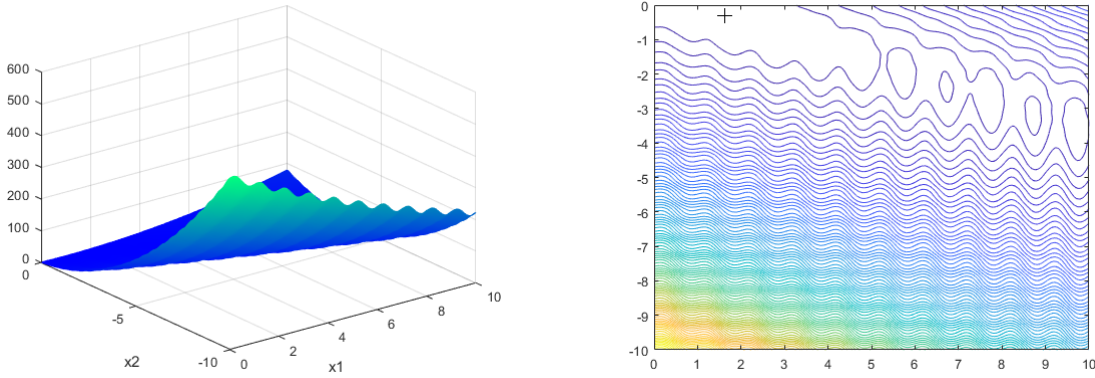
Tabela 3.8: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.8 com $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.3406e-16$.

Problema 3.9 (Problema A.4, [56]). *Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

$$f(x_1, x_2) = (1 - 2x_2 + 0.05\sin 4\pi x_2 - x_1)^2 + (x_2 - 0.5\sin(2\pi x_1))^2,$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 10; -10 \leq x_2 \leq 0\}$.

A Figura 3.16 ilustra o Problema 3.9. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.16: Problema 3.9

Na Tabela 3.9 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.9.

Problema 3.10 (Problema A.5, [56]). *Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

$$f(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10,$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid -5 \leq x_1 \leq 10; 0 \leq x_2 \leq 15\}$.

A Figura 3.17 ilustra o Problema 3.10. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.

Na Tabela 3.10 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.10.

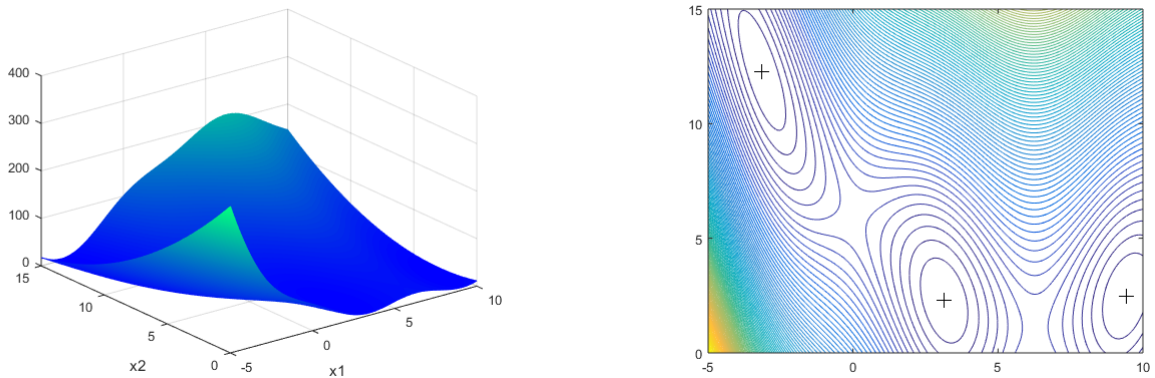
Problema 3.11 (Exemplo 5.1, [23]). *Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

$$f(x_1, x_2) = 100(x_2 - x_1^2) + (1 - x_1)^2,$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid -5 \leq x_1 \leq 5; -5 \leq x_2 \leq 5\}$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(1.8512e + 00, -4.0209e - 01)$	$2.6549e - 08$	5	$0.0000e + 00$
2	$(1.8513e + 00, -4.0209e - 01)$	$9.9755e - 13$	5	$3.7246e - 25$
3	$(1.8513e + 00, -4.0210e - 01)$	$1.1952e - 09$	6	$3.5896e - 22$
4	$(1.8513e + 00, -4.0228e - 01)$	$1.6403e - 07$	4	$1.7786e - 22$
5	$(1.8513e + 00, -4.0209e - 01)$	$2.4103e - 11$	4	$8.9068e - 22$
6	$(1.8513e + 00, -4.0209e - 01)$	$3.1062e - 11$	4	$1.8448e - 21$
7	$(1.8512e + 00, -4.0212e - 01)$	$3.7298e - 08$	5	$1.5277e - 23$
8	$(1.8513e + 00, -4.0206e - 01)$	$3.4213e - 09$	4	$2.9762e - 23$
9	$(1.8488e + 00, -4.0224e - 01)$	$2.6655e - 05$	3	$0.0000e + 00$
10	$(1.8572e + 00, -4.0191e - 01)$	$1.6318e - 04$	4	$0.0000e + 00$
Média	$(1.8516, -0.4021)$	$1.9007e - 05$	4.4	$3.3177e - 22$

Tabela 3.9: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.9 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $5.1344e - 05$.



Fonte: O autor

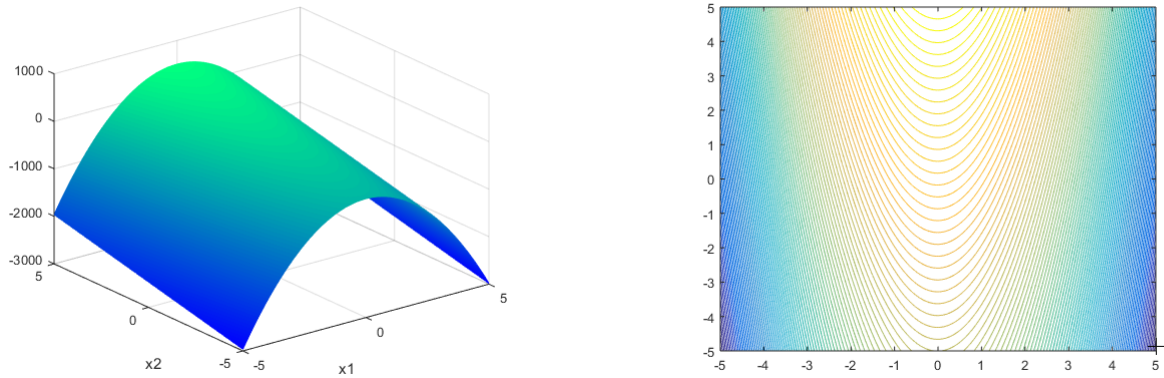
Figura 3.17: Problema 3.10

A Figura 3.18 ilustra o Problema 3.11. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.

Na Tabela 3.11 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.11.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(3.1413, 2.2767)	0.3979	16	$1.4677e - 27$
2	(9.4240, 2.4743)	0.3979	14	$3.2774e - 21$
3	(-3.1436, 12.2836)	0.3979	12	$4.9270e - 24$
4	(-3.1408, 12.2731)	0.3979	12	$1.4001e - 22$
5	(3.1417, 2.2749)	0.3979	76	$2.8365e - 21$
6	(9.4257, 2.4760)	0.3979	15	$0.0000e + 00$
7	(3.1414, 2.2779)	0.3979	16	$0.0000e + 00$
8	(3.1410, 2.2755)	0.3979	544	$4.3406e - 22$
9	(-3.1433, 12.2729)	0.3979	11	$7.3049e - 23$
10	(3.1416, 2.2750)	0.3979	15	$2.4928e - 22$
Média		0.3979	73.1	$7.0152e - 22$

Tabela 3.10: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.10 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$.



Fonte: O autor

Figura 3.18: Problema 3.11

Problema 3.12. Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

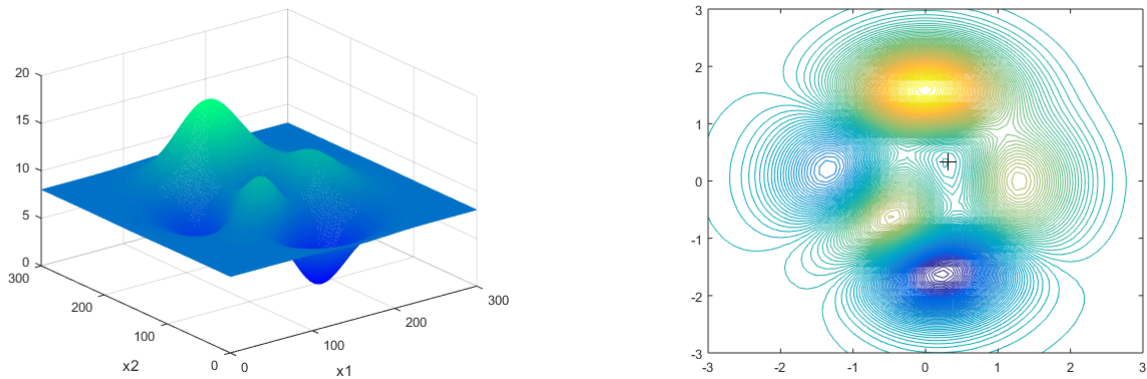
$$f(x_1, x_2) = 3(1 - x_1)^2 e^{-x_1^2 - (x_2+1)^2} - 10 \left(\frac{x_1}{5} - x_1^3 - x_2^5 \right) e^{-x_1^2 - x_2^2} - \frac{1}{3} e^{-(x_1+1)^2 - x_2^2},$$

e o domínio de busca $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 60; 0 \leq x_2 \leq 60\}$.

A Figura 3.19 ilustra o Problema 3.12. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(4.9881e + 00, -4.7029e + 00)$	$-2.9425e + 03$	21	$3.7454e - 21$
2	$(4.9346e + 00, -5.0330e + 00)$	$-2.9229e + 03$	18	$4.2359e - 22$
3	$(5.0560e + 00, -4.5887e + 00)$	$-2.9988e + 03$	20	$7.1171e - 21$
4	$(5.0268e + 00, -4.8892e + 00)$	$-2.9996e + 03$	21	$3.7804e - 22$
5	$(4.8231e + 00, -4.9427e + 00)$	$-2.8059e + 03$	16	$2.2241e - 22$
6	$(4.9302e + 00, -4.9426e + 00)$	$-2.9095e + 03$	26	$1.5571e - 21$
7	$(4.9405e + 00, -5.0258e + 00)$	$-2.9279e + 03$	15	$6.5402e - 22$
8	$(4.9918e + 00, -4.9612e + 00)$	$-2.9720e + 03$	18	$1.6716e - 22$
9	$(4.9410e + 00, -4.9722e + 00)$	$-2.9231e + 03$	19	$5.0023e - 21$
10	$(4.9623e + 00, -5.0051e + 00)$	$-2.9473e + 03$	20	$4.9480e - 21$
Média	$(4.9594, -4.9063)$	$-2.9349e + 03$	19.4	$2.4215e - 21$

Tabela 3.11: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.11 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 55.2526.



Fonte: O autor

Figura 3.19: Problema 3.12

Na Tabela 3.12 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.12.

Problema 3.13 (Exemplo 5.5, [23]). Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por

$$f(x_1, x_2) = 2(x_1^2 + x_2^2) - \lfloor x_1^2 + x_2^2 \rfloor,$$

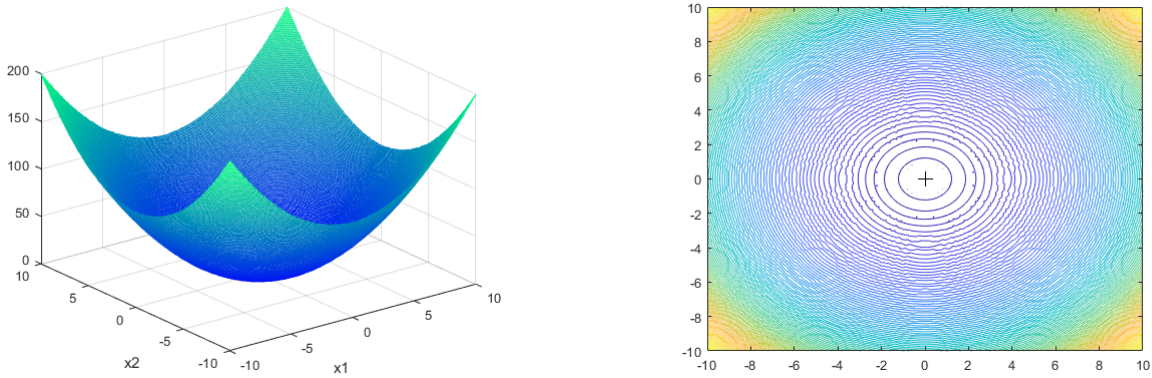
com $\lfloor z \rfloor$ sendo a parte inteira de z , e o domínio de busca

$$D = \{(x_1, x_2) \in \mathbb{R}^2 \mid -10 \leq x_1 \leq 10; -10 \leq x_2 \leq 10\}.$$

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	69	$7.0100e-21$
2	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	84	$9.7268e-21$
3	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	8	$2.7483e-25$
4	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	10	$3.8099e-27$
5	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	24	$6.3641e-25$
6	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	24	$2.7976e-26$
7	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	12	$1.3852e-25$
8	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	13	$4.6917e-25$
9	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	5	$1.6004e-25$
10	$(2.9645e-01, 3.2020e-01)$	$-6.4936e-02$	35	$3.9886e-25$
Média	$(0.2965, 0.3202)$	-0.0649	28.4	$1.6739e-21$

Tabela 3.12: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.12 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$.

A Figura 3.20 ilustra o Problema 3.13. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.20: Problema 3.13

Na Tabela 3.13 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.13.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-1.7657e-08, 7.3326e-08)$	$8.1261e-13$	4	$8.6351e-25$
2	$(4.7303e-08, -3.8427e-08)$	$3.7067e-13$	4	$1.7070e-25$
3	$(-2.0603e-08, 6.4068e-08)$	$2.6179e-12$	4	$9.5642e-24$
4	$(5.9956e-08, -1.5168e-08)$	$3.9009e-13$	4	$2.0387e-25$
5	$(9.0578e-08, 1.8388e-08)$	$9.2985e-13$	4	$1.1341e-24$
6	$(1.6687e-08, 5.0944e-09)$	$8.3703e-14$	4	$1.1138e-26$
7	$(6.0162e-08, 2.8935e-08)$	$1.9477e-13$	4	$4.1954e-26$
8	$(9.5187e-08, -2.8929e-08)$	$3.2783e-13$	4	$1.3377e-25$
9	$(-1.8487e-09, -2.2644e-08)$	$9.2949e-14$	4	$1.2025e-26$
10	$(3.3956e-08, 1.7310e-08)$	$1.1222e-13$	4	$1.8859e-26$
Média	$(0.3637e-07, 0.1020e-07)$	$5.9326e-13$	4	$1.2154e-24$

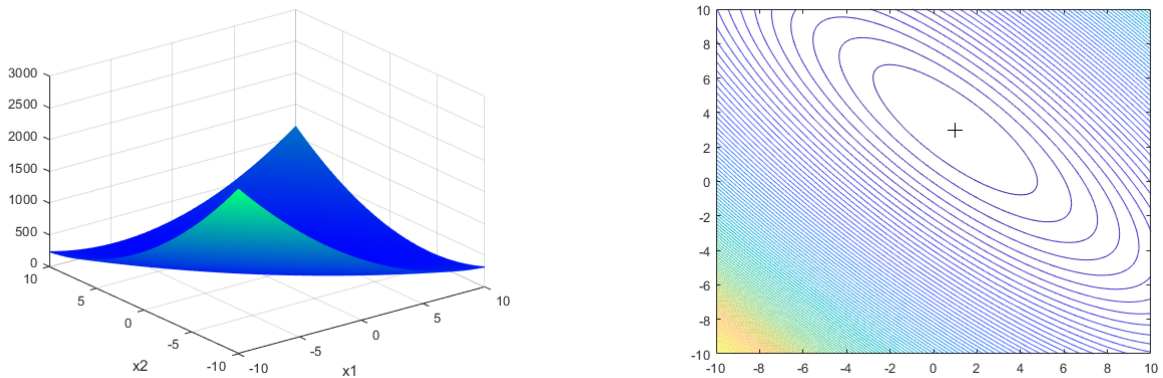
Tabela 3.13: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.13 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $7.6898e-13$.

