



Universidade do Estado do Rio de Janeiro
Centro de Tecnologia e Ciências
Instituto Politécnico

Gustavo Barbosa Libotte

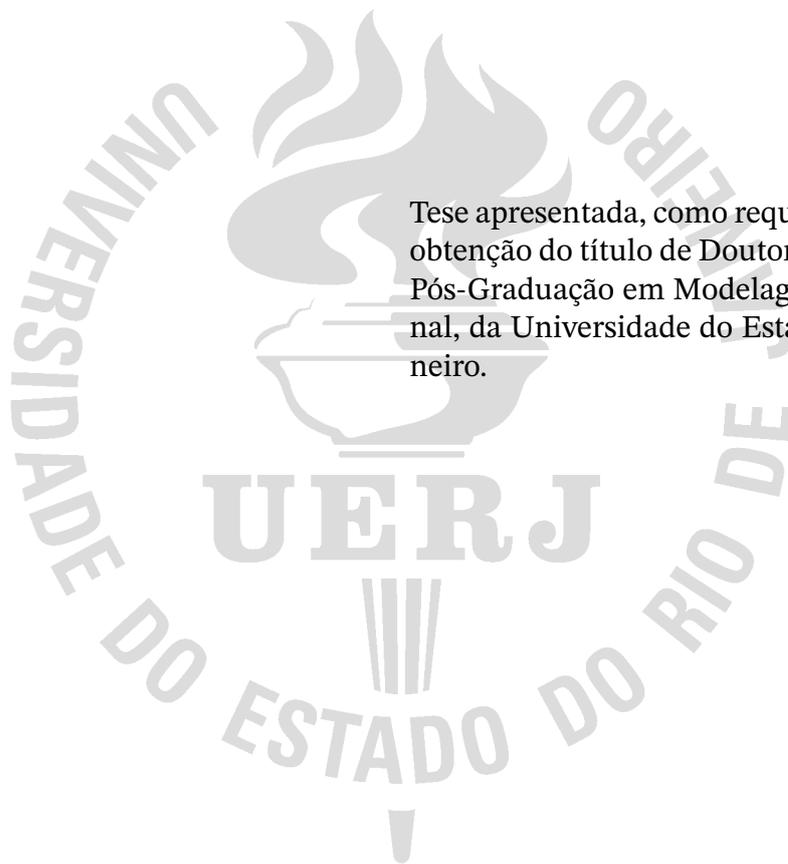
**Otimização multi-objetivo de sistemas de engenharia na presença
de incerteza**

Nova Friburgo

2020

Gustavo Barbosa Libotte

Otimização multi-objetivo de sistemas de engenharia na presença de incerteza



Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Modelagem Computacional, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Ph.D. Francisco Duarte Moura Neto

Orientador: Prof. Dr. Fran Sérgio Lobato

Nova Friburgo

2020

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CTC/E

L696 Libotte, Gustavo Barbosa.
Otimização multi-objetivo de sistemas de engenharia na
presença de incerteza / Gustavo Barbosa Libotte. – 2020.
218 f. : il

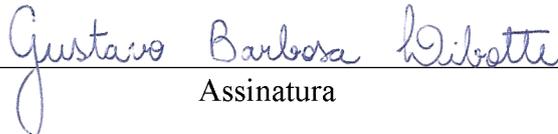
Orientadores: Francisco Duarte Moura Neto e Fran Sérgio Lobato.
Tese (doutorado) – Universidade do Estado do Rio de Janeiro,
Instituto Politécnico.

1. Otimização matemática – Teses. 2. Algoritmo numérico –
Teses. 3. Análise matemática – Teses. I. Moura Neto, Francisco
Duarte. II. Lobato, Fran Sérgio. III. Universidade do Estado do Rio de
Janeiro. Instituto Politécnico. IV. Título.

CDU 519.863

Bibliotecária Cleide Sancho CRB7/5843

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese,
desde que citada a fonte.


Assinatura

16/10/2020

Data

Otimização Multi-objetivo de Sistemas de Engenharia na Presença de Incerteza

Tese apresentada como requisito parcial para obtenção do título de Doutor ao Programa de Pós-Graduação em Modelagem Computacional do Instituto Politécnico, da Universidade do Estado do Rio de Janeiro.

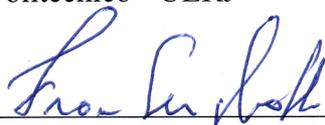
Aprovada em 17 de fevereiro de 2020.

Banca examinadora:



Prof. Francisco Duarte Moura Neto, Ph.D. (Orientador)

Instituto Politécnico - UERJ



Prof. Fran Sérgio Lobato, D.Sc. (Orientador)

Universidade Federal de Uberlândia



Prof. Roberto Pinheiro Domingos, D.Sc.

Instituto Politécnico - UERJ



Prof. Bernardo Sotto-Maior Peralva, D.Sc.

Instituto Politécnico - UERJ



Prof. Valder Steffen Junior, D.Sc.

Universidade Federal de Uberlândia



Prof. Gustavo Mendes Platt, D.Sc.

Universidade Federal do Rio Grande

DEDICATÓRIA

Ao meu pai (*in memoriam*).

AGRADECIMENTOS

Agradeço à minha esposa pelo apoio incondicional, mesmo nos momentos mais difíceis, por sempre me motivar a tentar ir além das expectativas e me lembrar a cada dia do meu valor e das minhas qualidades. Sei que compreende o esforço necessário para chegar até aqui e fico grato por valorizar cada conquista nossa. Também agradeço por ajudar a me manter firme em busca dos meus sonhos, que também são seus, me proporcionando a segurança e tranquilidade necessárias que a trajetória exige, mesmo que isso signifique a renúncia de vários momentos juntos.

Agradeço à minha mãe por ter me proporcionado a oportunidade de me dedicar aos estudos e por sempre ter acompanhado de perto a minha trajetória e evolução. Desde criança a educação foi colocada em primeiro lugar e, certamente, a base construída foi essencial para que eu pudesse alcançar esse objetivo. Além disso, muito obrigado pelos conselhos, por todo esforço para proporcionar sempre o melhor e por estar sempre por perto para nos ajudar e encorajar. Também agradeço ao meu irmão, que é parte de tudo isso, pela amizade e camaradagem de sempre, me ajudando sempre que preciso e estando presente em todos os momentos.

Agradeço ao orientador Francisco Moura Neto por ser tão presente e atuante, sempre paciente com as minhas dificuldades e disposto a me ajudar a ser um aluno melhor. Cada encontro semanal foi muito enriquecedor, tanto do ponto de vista acadêmico quanto pessoal. Muito obrigado pelo incentivo e encorajamento, além do empenho a cada página deste trabalho, sempre sem medir esforços para que pudesse ficar cada vez melhor. Certamente você contribuiu muito para o meu amadurecimento enquanto aluno.

Agradeço ao orientador Fran Sérgio Lobato por ter aceitado o desafio de me orientar mesmo de tão longe. A sua prestatividade e cordialidade são exemplares. Muito obrigado por ter me proporcionado uma orientação tão dedicada e participativa, a ponto de tornar imperceptível qualquer dificuldade que poderia ter sido ocasionada pela distância. Também não posso deixar de agradecer a amizade construída durante este período, além de todo estímulo para alcançar os meus objetivos.

Agradeço ao (sempre) orientador Gustavo Platt, que me acolheu desde a graduação, me ensinou muito do que sei, me proporcionou diversas oportunidades de aprendizado e foi determinante nas minhas escolhas. É gratificante poder continuar a produzir diversos projetos juntos, mesmo após a sua mudança. Também sou muito lisonjeado pelo seu constante apoio ao meu desejo de seguir a carreira acadêmica e por tudo que fez e tem feito para me ajudar a atingir esta meta.