Problema 3.14 (Appendix Booth, [35]). *Considere a função objetivo $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por*

$$f(x_1, x_2) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2,$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -10 \leq x \leq 10\}$.

A Figura 3.21 ilustra o Problema 3.14. A imagem da esquerda representa o gráfico da função objetivo, enquanto à direita são representadas algumas de suas curvas de nível.



Fonte: O autor

Figura 3.21: Problema 3.14

Na Tabela 3.14 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.14.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(1.0000e + 00, 3.0000e + 00)	6.5044e - 11	9	3.7247e - 21
2	(1.0000e + 00, 3.0000e + 00)	1.2770e - 11	9	2.1179e - 22
3	(1.0000e + 00, 3.0000e + 00)	5.8511e - 11	9	4.4424e - 21
4	(1.0000e + 00, 3.0000e + 00)	3.5046e - 12	8	3.5361e - 23
5	(1.0000e + 00, 3.0000e + 00)	1.7660e - 11	10	2.4052e - 22
6	(1.0000e + 00, 3.0000e + 00)	3.5572e - 11	9	1.5488e - 21
7	(1.0000e + 00, 3.0000e + 00)	1.3225e - 11	8	3.6029e - 22
8	(1.0000e + 00, 3.0000e + 00)	1.3242e - 11	9	1.9782e - 22
9	(1.0000e + 00, 3.0000e + 00)	1.5013e - 11	9	2.5859e - 22
10	(1.0000e + 00, 3.0000e + 00)	1.5833e - 11	10	2.6480e - 22
Média	(1, 3)	2.5037e - 11	9	1.1285e - 21

Tabela 3.14: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.14 com $N = 1000$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.0988e - 11$.

Problema 3.15 (Problema A.7, [56]). *Dados $m, n \geq 3$ e considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por*

$$f(x) = - \sum_{i=1}^m v_i \exp \left(- \sum_{j=1}^n A_{ij} (x_j - P_{ij})^2 \right),$$

onde os parâmetros $v \in \mathbb{R}^m$, A e $P \in \mathbb{R}^{m \times n}$ são dados nas Tabelas 3.15 e 3.17 para diferentes valores de n e m . Foi considerado o domínio de busca $D = \{x \in \mathbb{R}^n \mid 0 \leq x \leq 1\}$.

i	A_{ij}			v_i	P_{ij}		
1	3.0	10.0	30.0	1.0	0.3689	0.1170	0.2673
2	0.1	10.0	35.0	1.2	0.4699	0.4387	0.7470
3	3.0	10.0	30.0	3.0	0.1091	0.8732	0.5547
4	0.1	10.0	35.0	3.2	0.03815	0.5743	0.8828

Tabela 3.15: Parâmetros para o Problema 3.15 com $m = 4$ e $n = 3$.

Na Tabela 3.16 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.15.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(1.1519, 5.4303, 8.5469) \cdot e - 01$	$-3.8566e + 00$	2	$0.0000e + 00$
2	$(1.1550, 5.5546, 8.5170) \cdot e - 01$	$-3.8627e + 00$	2	$0.0000e + 00$
3	$(1.1646, 5.4810, 8.5227) \cdot e - 01$	$-3.8608e + 00$	2	$0.0000e + 00$
4	$(1.1781, 5.5502, 8.4483) \cdot e - 01$	$-3.8571e + 00$	2	$0.0000e + 00$
5	$(1.1430, 5.5402, 8.5278) \cdot e - 01$	$-3.8627e + 00$	2	$0.0000e + 00$
6	$(1.1623, 5.5631, 8.4972) \cdot e - 01$	$-3.8620e + 00$	2	$0.0000e + 00$
7	$(1.1442, 5.5943, 8.5192) \cdot e - 01$	$-3.8622e + 00$	3	$0.0000e + 00$
8	$(1.1345, 5.6598, 8.5224) \cdot e - 01$	$-3.8590e + 00$	3	$0.0000e + 00$
9	$(1.1455, 5.5565, 8.5255) \cdot e - 01$	$-3.8628e + 00$	4	$8.3295e - 25$
10	$(1.2077, 5.5297, 8.3559) \cdot e - 01$	$-3.8357e + 00$	2	$0.0000e + 00$
Média	$(0.1159, 0.5546, 0.8498)$	-3.8582	2.4	$8.3295e - 26$

Tabela 3.16: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.15, $N = 100$ e $t = 20$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0082.

i	A_{ij}						v_i	P_{ij}					
1	10.0	3.0	17.0	3.5	1.7	8.0	1.0	0.1312	0.1696	0.5569	0.0124	0.8283	0.5886
2	0.05	10.0	17.0	0.1	8.0	14.0	1.2	0.2329	0.4135	0.8307	0.3736	0.1004	0.9991
3	3.0	3.5	1.7	10.0	17.0	8.0	3.0	0.2348	0.1451	0.3522	0.2883	0.3047	0.6650
4	17.0	8.0	0.05	10.0	0.1	14.0	3.2	0.4047	0.8828	0.8732	0.5743	0.1091	0.0381

Tabela 3.17: Parâmetros para o Problema 3.15 com $m = 4$ e $n = 6$.

Na Tabela 3.18 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.17.

Problema 3.16 (Problema A.6, [56]). Considere $D = \{(x_1, \dots, x_4) \in \mathbb{R}^4 \mid 0 \leq x \leq 10\}$ e a função objetivo $f : D \rightarrow \mathbb{R}$ dada por

$$f(x) = - \sum_{i=1}^m \left(\frac{1}{(x - A_i)^T (x - A_i) + v_i} \right),$$

onde os parâmetros A_i e v_i são dados na Tabela 3.19.

Na Tabela 3.20 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19 e $m = 5$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(0.2926, 0.1491, 0.4637, 0.2814, 0.3101, 0.6522)	-3.2253	11	$1.2846e - 23$
2	(0.2009, 0.1512, 0.5090, 0.2904, 0.3132, 0.6576)	-3.3043	8	$0.0000e + 00$
3	(0.2093, 0.1498, 0.4705, 0.2765, 0.2949, 0.6573)	-3.3077	7	$2.8725e - 23$
4	(0.2025, 0.1499, 0.4735, 0.2758, 0.2990, 0.6576)	-3.3145	7	$0.0000e + 00$
5	(0.2033, 0.1506, 0.4743, 0.2762, 0.2636, 0.6582)	-3.2093	5	$0.0000e + 00$
6	(0.2061, 0.1335, 0.4722, 0.3019, 0.3112, 0.6576)	-3.2983	3	$0.0000e + 00$
7	(0.1998, 0.1503, 0.4814, 0.2522, 0.3167, 0.6567)	-3.3057	6	$3.6628e - 22$
8	(0.2018, 0.1497, 0.4764, 0.2752, 0.3042, 0.6431)	-3.3143	5	$0.0000e + 00$
9	(0.2092, 0.1512, 0.4840, 0.2755, 0.3111, 0.6574)	-3.3211	4	$0.0000e + 00$
10	(0.2021, 0.1428, 0.4725, 0.2753, 0.3113, 0.6578)	-3.3216	4	$0.0000e + 00$
Média	(0.2128, 0.1478, 0.4778, 0.2781, 0.3035, 0.6556)	-3.2922	6	$4.0785e - 23$

Tabela 3.18: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.15 com os parâmetros dispostos na Tabela 3.17, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0403.

i	A_i				v_i
1	4.0	4.0	4.0	4.0	0.1
2	1.0	1.0	1.0	1.0	0.2
3	8.0	8.0	8.0	8.0	0.2
4	6.0	6.0	6.0	6.0	0.4
5	3.0	7.0	3.0	7.0	0.4
6	2.0	9.0	2.0	9.0	0.6
7	5.0	5.0	3.0	3.0	0.3
8	8.0	1.0	8.0	1.0	0.7
9	6.0	2.0	6.0	2.0	0.5
10	7.0	3.6	7.0	3.6	0.5

Tabela 3.19: Parâmetros para o Problema 3.16

Na Tabela 3.21 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19 e $m = 7$.

Na Tabela 3.22 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19 e $m = 10$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(4.0000, 4.0046, 4.0001, 4.0001)	$-1.0151e + 01$	8	$0.0000e + 00$
2	(3.9892, 4.0001, 4.0000, 4.0001)	$-1.0142e + 01$	11	$3.4307e - 24$
3	(4.0000, 4.0001, 3.9994, 4.0001)	$-1.0153e + 01$	8	$6.0767e - 21$
4	(3.9788, 4.0001, 4.0006, 4.0001)	$-1.0108e + 01$	9	$3.8551e - 22$
5	(3.9999, 4.0001, 4.0000, 4.0001)	$-1.0153e + 01$	8	$2.5686e - 21$
6	(4.0000, 4.0001, 4.0000, 3.9995)	$-1.0153e + 01$	8	$0.0000e + 00$
7	(4.0001, 3.9958, 4.0007, 4.0001)	$-1.0151e + 01$	5	$0.0000e + 00$
8	(4.0000, 4.0001, 4.0000, 4.0001)	$-1.0153e + 01$	23	$2.7474e - 21$
9	(4.0065, 4.0001, 4.0168, 4.0001)	$-1.0121e + 01$	8	$0.0000e + 00$
10	(4.0000, 4.0001, 4.0000, 4.0001)	$-1.0153e + 01$	9	$1.1507e - 22$
Média	(3.9974, 4.0001, 4.0018, 4.0000)	-10.1438	9.7	$1.1897e - 21$

Tabela 3.20: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 5$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0161.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(4.0006, 4.0006, 3.9995, 4.0006)	$-1.0403e + 01$	6	$2.6456e - 25$
2	(4.0006, 4.0007, 3.9999, 3.9998)	$-1.0403e + 01$	9	$9.4877e - 22$
3	(4.0006, 4.0005, 3.9995, 3.9996)	$-1.0403e + 01$	7	$4.4765e - 21$
4	(4.0006, 4.0007, 3.9995, 3.9995)	$-1.0403e + 01$	7	$1.3024e - 25$
5	(4.0051, 4.0007, 3.9995, 3.9996)	$-1.0401e + 01$	7	$0.0000e + 00$
6	(4.0009, 4.0007, 3.9975, 3.9996)	$-1.0403e + 01$	6	$0.0000e + 00$
7	(4.0001, 4.0007, 3.9995, 4.0039)	$-1.0401e + 01$	8	$5.1978e - 23$
8	(4.0006, 4.0007, 3.9995, 3.9996)	$-1.0403e + 01$	7	$4.5246e - 22$
9	(4.0006, 4.0007, 3.9995, 3.9996)	$-1.0403e + 01$	7	$1.0059e - 22$
10	(4.0021, 4.0020, 3.9978, 3.9996)	$-1.0402e + 01$	11	$1.4192e - 21$
Média	(4.0012, 4.0008, 3.9992, 4.0001)	-10.4025	7.5	$7.4499e - 22$

Tabela 3.21: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 7$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $8.4984e - 04$.

Problema 3.17 (Problema A.8, [56]). *Considere $D = \{(x_1, \dots, x_4) \in \mathbb{R}^4 \mid -0.3 \leq x \leq 0.3\}$ e a função*

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(4.0007, 4.0006, 4.0049, 3.9995)	$-1.0534e + 01$	8	$9.0499e - 24$
2	(4.0007, 4.0006, 3.9997, 3.9995)	$-1.0536e + 01$	7	$1.3764e - 21$
3	(4.0007, 3.9970, 3.9997, 3.9995)	$-1.0535e + 01$	7	$1.5184e - 21$
4	(4.0007, 4.0006, 3.9997, 3.9995)	$-1.0536e + 01$	7	$5.4574e - 21$
5	(4.0083, 4.0006, 3.9994, 4.0007)	$-1.0531e + 01$	10	$0.0000e + 00$
6	(4.0005, 4.0006, 3.9997, 3.9995)	$-1.0536e + 01$	7	$7.0287e - 22$
7	(4.0008, 4.0003, 4.0089, 3.9740)	$-1.0463e + 01$	5	$0.0000e + 00$
8	(4.0009, 3.9955, 3.9997, 3.9995)	$-1.0534e + 01$	9	$0.0000e + 00$
9	(4.0007, 4.0006, 3.9992, 3.9995)	$-1.0536e + 01$	8	$5.3428e - 23$
10	(4.0007, 4.0006, 4.0080, 3.9995)	$-1.0529e + 01$	7	$4.5331e - 23$
Média	(4.0015, 3.9997, 4.0019, 3.9971)	-10.5270	7.5	$9.1629e - 22$

Tabela 3.22: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.16 com os parâmetros dispostos na Tabela 3.19, $m = 10$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi 0.0226.

objetivo $f : D \rightarrow \mathbb{R}$ dada por

$$f(x) = - \sum_{i=1}^{11} \left(e_i - x_1 \frac{v_i^2 + v_i x_2}{v_i^2 + v_i x_3 + x_4} \right)^2,$$

onde os parâmetros e_i e v_i são dados na Tabela 3.23.

i	1	2	3	4	5	6	7	8	9	10	11
e_i	0.1957	0.1947	0.1735	0.1600	0.08440	0.0627	0.0456	0.0342	0.0323	0.0235	0.0246
$1/v_i$	0.25	0.5	1	2	4	6	8	10	12	14	16

Tabela 3.23: Parâmetros para o Problema 3.17.

Na Tabela 3.24 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.17.

Problema 3.18 (Problema A.11, [56]). Dado $n \geq 2$ considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = \frac{\pi}{n} \{ \text{sen}^2(\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + 10 \text{sen}^2(\pi x_{i+1})] + (x_n - 1)^2 \}.$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -10 \leq x \leq 10\}$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(1.9284, 1.9084, 1.2313, 1.3577) \cdot e - 01$	$3.0749e - 04$	24	$3.3955e - 21$
2	$(1.9364, 1.7211, 1.1885, 1.2734) \cdot e - 01$	$3.0818e - 04$	34	$5.8526e - 21$
3	$(1.9283, 1.9080, 1.2310, 1.3575) \cdot e - 01$	$3.0749e - 04$	10	$1.8073e - 21$
4	$(1.9284, 1.9080, 1.2311, 1.3575) \cdot e - 01$	$3.0749e - 04$	11	$6.2527e - 22$
5	$(1.9274, 1.9289, 1.2351, 1.3672) \cdot e - 01$	$3.0749e - 04$	11	$3.8317e - 21$
6	$(1.9284, 1.9080, 1.2312, 1.3575) \cdot e - 01$	$3.0749e - 04$	10	$8.8743e - 22$
7	$(1.9283, 1.9090, 1.2312, 1.3579) \cdot e - 01$	$3.0749e - 04$	13	$1.6668e - 21$
8	$(1.9283, 1.9085, 1.2313, 1.3577) \cdot e - 01$	$3.0749e - 04$	7	$5.7578e - 21$
9	$(1.9283, 1.9084, 1.2309, 1.3577) \cdot e - 01$	$3.0749e - 04$	13	$2.6064e - 21$
10	$(1.9283, 1.9081, 1.2310, 1.3576) \cdot e - 01$	$3.0749e - 04$	12	$1.3764e - 21$
Média	$(0.1929, 0.1892, 0.1227, 0.1350)$	$3.0756e - 04$	14.5	$2.7807e - 21$

Tabela 3.24: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.17 com os parâmetros dispostos na Tabela 3.23, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.1820e - 07$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(1.0000e + 00, 1.0000e + 00)$	$1.6276e - 11$	8	$3.0816e - 22$
2	$(1.0000e + 00, 1.0000e + 00)$	$5.5946e - 11$	8	$3.7215e - 21$
3	$(1.0000e + 00, 1.0000e + 00)$	$4.3361e - 12$	8	$2.2813e - 23$
4	$(1.0000e + 00, 1.0000e + 00)$	$8.1157e - 11$	9	$8.6478e - 21$
5	$(1.0000e + 00, 1.0000e + 00)$	$4.3026e - 11$	9	$2.2931e - 21$
6	$(1.0000e + 00, 1.0000e + 00)$	$1.3773e - 11$	10	$2.2280e - 22$
7	$(1.0000e + 00, 1.0000e + 00)$	$2.2765e - 11$	7	$6.4178e - 22$
8	$(1.0000e + 00, 1.0000e + 00)$	$5.0106e - 12$	11	$3.0942e - 23$
9	$(1.0000e + 00, 1.0000e + 00)$	$3.0375e - 11$	9	$9.7049e - 22$
10	$(1.0000e + 00, 1.0000e + 00)$	$2.0346e - 12$	9	$5.5912e - 24$
Média	$(1, 1)$	$2.7470e - 11$	8.8	$1.6865e - 21$

Tabela 3.25: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.5745e - 11$.

Na Tabela 3.25 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 2$.

Na Tabela 3.26 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para

o Problema 3.18 com $n = 3$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(1.0000e + 00, 9.9794e - 01, 1.0000e + 00)	4.4366e - 06	9	4.0455e - 21
2	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	1.0888e - 11	6	8.9378e - 23
3	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	1.0696e - 11	7	1.4292e - 22
4	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	3.6946e - 12	8	1.3839e - 23
5	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	4.5249e - 12	8	1.2420e - 23
6	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	6.5005e - 12	7	4.2235e - 23
7	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	7.4982e - 11	8	6.4797e - 21
8	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	2.6769e - 11	7	5.5239e - 22
9	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	1.7816e - 12	7	3.6190e - 24
10	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	7.2555e - 12	7	5.0355e - 23
Média	(1.0000, 0.9998, 1.0000)	4.4367e - 07	7.4	1.1432e - 21

Tabela 3.26: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.4030e - 06$.

Na Tabela 3.27 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 4$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(1.0000e + 00, 1.0000e + 00, 9.9998e - 01, 1.0000e + 00)	3.1416e - 10	6	3.9642e - 22
2	(1.0000e + 00, 1.0000e + 00, 9.6289e - 01, 1.0000e + 00)	1.0816e - 03	7	5.0045e - 24
3	(1.0000e + 00, 1.0000e + 00, 1.0002e + 00, 1.0000e + 00)	3.1416e - 08	6	4.3625e - 24
4	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	1.1779e - 32	7	5.8148e - 24
5	(1.0000e + 00, 1.0000e + 00, 1.0061e + 00, 1.0000e + 00)	2.9225e - 05	5	0.0000e + 00
6	(1.0000e + 00, 1.0000e + 00, 9.9783e - 01, 1.0000e + 00)	3.6984e - 06	6	1.3750e - 23
7	(9.8967e - 01, 9.9982e - 01, 1.0000e + 00, 9.9973e - 01)	9.1076e - 04	5	0.0000e + 00
8	(1.0000e + 00, 1.0106e + 00, 1.0000e + 00, 1.0271e + 00)	6.6505e - 04	8	3.2581e - 22
9	(9.9936e - 01, 1.0000e + 00, 1.0000e + 00, 1.0713e + 00)	3.9962e - 03	7	9.4796e - 24
10	(9.9102e - 01, 1.0000e + 00, 1.0000e + 00, 1.0001e + 00)	6.8827e - 04	5	0.0000e + 00
Média	(0.9980, 1.0010, 0.9967, 1.0098)	7.3748e - 04	6.2	7.6064e - 23

Tabela 3.27: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.18 com $n = 4$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi de 0.0012.

Problema 3.19 (Problema A.12, [56]). Dado $n \geq 2$ considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = \xi(x) + \frac{\lfloor \xi(x) \rfloor}{n},$$

onde

$$\xi(x) = \sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)],$$

com $\lfloor z \rfloor$ sendo a parte inteira de z e o domínio de busca

$$D = \{x \in \mathbb{R}^n \mid -10 \leq x \leq 10\}.$$

Na Tabela 3.28 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 2$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(1.0000e + 00, 1.0000e + 00)	3.5199e - 12	9	1.5961e - 23
2	(1.0000e + 00, 1.0000e + 00)	9.3331e - 12	8	1.2979e - 22
3	(1.0000e + 00, 1.0000e + 00)	4.8074e - 11	10	2.7071e - 21
4	(1.0000e + 00, 1.0000e + 00)	8.3133e - 12	8	8.6645e - 23
5	(1.0000e + 00, 1.0000e + 00)	8.3606e - 11	8	8.9929e - 21
6	(1.0000e + 00, 1.0000e + 00)	5.1571e - 11	8	3.0791e - 21
7	(1.0000e + 00, 1.0000e + 00)	8.5327e - 11	8	9.5136e - 21
8	(1.0000e + 00, 1.0000e + 00)	4.6796e - 11	8	2.4330e - 21
9	(1.0000e + 00, 1.0000e + 00)	2.5493e - 11	8	8.0139e - 22
10	(1.0000e + 00, 1.0000e + 00)	3.6065e - 12	9	1.5885e - 23
Média	(1, 1)	3.6564e - 11	8.4	2.7775e - 21

Tabela 3.28: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $3.1435e - 11$.

Na Tabela 3.29 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 3$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	2.3435e - 11	8	5.5929e - 22
2	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	5.7187e - 11	9	3.8090e - 21
3	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	8.8040e - 12	8	1.5395e - 23
4	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	7.6588e - 11	7	6.8397e - 21
5	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	8.3526e - 11	7	7.5180e - 21
6	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	1.3791e - 11	7	2.3838e - 22
7	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	4.0169e - 12	8	1.5703e - 23
8	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	3.9271e - 12	8	1.8107e - 23
9	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	2.2043e - 12	8	3.4226e - 24
10	(1.0000e + 00, 1.0000e + 00, 1.0000e + 00)	2.6740e - 12	10	8.0534e - 24
Média	(1, 1, 1)	2.7615e - 11	8	1.9025e - 21

Tabela 3.29: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.19 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $3.2227e - 11$.

Problema 3.20 (Exemplo 5.6, [23]). Dado $n \geq 2$ considere $D = \{x \in \mathbb{R}^n \mid -1 \leq x \leq 1\}$ e a função objetivo $f : D \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} 1 + \left(\sum_{i=1}^n |x_i| \right)^{\frac{1}{2}} + \text{senal} \left(\text{sen} \left(\left(\sum_{i=1}^n |x_i| \right)^{\frac{1}{2}} - 0.5 \right) \right), & \text{se } x \neq 0, \\ 0, & \text{se } x = 0, \end{cases}$$

Na Tabela 3.30 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 2$.

Na Tabela 3.31 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 3$.

Problema 3.21 (Appendix Rastrigin, [35]). Dado n considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = 10n + \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i)).$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -2.56 \leq x \leq 5.12\}$.

Na Tabela 3.32 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 2$.

Na Tabela 3.33 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 3$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-2.7003e - 12, 9.3639e - 11)$	$0.0000e + 00$	17	$0.0000e + 00$
2	$(4.1739e - 11, 6.8838e - 11)$	$0.0000e + 00$	17	$0.0000e + 00$
3	$(-5.9117e - 11, -6.7090e - 11)$	$0.0000e + 00$	17	$0.0000e + 00$
4	$(-4.9266e - 11, 3.6661e - 11)$	$0.0000e + 00$	16	$0.0000e + 00$
5	$(8.2839e - 12, -9.8395e - 11)$	$0.0000e + 00$	13	$0.0000e + 00$
6	$(-1.2873e - 11, 9.4907e - 11)$	$0.0000e + 00$	14	$0.0000e + 00$
7	$(6.6900e - 11, 5.8530e - 11)$	$0.0000e + 00$	15	$0.0000e + 00$
8	$(-2.3677e - 12, 6.6074e - 12)$	$0.0000e + 00$	14	$0.0000e + 00$
9	$(2.1813e - 11, 4.8951e - 11)$	$0.0000e + 00$	18	$0.0000e + 00$
10	$(-7.8231e - 12, -9.2237e - 11)$	$0.0000e + 00$	19	$0.0000e + 00$
Média	$(0.0046, 0.1504) \cdot e - 10$	0	16	0

Tabela 3.30: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(3.2470e - 11, -4.8005e - 11, 7.3417e - 12)$	$0.0000e + 00$	15	$0.0000e + 00$
2	$(-3.2835e - 11, 6.8137e - 11, -6.2313e - 11)$	$0.0000e + 00$	14	$0.0000e + 00$
3	$(-6.9527e - 11, -2.2277e - 11, 3.3847e - 11)$	$0.0000e + 00$	16	$0.0000e + 00$
4	$(2.0919e - 11, 2.1644e - 11, -3.1782e - 11)$	$0.0000e + 00$	16	$0.0000e + 00$
5	$(7.9517e - 11, -5.0738e - 11, 2.6387e - 11)$	$0.0000e + 00$	13	$0.0000e + 00$
6	$(-4.5148e - 11, 1.3767e - 11, -8.4493e - 11)$	$0.0000e + 00$	13	$0.0000e + 00$
7	$(8.2652e - 11, -2.4006e - 12, -4.5499e - 11)$	$0.0000e + 00$	15	$0.0000e + 00$
8	$(1.5212e - 11, 5.2716e - 12, 2.7823e - 11)$	$0.0000e + 00$	12	$0.0000e + 00$
9	$(-7.7068e - 12, 4.2668e - 11, -5.2584e - 11)$	$0.0000e + 00$	13	$0.0000e + 00$
10	$(-1.1627e - 11, -1.8278e - 11, 8.0517e - 11)$	$0.0000e + 00$	16	$0.0000e + 00$
Média	$(0.0639, 0.0098, -0.1008) \cdot e - 10$	0	14.3	0

Tabela 3.31: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.20 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $0.0000e + 00$.

Problema 3.22 (Appendix Sum Squares, [35]). *Dado n considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por*

$$f(x) = \sum_{i=1}^n ix_i^2,$$

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-3.0128e-09, -7.0866e-09)$	$3.7613e-11$	15	$2.0841e-21$
2	$(-6.2751e-08, -1.1938e-08)$	$4.1183e-11$	13	$2.4586e-21$
3	$(-2.9578e-09, -7.6530e-09)$	$2.1020e-11$	14	$6.2819e-22$
4	$(7.5736e-09, 8.8968e-09)$	$1.3598e-12$	12	$2.2761e-24$
5	$(1.2333e-08, -9.7119e-08)$	$2.5045e-11$	11	$7.9932e-22$
6	$(4.3969e-08, -3.3267e-09)$	$1.7078e-11$	12	$3.5847e-22$
7	$(2.1673e-08, -2.5749e-09)$	$4.7066e-12$	11	$2.7710e-23$
8	$(-5.7374e-08, 7.1703e-09)$	$3.6200e-11$	12	$1.8863e-21$
9	$(4.2057e-09, -1.2038e-08)$	$1.9591e-12$	15	$5.8848e-24$
10	$(5.7467e-09, 2.0330e-08)$	$3.3868e-12$	12	$1.6060e-23$
Média	$(-0.0306, -0.1053) \cdot e - 07$	$1.8955e-11$	12.7	$8.2669e-22$

Tabela 3.32: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.5735e-11$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(2.0572e-09, 1.9775e-08, 2.1108e-08)$	$2.5827e-12$	9	$7.0074e-24$
2	$(2.0484e-08, -1.8370e-08, -1.7106e-07)$	$2.7203e-11$	12	$5.9064e-22$
3	$(-7.1441e-08, -1.0265e-09, -3.4594e-09)$	$2.4253e-12$	10	$2.5815e-24$
4	$(1.4483e-07, -1.3189e-08, -1.6499e-08)$	$4.7248e-11$	16	$2.2156e-21$
5	$(2.2610e-08, 1.8428e-08, 2.2797e-09)$	$4.2963e-12$	16	$2.1404e-23$
6	$(1.4510e-08, 1.2560e-08, 1.0929e-08)$	$2.4209e-12$	17	$6.3047e-24$
7	$(1.5590e-08, 7.7383e-10, 1.0855e-08)$	$4.5512e-12$	15	$2.4564e-23$
8	$(4.9113e-08, -4.3567e-09, 9.5088e-09)$	$5.4365e-12$	62	$3.0240e-23$
9	$(-2.9430e-08, -3.7408e-09, -2.8987e-07)$	$8.8251e-11$	18	$6.4199e-21$
10	$(-1.1816e-09, 1.4971e-08, 4.0739e-09)$	$2.4425e-12$	15	$7.1621e-24$
Média	$(0.1671, 0.0258, -0.4221) \cdot e - 07$	$1.8686e-11$	19	$9.3254e-22$

Tabela 3.33: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.21 com $n = 3$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.8601e-11$.

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -5 \leq x \leq 5\}$.

Na Tabela 3.34 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 2$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-5.8428e-07, 1.3615e-07)$	$3.5998e-11$	7	$1.6270e-21$
2	$(-4.5820e-07, 1.8156e-07)$	$4.8605e-11$	7	$2.9805e-21$
3	$(-2.1420e-07, 1.6971e-07)$	$5.6262e-12$	8	$3.9020e-23$
4	$(-3.8892e-07, 3.4048e-08)$	$2.5054e-11$	8	$7.5785e-22$
5	$(-1.9154e-07, -5.4210e-08)$	$2.5909e-11$	8	$9.5034e-22$
6	$(1.2179e-07, -5.1219e-07)$	$5.8673e-11$	7	$5.2573e-21$
7	$(1.4622e-07, 1.3925e-07)$	$2.4128e-11$	8	$7.9196e-22$
8	$(-3.5667e-08, 3.7571e-07)$	$3.3109e-12$	8	$1.2359e-23$
9	$(-6.9408e-07, 5.5433e-08)$	$5.2536e-11$	7	$3.5883e-21$
10	$(-1.7420e-07, -9.7231e-08)$	$1.1982e-11$	8	$1.8613e-22$
Média	$(-0.2473, 0.0428) \cdot e - 06$	$2.9182e-11$	7.6	$1.6191e-21$

Tabela 3.34: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.9471e-11$.