Agradeço ao IPRJ, na figura dos professores e funcionários que, em quase 10 anos, foram essenciais para a minha formação profissional, moral, ética e como cidadão. Seria impossível citar cada um, mas muitos contribuíram, direta e indiretamente, para que eu fosse capaz de fechar esse ciclo.

A ciência nunca resolve um problema sem criar pelo menos outros dez.

George Bernard Shaw

RESUMO

LIBOTTE, G. B. *Otimização multi-objetivo de sistemas de engenharia na presença de incerteza*. 2020. 218 f. Tese (Doutorado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2020.

Modelos matemáticos são capazes de simular diversos eventos, sob diferentes condições, possibilitando obter um panorama antecipado do sistema a ser implementado na prática, com desperdício reduzido de recursos e em menor tempo. Na otimização de projetos, estes modelos desempenham papel fundamental, pois permitem que sejam obtidos parâmetros e atributos capazes de potencializar o desempenho do produto, propiciando a redução de custos e de tempo de operação. Estas possibilidades estão sujeitas a diversos fatores, dentre eles a modelagem computacional mais precisa possível das características inerentes ao sistema. Em geral, modelos computacionais capazes de descrever fenômenos físicos com precisão satisfatória consideram possíveis incertezas em suas formulações matemáticas. A inclusão de incertezas ao modelo computacional também afeta a viabilidade dos resultados e sua implementação prática. Neste trabalho, são consideradas duas abordagens distintas capazes de quantificar incertezas durante a otimização de modelos matemáticos: na primeira, a otimização robusta, é avaliada a sensibilidade das variáveis de decisão em relação a perturbações ocasionadas por fatores externos. Soluções robustas representam a redução da possibilidade de desvio em relação a eventuais alterações no sistema. O propósito da segunda abordagem, a otimização baseada em confiabilidade, é mensurar a probabilidade de falha do sistema e obter parâmetros para o modelo capazes de assegurar um nível de confiabilidade estabelecido. Neste contexto, o objetivo fundamental é formular um problema de otimização multi-objetivo capaz de considerar a otimização robusta e a otimização baseada em confiabilidade em conjunto, para obtenção de soluções minimamente sensíveis a perturbações externas e que satisfaçam níveis de confiabilidade prescritos. Neste trabalho, também é proposto um algoritmo para análise de confiabilidade inversa, apropriado para ser empregado durante o procedimento de otimização baseada em confiabilidade. O algoritmo adota um conceito de cálculo de tamanho de passo de segunda ordem adaptativo, que possibilita notável redução do uso de recursos computacionais para obtenção de soluções confiáveis.

Palavras-chave: Otimização multi-objetivo. Otimização robusta. Otimização baseada em confiabilidade. Análise de confiabilidade inversa.

ABSTRACT

LIBOTTE, G. B. *Multi-objective optimization of engineering systems in the presence of uncertainty*. 2020. 218 f. Tese (Doutorado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2020.

Mathematical models are applied to simulate various events under different conditions, enabling an early overview of the system to be implemented in practice, with reduced waste of resources and in less time. In project optimization, these models play a fundamental role, as they allow obtaining parameters and attributes capable of enhancing product performance, reducing costs and operating time. These possibilities are subject to several factors, including a most accurate computational modeling of the inherent characteristics of the system. In general, computational models capable of describing physical phenomena with satisfactory accuracy consider possible uncertainties in their mathematical formulations. The inclusion of uncertainties in the computational model also affects the feasibility of the results and their practical implementation. In this work, two different approaches capable of quantifying uncertainties during the optimization of mathematical models are considered: in the first, robust optimization, the sensitivity of decision variables in relation to perturbations caused by external factors is assessed. Robust solutions reduce the possibility of deviation from possible system changes. The purpose of the second approach, reliability-based optimization, is to measure the probability of failure of a system and to obtain model parameters that can ensure an established level of reliability. In this context, the fundamental objective is to formulate a multi-objective optimization problem capable of considering together robust optimization and reliability-based optimization, to obtain solutions that are least sensitive to external noise and that satisfy prescribed reliability levels. In this work, it is also proposed an algorithm for inverse reliability analysis, suitable to be employed during the reliability-based optimization procedure. The algorithm adopts an adaptive second order step length calculation concept, which enables a remarkable reduction in the use of computational resources to obtain reliable solutions.