Na Tabela 3.35 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 3$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(3.1026e-07, 5.2664e-07, 2.8283e-07)$	$5.2202e-12$	7	$1.9700e-23$
2	$(3.7396e-07, 2.0032e-07, 9.3772e-07)$	$3.0607e-11$	7	$8.7560e-22$
3	$(-1.2812e-09, 2.3649e-08, 2.3885e-06)$	$2.2028e-11$	7	$2.9075e-23$
4	$(3.4872e-09, -9.9595e-08, -2.2885e-07)$	$4.0690e-12$	8	$1.8303e-23$
5	$(-9.7656e-07, 1.0095e-07, 1.1524e-06)$	$3.3899e-11$	7	$9.7319e-22$
6	$(3.5104e-06, 1.1512e-07, -1.0157e-06)$	$6.1253e-11$	7	$5.3673e-22$
7	$(4.2655e-07, 4.8010e-07, 8.9577e-08)$	$5.9976e-12$	7	$2.6573e-23$
8	$(-2.3254e-07, 9.2132e-08, -7.0891e-08)$	$1.2031e-11$	7	$1.6453e-22$
9	$(-1.7903e-07, -1.1390e-07, -7.9258e-06)$	$1.9386e-10$	7	$3.5157e-23$
10	$(1.7958e-07, 2.7224e-07, -3.1347e-08)$	$2.2193e-12$	7	$4.1420e-24$
Média	$(0.3415, 0.1598, -0.4422) \cdot e - 06$	$3.7118e-11$	7.1	$2.6830e-22$

Tabela 3.35: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.22 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $5.8077e-11$.

Problema 3.23 (Appendix Sphere, [35]). Dado n considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = \sum_{i=1}^n x_i^2,$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -2.56 \leq x \leq 5.12\}$.

Na Tabela 3.36 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 2$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(3.7314e - 08, -1.4844e - 07)$	$9.3854e - 12$	7	$1.0516e - 22$
2	$(-5.6527e - 08, 1.1707e - 06)$	$2.6084e - 11$	7	$8.3641e - 22$
3	$(-1.2186e - 08, -2.7571e - 07)$	$1.4165e - 11$	7	$2.7409e - 22$
4	$(-4.2219e - 08, 1.9477e - 08)$	$2.7078e - 12$	7	$9.6243e - 24$
5	$(6.9105e - 08, 1.5484e - 08)$	$2.2347e - 12$	7	$6.4235e - 24$
6	$(-2.8641e - 07, -4.1767e - 07)$	$9.8431e - 12$	7	$1.2816e - 22$
7	$(-2.8472e - 07, -1.8712e - 06)$	$7.4353e - 11$	6	$6.2670e - 21$
8	$(-1.2799e - 06, 8.1934e - 07)$	$5.8954e - 11$	7	$3.9825e - 21$
9	$(6.4146e - 07, -1.1219e - 07)$	$8.1545e - 11$	7	$8.9064e - 21$
10	$(-8.5170e - 07, 5.4943e - 08)$	$3.0926e - 11$	7	$1.1333e - 21$
Média	$(-0.2066, -0.0745) \cdot e - 06$	$3.1020e - 11$	6.9	$2.1649e - 21$

Tabela 3.36: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.9937e - 11$.

Na Tabela 3.37 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 3$.

Na Tabela 3.38 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 4$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-4.1778e-07, -1.6440e-08, -2.2199e-07)$	$3.0815e-12$	5	$9.6391e-24$
2	$(-1.7198e-07, -9.2310e-07, 2.7837e-08)$	$1.2827e-11$	5	$1.6107e-22$
3	$(-3.6250e-07, -1.1795e-06, -9.4995e-07)$	$2.3206e-11$	5	$4.8710e-22$
4	$(-5.3929e-07, 1.3038e-06, -9.1259e-07)$	$7.1258e-11$	5	$5.5985e-21$
5	$(6.0442e-08, 3.9880e-07, -1.2907e-07)$	$2.8396e-12$	5	$9.2046e-24$
6	$(8.9299e-07, 5.3806e-07, 1.4366e-05)$	$2.2517e-10$	5	$4.0876e-22$
7	$(3.3323e-07, 4.7065e-08, 2.3916e-07)$	$2.3372e-12$	5	$5.2813e-24$
8	$(-9.6500e-07, 1.1169e-06, 3.3531e-06)$	$5.8711e-11$	5	$2.4274e-21$
9	$(7.6771e-08, -9.4358e-08, 5.2029e-08)$	$4.1890e-13$	5	$2.1293e-25$
10	$(-1.9202e-07, -3.7581e-06, 4.6258e-07)$	$2.0575e-11$	5	$4.7179e-23$
Média	$(-0.0129, -0.0257, 0.1629) \cdot e-05$	$4.2042e-11$	5	$9.1543e-22$

Tabela 3.37: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 3$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $6.8880e-11$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-0.0094, 0.0072, 0.0212, -0.0034) \cdot e-04$	$1.4555e-11$	8	$8.1771e-23$
2	$(-0.0032, -0.0117, -0.0176, 0.0232) \cdot e-04$	$7.4520e-11$	8	$5.1598e-21$
3	$(-0.0018, 0.0304, 0.0176, 0.0022) \cdot e-04$	$4.2274e-11$	8	$1.0039e-21$
4	$(-0.0143, 0.0042, 0.0002, 0.0015) \cdot e-04$	$5.7350e-12$	9	$1.4932e-23$
5	$(0.0012, -0.0025, 0.0044, -0.0050) \cdot e-04$	$1.0336e-11$	9	$1.1644e-22$
6	$(0.0014, 0.0011, -0.0284, 0.0130) \cdot e-04$	$2.0099e-11$	8	$1.2560e-22$
7	$(0.1847, 0.0241, -0.0178, -0.0136) \cdot e-04$	$4.1873e-10$	8	$5.1044e-21$
8	$(-0.0210, 0.0289, 0.0297, -0.0526) \cdot e-04$	$1.3511e-10$	7	$8.7250e-21$
9	$(-0.0104, 0.0035, 0.0389, -0.0084) \cdot e-04$	$5.7185e-11$	8	$1.8695e-21$
10	$(-0.0015, 0.2739, -0.0079, -0.0117) \cdot e-04$	$7.8132e-10$	8	$1.0606e-21$
Média	$(0.1258, 0.3591, 0.0403, -0.0547) \cdot e-05$	$1.5599e-10$	8.1	$2.3262e-21$

Tabela 3.38: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.23 com $n = 4$, $N = 20$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $2.5212e-10$.

Problema 3.24 (Problema A.13, [56]). Dado n considere a função objetivo $f : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} 1 + \frac{\sum_{i=1}^n |x_i|}{n} + \text{sinal} \left(\sin \left(\frac{n}{\sum_{i=1}^n |x_i|} \right) - 0.5 \right), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -1 \leq x \leq 1\}$.

Na Tabela 3.39 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 2$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(-4.9791e - 12, 3.8855e - 11)$	$1.1157e - 10$	14	$7.7423e - 21$
2	$(-6.1627e - 11, 1.4777e - 11)$	$9.6735e - 11$	15	$6.8683e - 21$
3	$(6.9810e - 11, 6.0344e - 11)$	$0.0000e + 00$	14	$0.0000e + 00$
4	$(-4.2016e - 11, 4.7452e - 11)$	$3.2404e - 11$	13	$5.4475e - 21$
5	$(-2.5637e - 11, 8.3556e - 11)$	$0.0000e + 00$	14	$0.0000e + 00$
6	$(2.2689e - 11, 5.9128e - 11)$	$2.0258e - 11$	14	$4.6293e - 21$
7	$(4.6936e - 11, 3.3000e - 11)$	$6.0816e - 11$	13	$7.1450e - 21$
8	$(-2.5526e - 11, 2.3261e - 11)$	$0.0000e + 00$	15	$0.0000e + 00$
9	$(4.4444e - 11, -1.4592e - 11)$	$0.0000e + 00$	16	$0.0000e + 00$
10	$(-9.6016e - 11, -3.2923e - 12)$	$5.5744e - 11$	14	$5.4171e - 21$
Média	$(-0.0719, 0.3425) \cdot e - 10$	$3.7753e - 11$	14.2	$3.7249e - 21$

Tabela 3.39: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 2$, $N = 100$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $4.1925e - 11$.

Na Tabela 3.40 são apresentados os resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 3$.

Rodada	$\bar{x}$	$f(\bar{x})$	Iterações	VF
1	$(8.1981e - 11, 1.3207e - 11, -4.1253e - 11)$	$1.0787e - 10$	13	$3.5615e - 21$
2	$(-5.5979e - 11, 2.9209e - 11, 4.4464e - 11)$	$1.5296e - 10$	13	$3.2015e - 21$
3	$(7.6581e - 11, 3.0246e - 11, -5.4594e - 11)$	$7.1380e - 11$	12	$2.7498e - 21$
4	$(-9.8301e - 11, 1.8411e - 11, 2.8841e - 11)$	$1.4488e - 10$	12	$3.4204e - 21$
5	$(-9.3302e - 11, -2.2592e - 11, -1.8966e - 12)$	$1.5608e - 10$	14	$6.2739e - 21$
6	$(1.1192e - 11, 2.4383e - 12, 1.0674e - 10)$	$1.3588e - 10$	15	$3.0923e - 21$
7	$(5.5229e - 11, -6.1423e - 13, 7.1478e - 11)$	$1.7768e - 10$	12	$3.6715e - 21$
8	$(-9.8736e - 11, -1.1872e - 08, -1.7750e - 11)$	$4.1001e - 09$	11	$4.9715e - 21$
9	$(-8.0546e - 11, -1.0350e - 10, 2.5095e - 12)$	$1.8056e - 10$	12	$2.5510e - 21$
10	$(3.2757e - 11, 3.6604e - 11, -5.1827e - 11)$	$2.1199e - 10$	13	$4.6967e - 21$
Média	$(-0.0017, -0.1187, 0.0009) \cdot e - 08$	$5.4394e - 10$	12.7	$3.8190e - 21$

Tabela 3.40: Resultados obtidos para 10 rodadas do Algoritmo 3.5 para o Problema 3.24 com $n = 3$, $N = 50$ e $t = 200$. O desvio padrão dos valores obtidos para $f(\bar{x})$ foi $1.2501e - 09$.

3.3 Problema Multiobjetivo

Nesta seção, propomos um algoritmo para problemas de otimização multiobjetivo, da forma:

$$\begin{aligned} &\text{Minimize } F(x) = (f_1(x), \dots, f_r(x)) \\ &x \in X, \end{aligned} \tag{VOP}$$

onde: $f_\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, para $\ell \in L := \{1, \dots, r\}$, são as funções objetivo e $X \subseteq \mathbb{R}^n$.

3.3.1 Algoritmo

Consideremos primeiramente o conjunto W de pesos, dado por

$$W^* = \{w \in \mathbb{R}^r \mid w_\ell > 0, \ell \in L, \text{ e } \sum_{\ell=1}^r w_\ell = 1\}.$$

Dado $w \in W$, temos então o problema de Chebyshev ponderado (WCS_w) correspondente a (VOP):

$$\begin{aligned} &\text{Minimizar } \|f(x) - f^*\|_\infty^w = \Psi_w(x) \\ &x \in X \end{aligned}$$

onde $\|f(x) - f^*\|_\infty^w = \max_\ell \{w_\ell(f_\ell(x) - f_\ell^*)\}$ e as componentes f_ℓ^* são obtidas minimizando cada uma das funções objetivos sujeitas individualmente às restrições, isto é, resolvendo $f_\ell^* = \min_{x \in X} f_\ell(x)$ para $\ell = 1, \dots, r$.

Apesar desta técnica de escalarização ser simples, implica em uma decisão que pode não ser, pois para determinar a fronteira de Pareto, o problema deve ser resolvido iterativamente utilizando diferentes conjuntos de pesos em cada iteração.

Dos resultados teóricos obtidos na Seção 2.2, temos que os pesos devem ser necessariamente todos positivos e não nulos. Em nosso algoritmo tomamos os pesos de forma aleatória. Portanto, para definir tal conjunto de pesos, consideramos primeiramente um conjunto $\overline{W} = \{\overline{w} \in \mathbb{R}^r \mid (\overline{w}_1, \dots, \overline{w}_r)\}$ de vetores com as componentes sendo números aleatórios uniformemente distribuídos sobre o intervalo $(0, 1)$. Sendo assim, o conjunto de pesos é dado por:

$$W = \left\{ \frac{\overline{w}_1}{\|\overline{w}\|_s}, \dots, \frac{\overline{w}_r}{\|\overline{w}\|_s} \right\}, \quad (3.16)$$

onde $\|\cdot\|_s$ denota a norma da soma em $\mathbb{R}^r$.

Temos então o seguinte algoritmo.

Algoritmo 3.6. Método do Valor Médio dos conjuntos de nível para problemas Multiobjetivos

Passo 0 - Escalarização

(a) Para cada $\ell \in L$ encontrar x_ℓ^* que resolve o seguinte problema escalar:

$$\min_{x \in X} f_\ell(x). \quad (3.17)$$

Defina o vetor ideal $f^* = (f_1(x_1^*), \dots, f_r(x_r^*))$.

(b) Dado $w \in W$ como em (3.16), defina

$$\Psi_w(x) = \max_{\ell} \{w_\ell(f_\ell(x) - f_\ell^*)\}.$$

Passo 1 - Tome um ponto $x_0 \in X$ e $\varepsilon > 0$. Seja

$$c_0 = \Psi(x_0) \quad \text{e} \quad H_{c_0} = \{x \in X \mid \Psi(x) \leq c_0\}.$$

Defina $k := 0$.

Passo 2 - Calcule o valor médio

$$c_{k+1} = M(\Psi, c_k, X) = \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} \Psi(x) d\mu$$

e seja

$$H_{c_{k+1}} = \{x \in X \mid \Psi(x) \leq c_{k+1}\}.$$

Passo 3 - Calcule a variância modificada

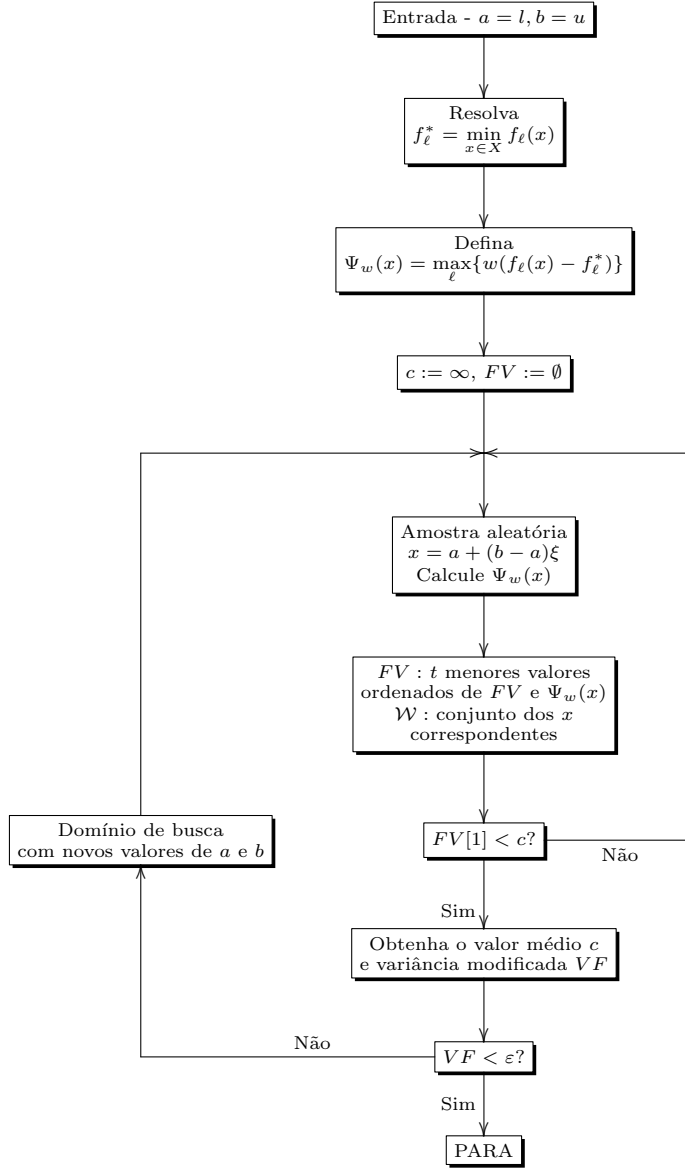
$$VF = V_1(\Psi, c_k, X) = \frac{1}{\mu(H_{c_k} \cap X)} \int_{H_{c_k} \cap X} (\Psi(x) - c_k)^2 d\mu.$$

Passo 4 - Se $VF \geq \varepsilon$, seja $k := k + 1$ e volte para o passo 2, caso contrário vá para o passo 5.