Keywords: Multi-objective optimization. Robust optimization. Reliability-based optimization. Inverse reliability analysis.

LISTA DE FIGURAS

Figura 1	- Exemplo de variação geométrica no encaixe de peças.	19
Figura 2	- Relação entre o espaço das variáveis de decisão e o espaço objetivo em um problema de otimização.	28
Figura 3	- Exemplo de otimização multi-objetivo de um problema genérico, que relaciona custo de mão de obra e matéria-prima.	30
Figura 4	- Tipos de conjuntos de Pareto, dependendo da finalidade de cada um dos objetivos.	32
Figura 5	- Valores ótimos locais de Pareto, no interior de $B(\mathbf{f}(\mathbf{x}^*), \delta)$	33
Figura 6	- Exemplos de funções-objetivo convexas.	34
Figura 7	- Exemplo da relação de convergência e diversidade de soluções em um problema de $\min f_1(\mathbf{x})$ e $\min f_2(\mathbf{x})$	35
Figura 8	- Aplicação do Método da Soma Ponderada sob a perspectiva de dois tipos de problemas distintos.	39
Figura 9	- Aplicação do Método da Negociação para diferentes valores de ε	40
Figura 10	- Método de Otimização de Compromisso aplicado a um problema genérico, com diferentes valores do parâmetro p	42
Figura 11	- Esquema geométrico do procedimento de aglomeração de soluções para problema com dois objetivos.	51
Figura 12	- Funcionamento do processo evolutivo do NSGA-II.	53
Figura 13	- Fluxograma do processo iterativo simplificado do SA-MODE.	58
Figura 14	- Distribuição de amostras aleatórias na vizinhança dos pontos mínimos do problema, com $\delta = 0,1$	77
Figura 15	- Noção fundamental da avaliação de riscos.	81
Figura 16	- Esquema genérico do cálculo da probabilidade de sucesso através da integração da função densidade de probabilidade em $g(\mathbf{X}) > 0$	84
Figura 17	- Esquema conceitual do Método de Amostragem de Importância de Harbitz.	88
Figura 18	- Transformação da variável aleatória X para o espaço normal padrão U	95
Figura 19	- Obtenção do ponto mais provável de falha (\mathbf{u}^*) no FORM.	97
Figura 20	- Fluxograma do Método de Hasofer-Lind-Rackwitz-Fiessler.	102
Figura 21	- Erros relacionados às aproximações de primeira e segunda ordens.	105
Figura 22	- Procedimento de rotação dos eixos no SORM, para um problema no plano.	107
Figura 23	- Função não-decrescente do índice de probabilidade generalizado $\beta_g(\gamma) = -\Phi^{-1}(F_x(\gamma))$	110
Figura 24	- Busca pelo ponto mais provável de falha na abordagem PMA.	111

Figura 25	- Fluxograma do Método HMV, destacando as principais etapas do AMV e CMV.	114
Figura 26	- Comportamento do método SLA no problema fornecido pela Eq. (70). Nesta simulação, $\beta_t = 2,5$	120
Figura 27	- Obtenção do minimizador de uma função convexa.	123
Figura 28	- Extensão do limite superior ($\bar{t}^{(k)}$) para o intervalo de busca, quando a aproximação polinomial de segunda ordem é não-convexa.	124
Figura 29	- Fluxograma do Método de Tamanho de Passo de Segunda Ordem Adaptativo.	127
Figura 30	- Comportamento do método ASOSL no problema dado pela Eq. (70). . .	129
Figura 31	- Histórico de iterações do método ASOSL na solução do problema $G_2(\mathbf{U})$, com $\delta_\eta = 1$	132
Figura 32	- Iterações de SLA e ASOSL na resolução do problema $G_3(\mathbf{U})$	134
Figura 33	- Esquema de sistema de vibração dinâmico com dois graus de liberdade.	137
Figura 34	- Valor da função de desempenho em cada iteração k para cada método. . .	139
Figura 35	- Resultados obtidos pelo LDMV para diferentes combinações dos parâmetros de controle ρ_{LDMV} e δ_{LDMV}	140
Figura 36	- Relação entre resultados obtidos e quantidade de avaliações da função de desempenho do método DCC para diferentes valores do parâmetro c_{DCC}	141
Figura 37	- Resultados obtidos pelo ASOSL para diferentes combinações dos parâmetros de controle α_b e s_b , com $\delta_\eta = 1$	142
Figura 38	- Fluxograma da metodologia geral para incorporação de robustez e confiabilidade em problemas de otimização multiobjetivo.	148
Figura 39	- Soluções do problema da Eq. (88) para diferentes níveis de robustez das variáveis de projeto.	150
Figura 40	- Otimização baseada em confiabilidade do problema dado pela Eq. (89), para $\beta_t = 3$	152
Figura 41	- Conjuntos de valores ótimos obtidos para diferentes níveis de robustez, no problema dado pela Eq. (91).	153
Figura 42	- Rede de trocadores de calor com três estágios.	154
Figura 43	- Soluções do problema de otimização multi-objetivo de uma rede de trocadores de calor com incertezas, definido pela Eq. (96), para diferentes níveis de robustez.	158
Figura 44	- Esquema do projeto de rede de reatores.	160
Figura 45	- Conjuntos de resultados ótimos obtidos pelo MODE para o problema dado pela Eq. (100), e pela DE para um problema mono-objetivo equivalente, variando-se o índice de confiabilidade a cada execução.	163

Figura 46	- Conjunto de resultados do problema de otimização multi-objetivo de uma rede de reatores com incertezas, dado pela Eq. (101), para diferentes níveis de robustez.	164
Figura 47	- Conjunto de resultados do problema de otimização multi-objetivo de uma rede de reatores com incertezas, dado pela Eq. (101), utilizando as técnicas de Diferença Normalizada e de Penalização, com níveis de robustez em 5% e 10%.	167
Figura 48	- Conjunto de resultados do problema de otimização multi-objetivo da mistura de catalisadores com incertezas, dado pela Eq. (104), para diferentes níveis de robustez.	170
Figura 49	- Perfis de controle obtidos para dois níveis de robustez distintos, quando $\beta_t \approx 1,6$	171
Figura 50	- Conjunto de resultados do primeiro estudo de caso do problema dado pela Eq. (113), para diferentes níveis de robustez.	180
Figura 51	- Conjunto de resultados do segundo estudo de caso do problema dado pela Eq. (113), para diferentes níveis de robustez.	181
Figura 52	- Relação entre a função densidade de probabilidade e a função de distribuição acumulada associada.	206
Figura 53	- Mapeamento de experimento estocástico com variável aleatória bidimensional.	209
Figura 54	- Área sob a superfície da função densidade de probabilidade conjunta e a representação da respectiva probabilidade $P [X \leq x, Y \leq y]$ na função de distribuição acumulada conjunta.	210
Figura 55	- Função densidade de probabilidade e função de distribuição acumulada de uma variável aleatória X com distribuição normal.	215
Figura 56	- Função densidade de probabilidade da distribuição log-normal, com $\mu_x = 0,5$ e diferentes valores para σ_x	218