Passo 5 - Defina $\bar{c} = c_{k+1}$ e $\overline{H} = H_{c_{k+1}}$. PARE.

Observação 3.25. Na resolução do problema (3.17) para cada objetivo pode ser aplicado o Algoritmo 3.3 do Método do Valor Médio dos conjuntos de nível para problemas mono objetivo.

Fluxograma do Algoritmo 3.6



3.3.2 Testes Numéricos

Consideremos valor inicial de $c_0 = 10^{12}$, $\varepsilon = 10^{-20}$, $N = 100$ (N pontos aleatórios uniformemente distribuídos sobre o cuboide), $t = 20$ e o cuboide $D_0 = D$ fornecido em cada problema. Diferentemente do Algoritmo 3.3 para problemas mono objetivo no Algoritmo 3.6 não foi fixado o número de rodadas, ou seja, foi rodado m vezes para cada um dos 22 problemas considerados.

Para cada rodada o conjunto de pesos foi tomado de acordo com (3.16). Sendo assim, em cada rodada temos um conjunto de pesos distinto do anterior.

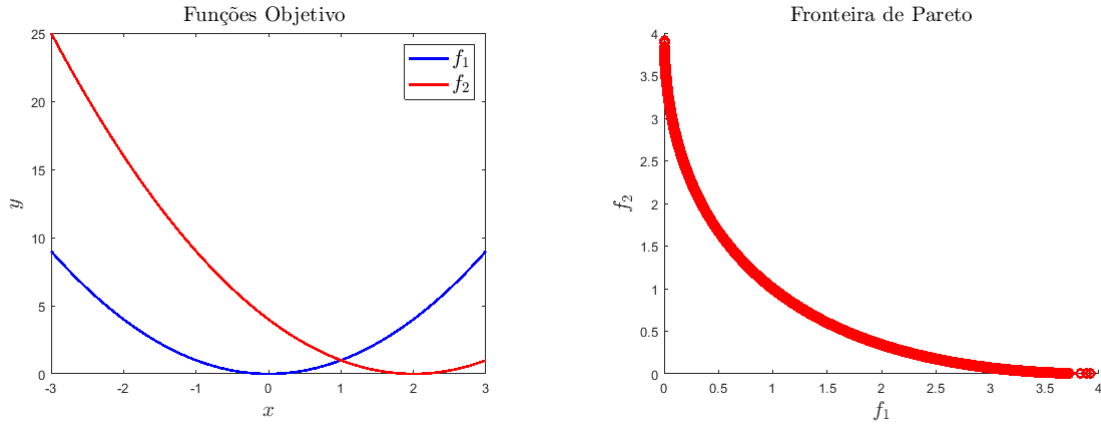
Em cada problema descrevemos quais são os objetivos e o domínio de busca.

Problema 3.26 (Problema SCH1, [49]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(x) &= x^2, \\ f_2(x) &= (x - 2)^2, \end{aligned}$$

e o domínio de busca $D = \mathbb{R}$.

A Figura 3.22 ilustra o Problema 3.26. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , sendo fácil identificar que as soluções Pareto-ótimas estão no intervalo $[0, 2]$, já que fora deste intervalo ambas funções crescem. Note que neste intervalo acontece uma compensação entre as funções objetivo, isto é, enquanto uma função cresce a outra diminui. À direita é representada a fronteira de Pareto, a qual foi gerada para 10.000 rodadas do Algoritmo 3.6.



Fonte: O autor

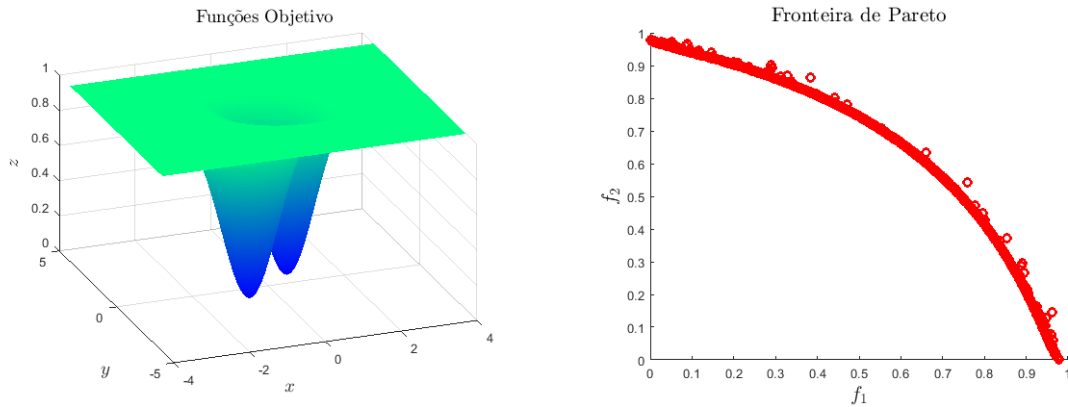
Figura 3.22: Problema 3.26

Problema 3.27 (Problema 6.1, [14]). *Dado n considere as funções objetivo $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$*

$$\begin{aligned} f_1(x) &= 1 - \exp\left(-\sum_{i=1}^n \left(x_i - \frac{1}{\sqrt{n}}\right)^2\right), \\ f_2(x) &= 1 - \exp\left(-\sum_{i=1}^n \left(x_i + \frac{1}{\sqrt{n}}\right)^2\right), \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^n \mid -4 \leq x \leq 4\}$.

A Figura 3.23 ilustra o Problema 3.27. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto. Consideremos $n = 2$, onde as funções não são convexas. A fronteira de Pareto foi gerada para 1.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.23: Problema 3.27

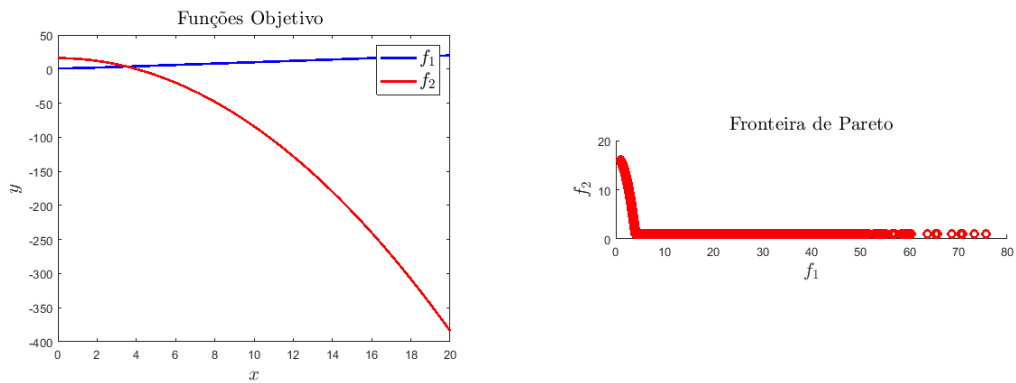
Problema 3.28 (Exemplo 9, [9]). Considere as funções objetivo $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ dadas por

$$f_1(x) = \sqrt{x^2 + 1}$$

$$f_2(x) = \begin{cases} -x^2 + 16, & 0 \leq x \leq \sqrt{15} \\ 1, & x > \sqrt{15} \end{cases},$$

e o domínio de busca $D = \{x \in \mathbb{R} \mid x \geq 0\}$

A Figura 3.24 ilustra o Problema 3.28. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto gerada para 5.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.24: Problema 3.28

Problema 3.29 (Exemplo 4.3.6, [9]). *Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

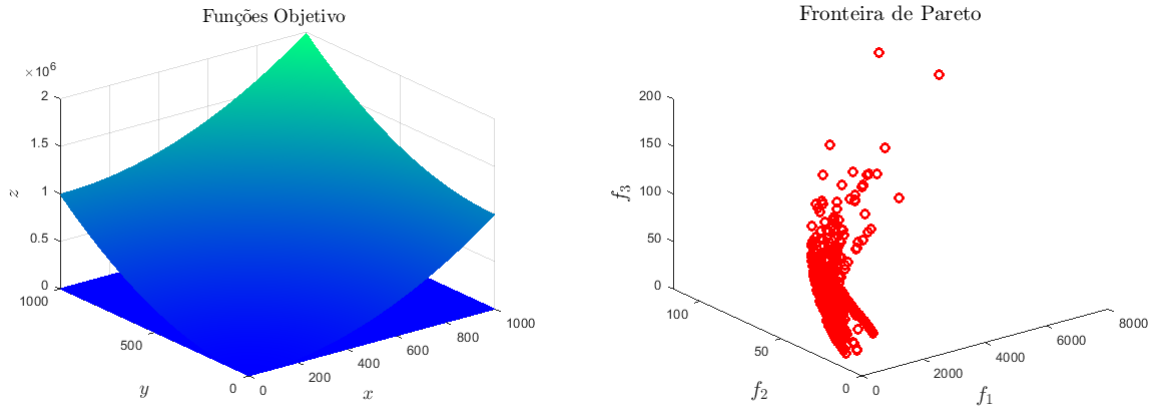
$$f_1(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2,$$

$$f_2(x_1, x_2) = x_1 + x_2,$$

$$f_3(x_1, x_2) = x_1 + 2x_2,$$

e o domínio de busca $X = \{x \in \mathbb{R}^2 \mid g_1(x) = -x_1 \leq 0 \text{ e } g_2(x) = -x_2 \leq 0\}$.

A Figura 3.25 ilustra o Problema 3.29. A imagem da esquerda representa o gráfico das funções objetivo f_1, f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.25: Problema 3.28

Problema 3.30 (Problema 4.7, [9]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ dadas por*

$$f_1(x) = x - 1,$$

$$f_2(x) = (x - 3)^2 + 1,$$

e o domínio de busca $X = \{x \in \mathbb{R} \mid x \geq 0\}$.

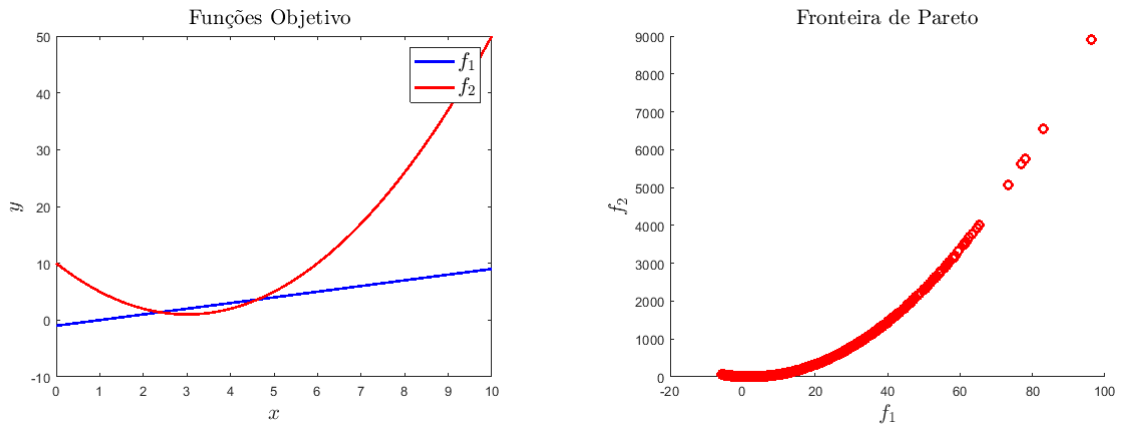
A Figura 3.26 ilustra o Problema 3.30. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 10.000 rodadas do Algoritmo 3.6.

Problema 3.31 (MOP 2, [51]). *Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$f_1(x_1, x_2) = x_1 - 10^6,$$

$$f_2(x_1, x_2) = x_2 - 2 \cdot 10^{-6},$$

$$f_3(x_1, x_2) = x_1 x_2 - 2,$$

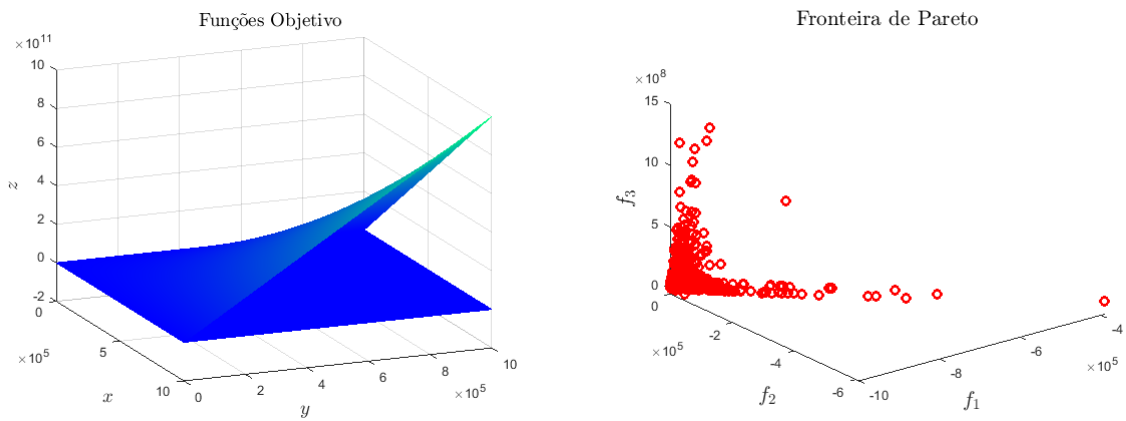


Fonte: O autor

Figura 3.26: Problema 3.30

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid 10^{-6} \leq x \leq 10^6\}$.

A Figura 3.27 ilustra o Problema 3.31. A imagem da esquerda representa o gráfico das funções objetivo f_1 , f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 1.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 7.0756 segundos.



Fonte: O autor

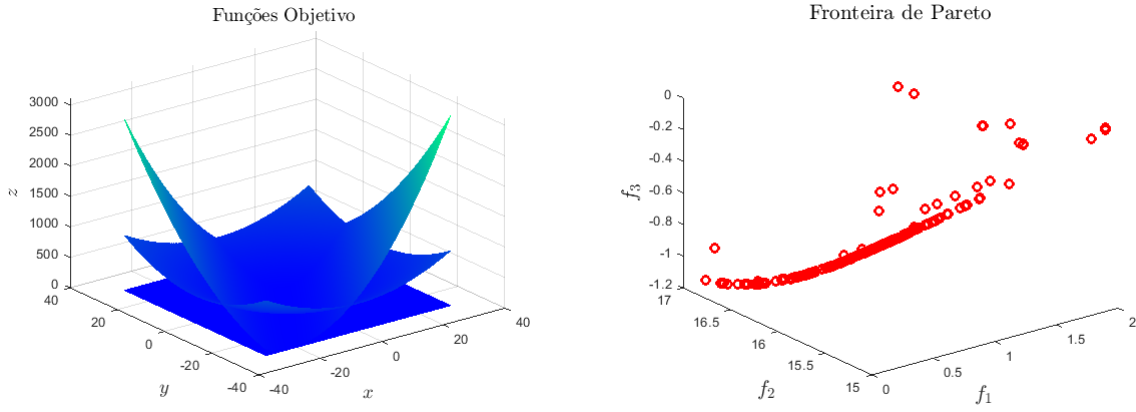
Figura 3.27: Problema 3.31

Problema 3.32 (MOP 18, [51]). Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= 0.5(x_1^2 + x_2^2) + \sin(x_1^2 + x_2^2), \\ f_2(x_1, x_2) &= \frac{1}{8}(3x_1 - 2x_2 + 4)^2 + \frac{1}{27}(x_1 - x_2 + 1)^2 + 15, \\ f_3(x_1, x_2) &= \frac{1}{(x_1^2 + x_2^2 + 1)} - 1.1e^{-x_1^2 - x_2^2}, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -3 \leq x \leq 3\}$.

A Figura 3.28 ilustra o Problema 3.32. A imagem da esquerda representa o gráfico das funções objetivo f_1 , f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 200 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.28: Problema 3.32

Problema 3.33 (MOP 1, [51]). Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

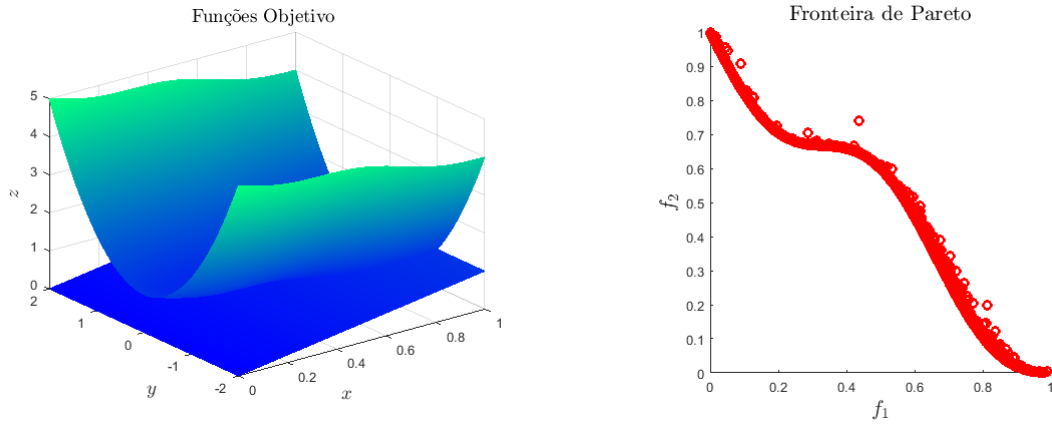
$$\begin{aligned} f_1(x_1, x_2) &= x_1, \\ f_2(x_1, x_2) &= 1 + x_2^2 - x_1 - 0.1\sin(3\pi x_1). \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1 \text{ e } -2 \leq x_2 \leq 2\}$.

A Figura 3.29 ilustra o Problema 3.33. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 5.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 1.8992 segundos.

Problema 3.34 (MOP 3, [51]). Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= 1 - e^{-(x_1-1)^2 - (x_2+1)^2}, \\ f_2(x_1, x_2) &= 1 - e^{-(x_1+1)^2 - (x_2-1)^2}, \end{aligned}$$

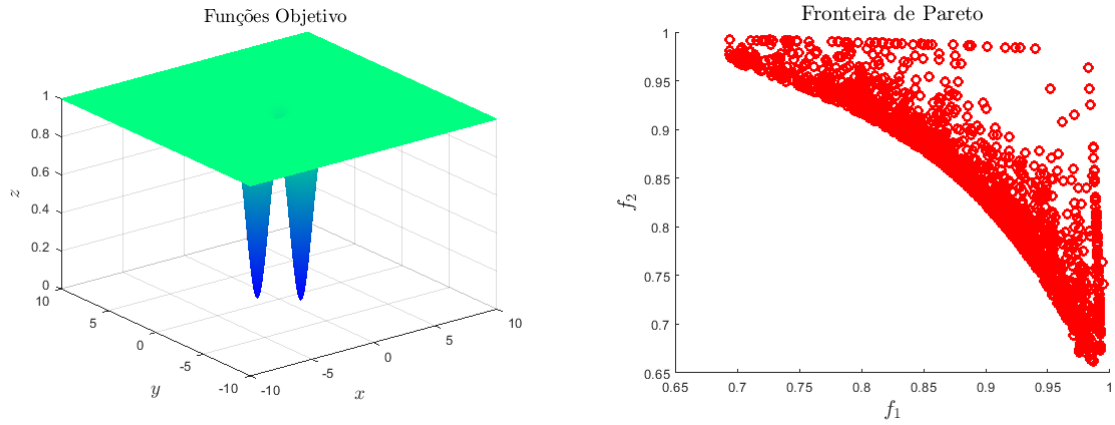


Fonte: O autor

Figura 3.29: Problema 3.33

e o domínio de busca $D = \mathbb{R}^2$.

A Figura 3.30 ilustra o Problema 3.34. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 7.6871 segundos.



Fonte: O autor

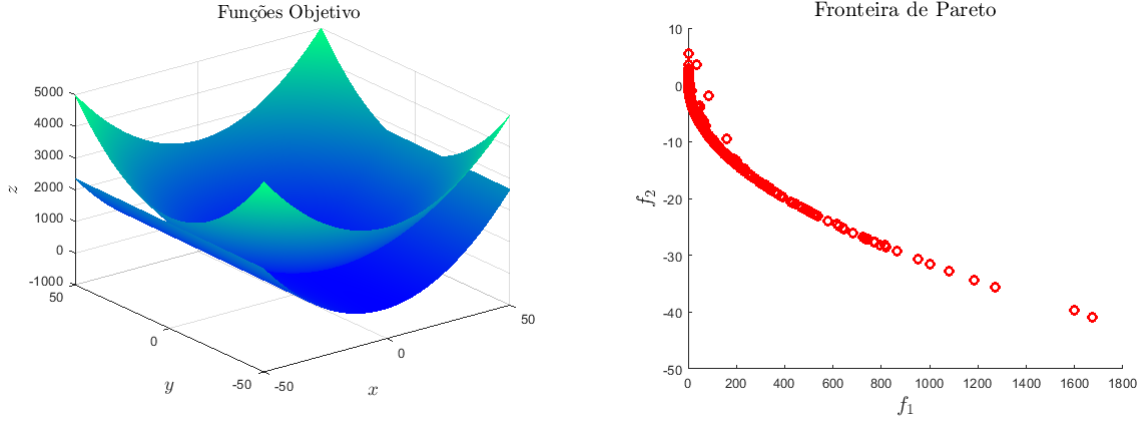
Figura 3.30: Problema 3.34

Problema 3.35 (MOP 6, [51]). Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2, \\ f_2(x_1, x_2) &= (x_1 + 2)^2 + x_2, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -50 \leq x \leq 50\}$.

A Figura 3.31 ilustra o Problema 3.35. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 1.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 88.3943 segundos.



Fonte: O autor

Figura 3.31: Problema 3.35

Problema 3.36 (MOP 7, [51]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(x_1, x_2) &= \sqrt[8]{x_1^2 + x_2^2}, \\ f_2(x_1, x_2) &= \sqrt[4]{(x_1 - 0.5)^2 + (x_2 - 0.5)^2}, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -5 \leq x \leq 10\}$.

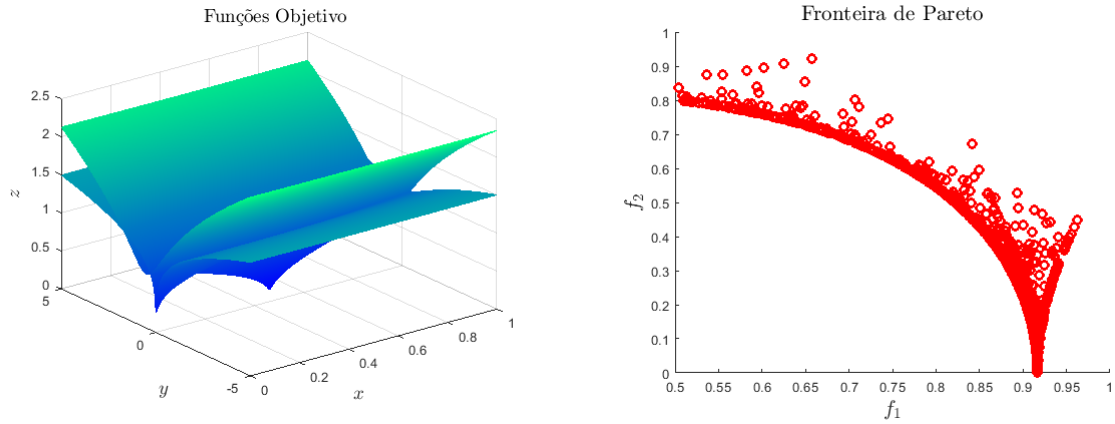
A Figura 3.32 ilustra o Problema 3.36. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 1.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 41.5521 segundos.

Problema 3.37 (MOP 8, [51]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(x_1, x_2) &= 2\sqrt{x_1}, \\ f_2(x_1, x_2) &= x_1(1 - x_2) + 5, \end{aligned}$$

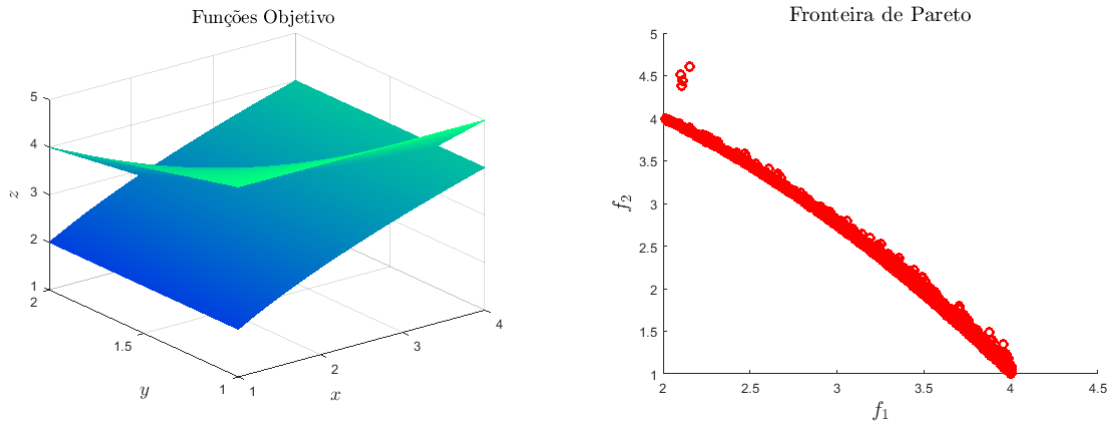
e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid 1 \leq x_1 \leq 4, 1 \leq x_2 \leq 2\}$.

A Figura 3.33 ilustra o Problema 3.37. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 2.6056 segundos.



Fonte: O autor

Figura 3.32: Problema 3.36



Fonte: O autor

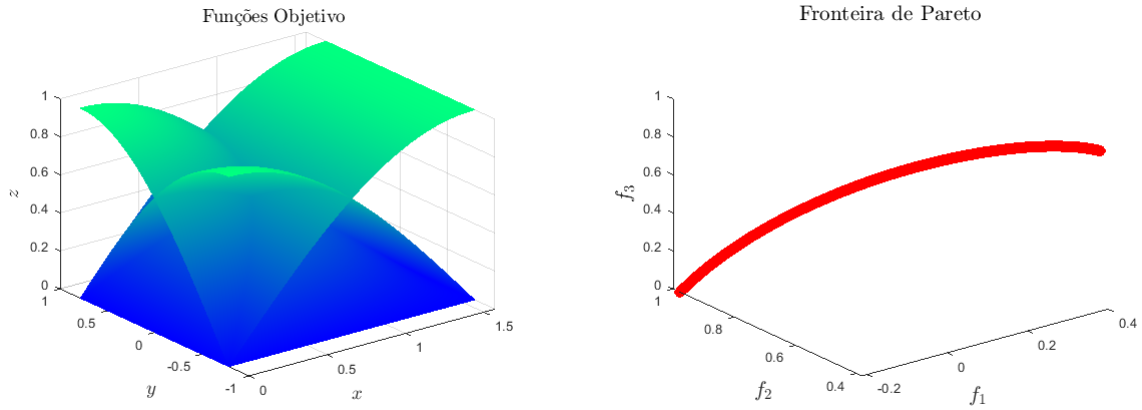
Figura 3.33: Problema 3.36

Problema 3.38 (Problema 6.2, [14]). *Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(\theta, \gamma) &= \cos \theta \cos \left(\gamma + \frac{\pi}{4} \right), \\ f_2(\theta, \gamma) &= \cos \theta \sin \left(\gamma + \frac{\pi}{4} \right), \\ f_3(\theta, \gamma) &= \sin \theta. \end{aligned}$$

e o domínio de busca $D = \left\{ (\theta, \gamma) \in \mathbb{R}^2 \mid 0 \leq \theta \leq \frac{\pi}{2} \text{ e } -\frac{\pi}{4} \leq \gamma \leq \frac{\pi}{4} \right\}$.

A Figura 3.34 ilustra o Problema 3.38. A imagem da esquerda representa o gráfico das funções objetivo f_1, f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 15.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.34: Problema 3.38

Problema 3.39 (Problema 6.28, [14]). *Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^3 \rightarrow \mathbb{R}$ dadas por*

$$f_1(x_1, x_2, x_3) = (1 + x_3)(x_1^3 x_2^2 - 10x_1 - 4x_2),$$

$$f_2(x_1, x_2, x_3) = (1 + x_3)(x_1^3 x_2^2 - 10x_1 + 4x_2),$$

$$f_3(x_1, x_2, x_3) = 3(1 + x_3)x_1^2;$$

e o domínio de busca $D = \{x \in \mathbb{R}^3 \mid -5 \leq x \leq 10\}$.

A Figura 3.35 ilustra o Problema 3.39, onde é representada a fronteira de Pareto, a qual foi gerada para 2.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 50.6280 segundos.

Problema 3.40 (MOP 9, [51]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$f_1(x_1, x_2) = -(1 + (A_1 - B_1)^2 + (A_2 - B_2)^2),$$

$$f_2(x_1, x_2) = -((x_1 + 3)^2 + (x_2 + 1)^2),$$

onde

$$A_1 = 0.5\sin(1) - 2\cos(1) + \sin(2) - 1.5\cos(2),$$

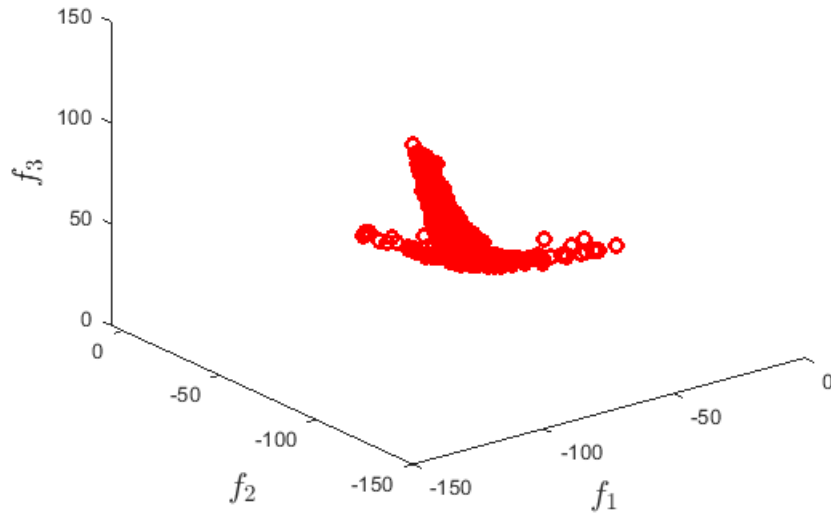
$$A_2 = 1.5\sin(1) - \cos(1) + 2\sin(2) - 0.5\cos(2),$$

$$B_1 = 0.5\sin(x_1) - 2\cos(x_1) + \sin(x_2) - 1.5\cos(x_2),$$

$$B_2 = 1.5\sin(x_1) - \cos(x_1) + 2\sin(x_2) - 0.5\cos(x_2),$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -\pi \leq x \leq \pi\}$.

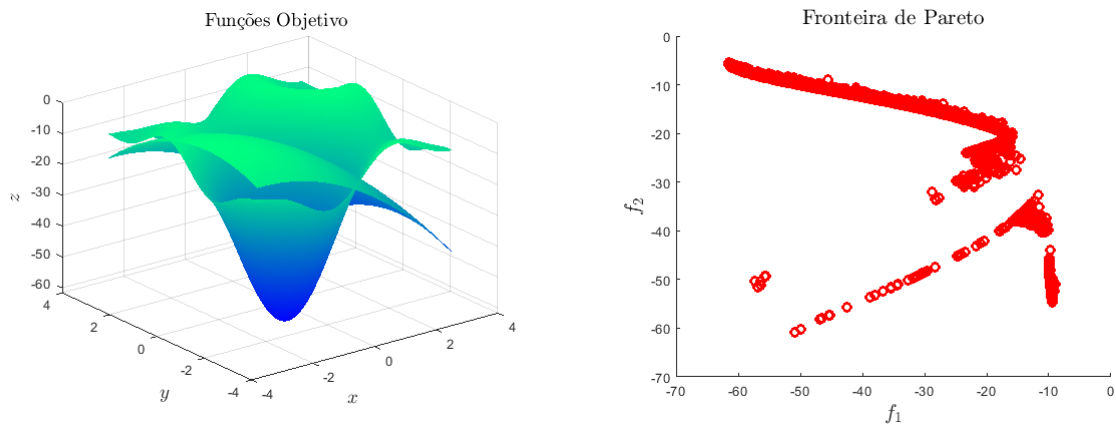
Fronteira de Pareto



Fonte: O autor

Figura 3.35: Problema 3.39

A Figura 3.36 ilustra o Problema 3.40. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 3.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 10.93353 segundos.



Fonte: O autor

Figura 3.36: Problema 3.40

Problema 3.41 (MOP 10, [51]). Dado n considere as funções objetivo $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ dadas por

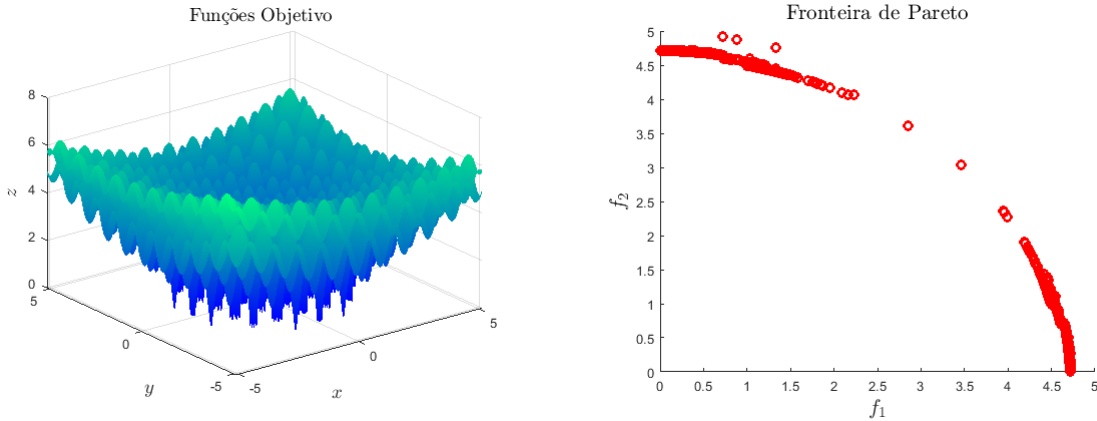
$$\begin{aligned} f_1(x) &= \sqrt{\frac{A_1}{n}}, \\ f_2(x) &= \sqrt{\frac{A_2}{n}}, \end{aligned}$$

onde

$$\begin{aligned} A_1 &= \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10), \\ A_2 &= \sum_{i=1}^n ((x_i - 1.5)^2 - 10 \cos(2\pi(x_i - 1.5)) + 10), \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -5.12 \leq x \leq 5.12\}$.

A Figura 3.37 ilustra o Problema 3.41. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 para $n = 2$, enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 500 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.37: Problema 3.41

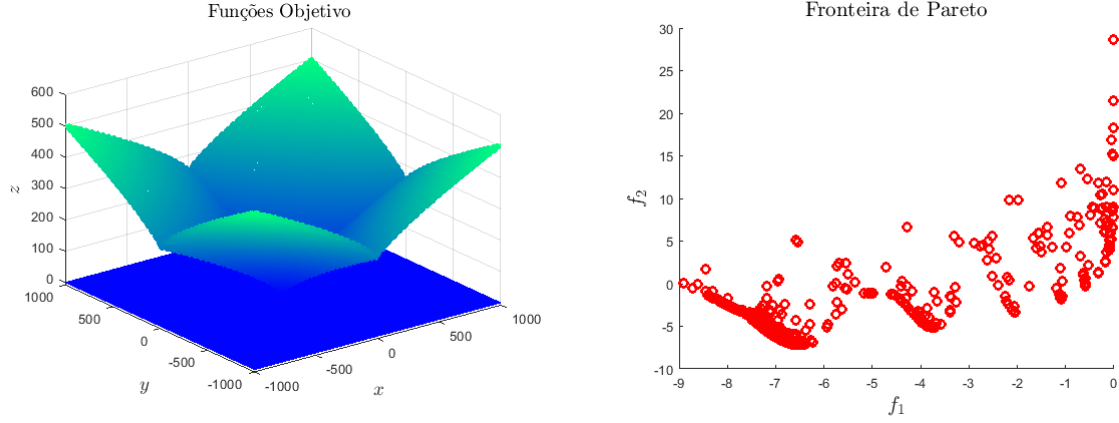
Problema 3.42 (MOP 5, [51]). Dado n considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= \sum_{i=1}^{n-1} -10e^{-0.2 \cdot \sqrt{x_i^2 + x_{i+1}^2}}, \\ f_2(x_1, x_2) &= \sum_{i=1}^n (|x_i|^{0.8} + 5\sin^3(x_i)), \end{aligned}$$

e o domínio de busca $D = \mathbb{R}^2$.

A Figura 3.38 ilustra o Problema 3.42. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000

rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 36.18945 segundos.



Fonte: O autor

Figura 3.38: Problema 3.42

Problema 3.43 (MOP 11, [51]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(x_1, x_2) &= \frac{1}{x_1^2 + x_2^2 + 1}, \\ f_2(x_1, x_2) &= x_1^2 + 3x_2^2 + 1, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -3 \leq x \leq 3\}$.

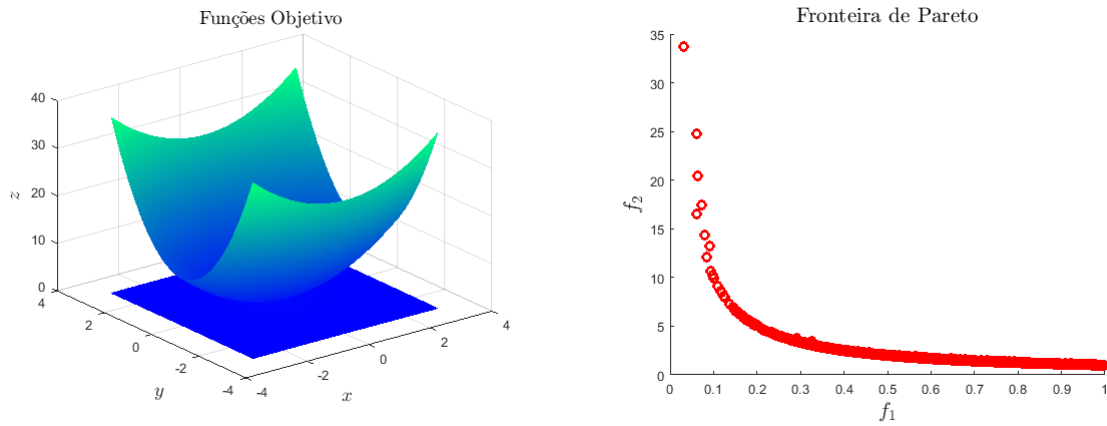
A Figura 3.39 ilustra o Problema 3.43. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 4.000 rodadas do Algoritmo 3.6. O tempo médio de execução do algoritmo em cada rodada foi de 1.65995 segundos.

Problema 3.44 (MOP 12, [51]). *Considere as funções objetivo $f_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por*

$$\begin{aligned} f_1(x_1, x_2) &= x_1 + x_2 + 1, \\ f_2(x_1, x_2) &= x_1^2 + 2x_2 - 1, \end{aligned}$$

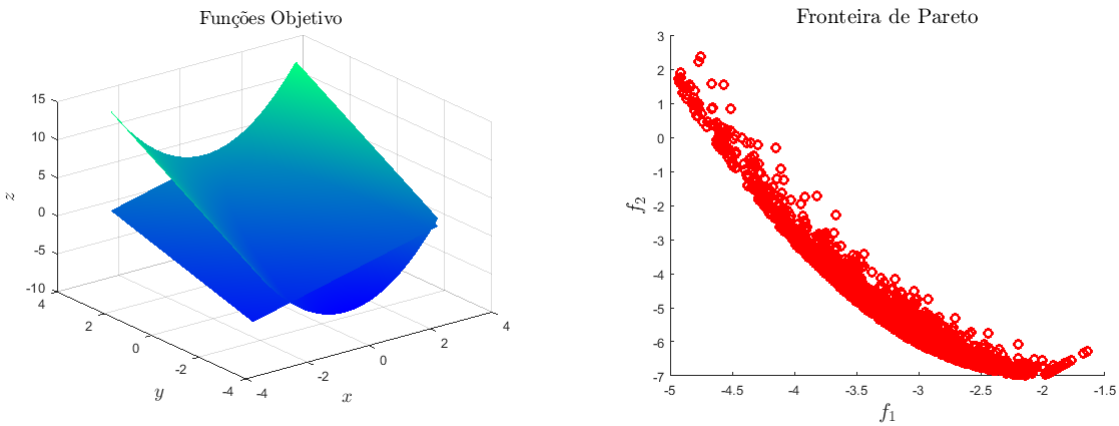
e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -3 \leq x \leq 3\}$.

A Figura 3.40 ilustra o Problema 3.44. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.39: Problema 3.43



Fonte: O autor

Figura 3.40: Problema 3.44

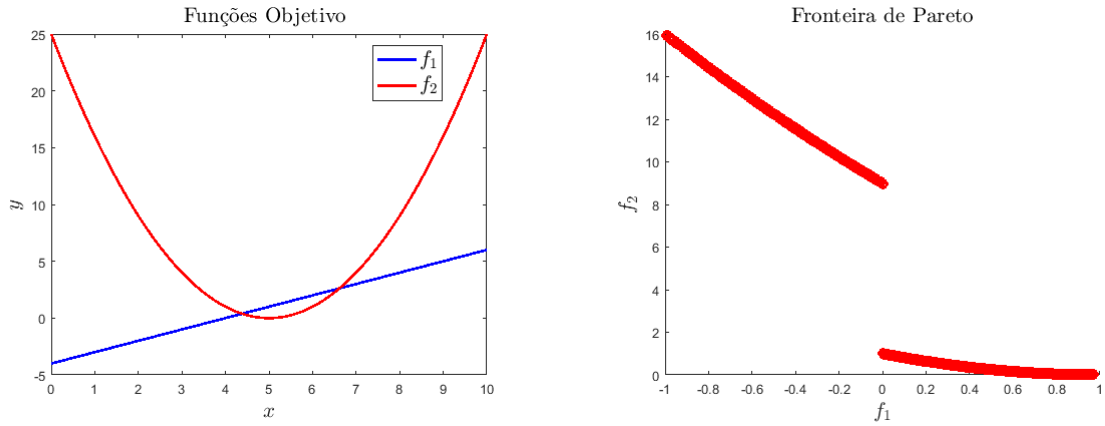
Problema 3.45 (MOP 14, [51]). Considere as funções objetivo $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ dadas por

$$f_1(x) = \begin{cases} -x, & x \leq 1 \\ -2 + x, & 1 < x \leq 3 \\ 4 - x, & 3 < x \leq 4 \\ -4 + x, & x > 4 \end{cases}$$

$$f_2(x) = (x - 5)^2,$$

e o domínio de busca $D = \{x \in \mathbb{R} \mid -5 \leq x \leq 10\}$.

A Figura 3.41 ilustra o Problema 3.45. A imagem da esquerda representa o gráfico das funções objetivo f_1 e f_2 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 20.000 rodadas do Algoritmo 3.6.



Fonte: O autor

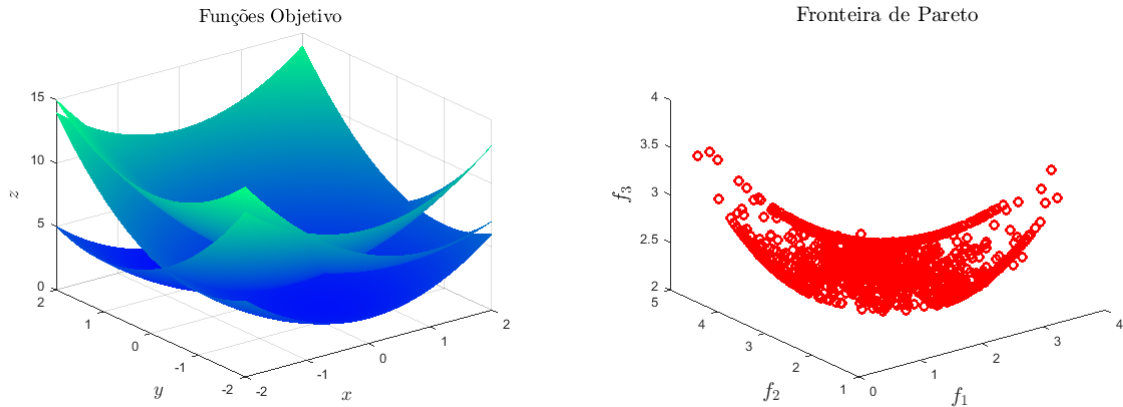
Figura 3.41: Problema 3.45

Problema 3.46 (MOP 16, [51]). Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + (x_2 - 1)^2, \\ f_2(x_1, x_2) &= x_1^2 + (x_2 + 1)^2 + 1, \\ f_3(x_1, x_2) &= (x_1 - 1)^2 + x_2^2 + 2, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -2 \leq x \leq 2\}$.

A Figura 3.42 ilustra o Problema 3.46. A imagem da esquerda representa o gráfico das funções objetivo f_1 , f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6.



Fonte: O autor

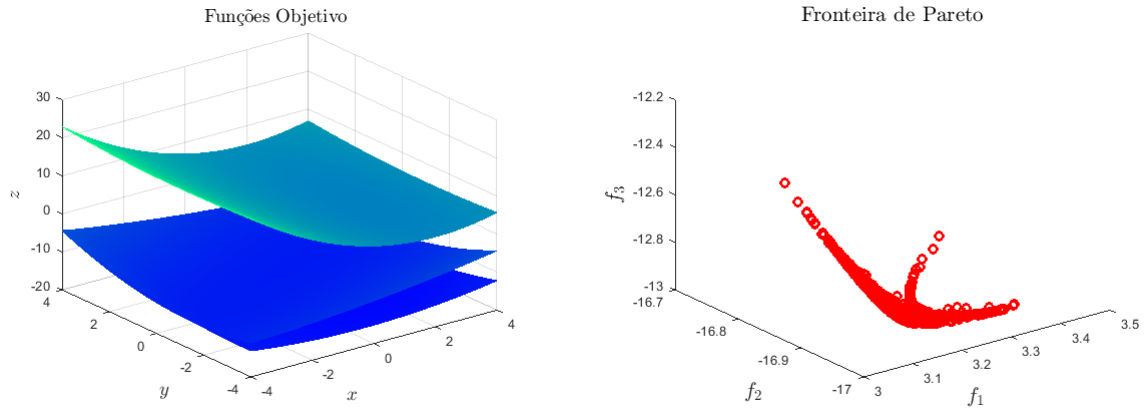
Figura 3.42: Problema 3.46

Problema 3.47 (MOP 17, [51]). Considere as funções objetivo $f_1, f_2, f_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ dadas por

$$\begin{aligned} f_1(x_1, x_2) &= \frac{1}{2}(x_1 - 2)^2 + \frac{1}{13}(x_2 + 1)^2 + 3, \\ f_2(x_1, x_2) &= \frac{1}{36}(x_1 + x_2 - 3)^2 + \frac{1}{8}(-x_1 + x_2 + 2)^2 - 17, \\ f_3(x_1, x_2) &= \frac{1}{175}(x_1 + 2x_2 - 1)^2 + \frac{1}{17}(2x_2 - x_1)^2 - 13, \end{aligned}$$

e o domínio de busca $D = \{x \in \mathbb{R}^2 \mid -4 \leq x \leq 4\}$.

A Figura 3.43 ilustra o Problema 3.47. A imagem da esquerda representa o gráfico das funções objetivo f_1 , f_2 e f_3 , enquanto à direita é representada a fronteira de Pareto a qual foi gerada para 2.000 rodadas do Algoritmo 3.6.



Fonte: O autor

Figura 3.43: Problema 3.47

Como estabelecido no Teorema 1.61, podemos determinar todas as soluções Pareto eficientes, sejam elas na parte convexa como na parte não convexa da fronteira. Contudo, algumas soluções fracamente eficientes (Pareto fracas), as quais não são Pareto eficientes, podem ser determinadas pela técnica de Chebyshev.

CONCLUSÃO

Nesta dissertação estudamos as caracterizações de otimalidade global propostas por Falk [16] e Hong e Quan [23] para problemas mono objetivos. Tais caracterizações são baseadas na Teoria da Medida e Integração. Vimos que tais condições são menos exigentes que as tradicionais, pois não exigem nem mesmo a continuidade da função objetivo. Algoritmos e testes numéricos também foram estudados.

Através das técnicas de escalarização obtivemos as devidas extensões para problemas de otimização multiobjetivo. Além dos resultados teóricos obtidos, também propomos um algoritmo para resolver problemas multiobjetivos, onde exigimos apenas a continuidade de cada função objetivo. Realizamos testes numéricos para ilustrar a eficácia do mesmo.

Está claro que estas condições de otimalidade via integração podem ser uma ferramenta poderosa para tratar de diversos problemas de otimização. Visto que abrange a grande maioria dos problemas de natureza prática que aparecem nas mais diversas áreas do conhecimento.

Uma possibilidade de trabalho futuro seria o estudo desta abordagem integral da otimalidade global, a uma classe de funções de penalidade descontínua propostas em [54, 59] para resolver problemas de otimização mono objetivo restritos. E estender estes resultados para problemas multiobjetivos.

Neste trabalho consideramos as técnicas de escalarização da soma ponderada e Chebyshev ponderado. Outra possibilidade é o estudo desta abordagem integral das condições de otimalidade global utilizando outras técnicas de escalarização, por exemplo as técnicas de escalarização: ε -restrição, cônica, Benson e Pascoletti–Serafini.

REFERÊNCIAS

- [1] AVRIEL, M.: *Nonlinear Programming: Analysis and Methods*. Prentice-Hall, New Jersey, 1976.
- [2] BECKER, L. C. *Introducción a la Optimización con Restricciones*. Universidad de Chile: 1984.
- [3] BERTSEKAS, Dimitri P. *Nonlinear Programming*. Belmont: Athena Scientific.
- [4] BOWMAN, V.J. *On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria objectives*. Lecture Notes in Economics and Mathematical Systems 130, 76–86, 1976.
- [5] BOYLE, P.: *Options: A monte carlo approach*. Journal of Financial Economics, v. 4, 1977.
- [6] BIGI, G. e CASTELLANI, M.: *Second order optimality conditions for differentiable multiobjective problems*. RAIRO – Operations Research 34, 411-426, (2000).
- [7] BIGI, G. e CASTELLANI, M.: *Uniqueness of KKT Multipliers in Multiobjective Optimization*. Applied Mathematics Letters, vol. 17. pag.1285-1290. ELSEVIER, 2004.
- [8] CRAVEN, B. D.: *Invex functions and constrained local minima*. Bulletin of the Australian Mathematical Society. Vol. 24, 357-366, 1981.
- [9] CHANKONG, V. e HAIMES, Y. : *Multiobjective Decision Making: Theory and Methodology*. North Holland, Series in System Science and Engineering, 1983.
- [10] CHARNES, A. e COOPER, W. W.: *Deterministic equivalents for optimizing and satisficing under chance constraints*. Operations research, vol. 11, no. 1, pp. 18–39, 1963.
- [11] CLARKE, F. H.: *Optimization and nonsmooth analysis*. Classics in applied mathematics. New York, Wiley, 1983.
- [12] CRANE, M.; HECQUET, D.; RUSKIN, H.J.: *Optimization and parallelisation strategies for Monte Carlo simulation of HIV infection*. Computers in Biology and Medicine, Elsevier Science, v. 37, 2007.
- [13] CUI, H.; WANG, C.; ZHENG Q.: *Optimality coditions and algorithms for integral global minimization*. Computers and mathematics with applications, vol. 52, pp. 55-64, 2006.
- [14] DEB, K.; THIELE L.; LAMANNS, M.; ZITZLER, E.: *Scalable Test Problems for Evolutionary Multiobjective Optimization*. In: Abraham A., Jain L., Goldberg R. (eds) Evolutionary Multiobjective Optimization. Advanced Information and Knowledge Processing. Springer, London, 2005.

- [15] EUSTÁQUIO, R. G. *Condições de otimalidade e de qualificação para problemas de programação não linear*. 136 f. Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Métodos Numéricos em Engenharia da Universidade Federal do Paraná, Curitiba, 2007.
- [16] FALK, J. E. *Technical Note – Conditions for Global Optimality in Nonlinear Programming*. Operations Research 21(1):337-340, 1973. <https://doi.org/10.1287/opre.21.1.337>
- [17] FILHO, A. A.: *Novas extensões de técnicas de escalarizações no problema de corte unidimensional inteiro multiobjetivo*. 204 f. Tese de doutorado apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas. Campinas, 2016.
- [18] GASS, S., Saaty, T.: *The computational algorithm for the parametric objective function*. Naval Res. Logist. Quart. 2, 39 (1955)
- [19] GAUVIN, J.: *A necessary and sufficient regularity condition to have bounded multipliers in non-convex programming*. Mathematical Programming, 12(1):136-138, 1977.
- [20] GALE, D., KUHN, H.W., TUCKER, A.W.: *Linear programming and the theory of games*. Koopmans, T.C. (ed.), Activity analysis of production and allocation, pp. 317–329. John Wiley, New York, 1951.
- [21] GOULD, F.J. e TOLLE, J.W.: *A necessary and sufficient qualification for constrained optimization*. SIAM Journal on Applied Mathematics, 20:164-172, 1969.
- [22] HANSON, M.A.: *On Sufficiency of the Kuhn-Tucker Conditions* - Journal of analysis and applications, Vol. 80, pp.545-550, 1981.
- [23] HONG, C. S.; QUAN, Z.: *Lecture Notes in Economics and Mathematical Systems – Integral Global Optimization, Theory, Implementation and Applications*. Springer-Verlag Berlin Heidelberg, 1988.
- [24] ISOTON, C.: *Condições necessárias e suficientes de otimalidade para problemas com um e com vários objetivos*. 65 f. Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Matemática Aplicada da Universidade Federal do Paraná – Curitiba, 2013.
- [25] IZMAILOV, A.; SOLODOV, M.: *Otimização: condições de otimalidade, elementos de análise convexa e de dualidade*. Vol.1, Rio de Janeiro:IMPA, 2005.
- [26] JOHN, F.: *Extremum Problems with Inequalities as Subsidiary Conditions*. In “Studies and Essays, Courant Anniversary Volume” (K. O. Friedrichs, O. E. Neugebauer e J. J. Stoker, editors), pp. 187-204. Wiley (Interscience), New York, 1948.

- [27] KALOS, M.H.; WHITLOCK, P.A.: *Monte Carlo Methods*. Courant Institute of Mathematical Sciences New York University. WILEY-VCH Verlag GmbH & Co. KGaA. Weinheim, 2004.
- [28] KARLIN, S.: *Mathematical methods and theory in games, programming, and economics*. Addison-Wesley, 1959.
- [29] KASIMBEYLI, R.; OZTURK, Z. K.; KASIMBEYLI, N.; YALCIN, G. D.; ERDEM, B. I.: *Comparison of Some Scalarization Methods in Multiobjective Optimization Comparison of Scalarization Methods*. Malaysian Mathematical Sciences Society and Penerbit Universiti Sains Malaysia, 2017.
- [30] KOSTREVA, M.M.; ZHENG, Q.: *Integral global optimization method for solution of nonlinear complementary problems*. Journal of global optimization, vol. 5, pp. 181-193, 1994.
- [31] KUHN, H.W.: *Nonlinear Programming: a historical view*. In “Nonlinear Programming” (R.W. Cottle e C.E. Lemke, editors) volume IX, pp. 1-26. SIAM-AMS Proceedings, American Mathematical Society, Providence, RI, New York, 1976.
- [32] KUHN, H. W.; TUCKER, A. W.: *Nonlinear programming*. In “Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability” (J. Neyman, ed.), pp. 481-492. University of California Press, Berkeley, California, 1951.
- [33] KYPARISIS, J.: *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*. Mathematical Programming 32, 242-246. North-Holland, 1985.
- [34] LEHRACH, H.; WIERLING, C.; KÜHN, A.; HACHE, H.; DASKALAKI, A.; MASCHKE-DUTZ, E.; PEYCHEVA, S.; LI, J.; HERWIG, R.: *Prediction in the face of uncertainty: A Monte Carlo-based approach for systems biology of cancer treatment*. Mutation Research\Genetic Toxicology and Environmental Mutagenesis, Elsevier Science, v. 746, 2012.
- [35] LAGUNA, M. e MARTÍ, R.: *Experimental Testing of Advanced Scatter Search Designs for Global Optimization of Multimodal Functions*. Journal of Global Optimization 33, 235–255. Springer, 2005. DOI 10.1007/s10898-004-1936-z.
- [36] MACIEL, M.C.; SANTOS, S.A.; SOTTOSANTO, G.N.: *On Second-Order Optimality Conditions for Vector Optimization*. Journal of optimization theory and applications 149, 332-351. Springer, 2011. DOI 10.1007/s10957-010-9793-z.
- [37] MANGASARIAN, O.L.: *Nonlinear Programming*, McGraw-Hill Book Company, New York, 1969.
- [38] MARTIN, D.H.: *The Essence of Inexactness* - Journal of optimization and applications: vol. 47, No. 1, September 1985.

- [39] McCormick, G.P.: *Second order conditions for constrained minima*. SIAM J. Appl.Math. 15(3), 641–652, 1967.
- [40] MIETTINEN, K.M.: *Nonlinear multiobjective optimization*. Kluwer Academic, Norwell, 1999.
- [41] MORETTI, A. C.; ROJAS-MEDAR, M. A.: *Condiciones suficientes de optimalidad en programación no lineal*. Campinas, 2000.
- [42] MOULIN, H.; SOULIÉ, F. F.: *La convexité dans les mathématiques de la décision*. HERMANN, Paris, 1979.
- [43] NORGAARD, R. L.: *A Monte Carlo simulation in insurance company portfolio management*. Journal of Risk and Insurance, John Wiley and Sons, v. 33, 1966.
- [44] OSUNA-GÓMEZ, R.; BEATO-MORENO, A.; RUFIAN-LIZANA, A.: *Generalized Convexity in Multiobjective Programming*. Journal of Mathematical Analysis and Applications 233, 205-220. Academic Press, 1999.
- [45] PARETO, V.: *Cours d'économie politique*. Rouge, Lausanni, Suíça, 1896.
- [46] RAUN, D. L.: *The application of Monte Carlo analysis to an inventory problem*. The Accounting Review, American Accounting Association, v. 38, 1963.
- [47] RIBEIRO, A. A.; KARAS, E. W.: *Otimização Contínua: aspectos teóricos e computacionais*. São Paulo: Cengage Learning, 2013.
- [48] RIBEIRO, F. B.; MOLINA, E. C.: *Uma introdução ao método de Monte Carlo*. Notas em Matemática Aplicada; v. 86. São Carlos, SP: SBMAC, 2017.
- [49] SCHAFFER, J.D.: *Some Experiments in Machine Learning Using Vector Evaluated Genetic Algorithms*. PhD thesis, Nashville, TN: Vanderbilt University, 1984.
- [50] SETZU, A.; ECCA, S.; MARCHESI, M.: *Modeling and Simulation of an Artificial Stock Option Market*. Computational Economics, Springer US, v. 32, 2008.
- [51] VELDHUIZEN, D.A.V.: *Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations*. 270 f .Dissertação apresentada à Faculdade de Pós-Graduação em Engenharia do Instituto de Tecnologia da Força Aérea da Air University, no cumprimento parcial dos requisitos para o grau de doutor em filosofia, EUA, 1999.
- [52] WACHSMUTH, G.: *On LICQ and the uniqueness of Lagrange multipliers*. Operations Research Letters, 41(1):78 - 80, 2013.

- [53] WANG, S.: *Second-order necessary and sufficient conditions in multiobjective programming*. Numerical Functional Analysis and Optimization, vol. 12, 237-252 (1991).
- [54] WU, B.; CUI, H.; ZHENG, Q.: *Integral global minimization of constrained problems with discontinuous penalty functions*. Journal of Shanghai University, vol. 9, No. 5, pp. 385-390, 2005.
- [55] ZHENG, Q.: *Robust analysis and global optimization*, Computers Math. Applic. Vol. 21, No. 6/7, pp. 17-24, Great Britain, 1991.
- [56] ZHENG, Q.; ZHUANG, D.: *Integral global minimization: algorithms, implementations and numerical tests*, Journal of Global Optimization, vol. 7, pp. 421-454, Netherlands, 1995.
- [57] ZHENG, Q.: *Optimality conditions for global optimization (I)*. Acta Mathematicae applicatae sinica, vol. 2, No.1, pp. 66-78, 1985.
- [58] ZHENG, Q.: *Optimality conditions for global optimization (II)*. Acta Mathematicae applicatae sinica, vol. 2, No.1, pp. 118-132, 1985.
- [59] ZHENG, Q.; ZHANG, L.: *Global minimization of constrained problems with discontinuous penalty functions*. Computers and mathematics with applications, vol. 37, pp. 41-58, 1999.
- [60] ZHENG, Q.: *Robust analysis and global minimization of a class of discontinuous functions (I)*. Acta mathematicae applicatae sinica, vol. 6, No. 3, pp. 205-223, 1990.
- [61] ZHENG, Q.: *Robust analysis and global minimization of a class of discontinuous functions (II)*. Acta mathematicae applicatae sinica, vol. 6, No. 4, pp. 317-337, 1990.