LISTA DE TABELAS

Tabela 1	- Análise de robustez das soluções $x = -1$ e $x = 1$ do problema dado pela Eq. (29) utilizando diferentes métodos.	78
Tabela 2	- Resultados do processo iterativo do método HMV para a função de desempenho fornecida pela Eq. (70).	118
Tabela 3	- Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_1(\mathbf{U})$ derivado da Eq. (81), para diferentes técnicas. . . .	130
Tabela 4	- Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_2(\mathbf{U})$ derivado da Eq. (82), para diferentes técnicas. . . .	131
Tabela 5	- Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_3(\mathbf{U})$ derivado da Eq. (83), para diferentes técnicas. . . .	133
Tabela 6	- Convergência do ASOSL para diferentes valores de δ_η , com $\alpha_b = 10^{-4}$ e $s_b = 0,5$	135
Tabela 7	- Propriedades estatísticas das variáveis aleatórias do problema de vibração com dois graus de liberdade.	137
Tabela 8	- Resultados estatísticos referentes ao custo computacional e à sensibilidade dos parâmetros de controle relacionados aos métodos ASOSL, LDMV e DCC, no problema do sistema dinâmico com dois graus de liberdade.	142
Tabela 9	- Propriedades estatísticas das variáveis aleatórias do problema da rede de trocadores de calor.	156
Tabela 10	- Resultados ótimos do problema de rede de trocadores de calor, considerando diferentes níveis de compromisso entre os objetivos.	159
Tabela 11	- Soluções ótimas do problema de projeto de rede de reatores, com a indicação do tipo de solução ótima (local ou global) de cada ponto.	161
Tabela 12	- Propriedades estatísticas das variáveis aleatórias do problema da rede de reatores.	162
Tabela 13	- Medidas de qualidade da curva ajustada em relação à capacidade de estimação dos valores ótimos computados para cada nível de robustez na Fig. 46.	165
Tabela 14	- Valores das propriedades físicas para ambos os estudos de caso do problema de custo de operação de um trocador de calor de casco e tubo. . .	178

LISTA DE ALGORITMOS

Algoritmo	1 - Ordenamento das frentes de Pareto não-dominadas.	50
Algoritmo	2 - Distância de aglomeração de soluções.	52
Algoritmo	3 - Laço de repetição principal do NSGA-II.	54

LISTA DE ABREVIATURAS E SIGLAS

AFOSM	<i>Advanced First-order Second-moment</i> (Primeira Ordem Segundo Momento Avançado)
ASOSL	<i>Adaptive Second Order Step Length</i> (Tamanho de Passo de Segunda Ordem Auto-adaptativo)
AMV	<i>Advanced Mean-Value</i> (Valor Médio Avançado)
CMV	<i>Conjugate Mean Value</i> (Valor Médio Conjugado)
CSTR	<i>Continuous Stirred Tank Reactor</i> (Reator Tanque Agitado Contínuo)
DE	<i>Differential Evolution</i> (Evolução Diferencial)
FOSM	<i>First-order Second-moment</i> (Primeira Ordem Segundo Momento)
FORM	<i>First Order Reliability Method</i> (Método de Confiabilidade de Primeira Ordem)
GA	<i>Genetic Algorithm</i> (Algoritmo Genético)
HL-RF	Hasofer-Lind-Rackwitz-Fiessler
HMV	<i>Hybrid Mean Value</i> (Valor Médio Híbrido)
IRA	<i>Inverse Reliability Analysis</i> (Análise de Confiabilidade Inversa)
MODE	<i>Multi-objective Differential Evolution</i> (Evolução Diferencial Multi-objetivo)
MVFOSM	<i>Mean Value First-order Second-moment</i> (Primeira Ordem Segundo Momento de Valores Médios)
PMA	<i>Performance Measure Approach</i> (Abordagem de Medida de Desempenho)
RIA	<i>Reliability Index Approach</i> (Abordagem de Índice de Confiabilidade)
SLA	<i>Step Length Adjustment Method</i> (Método de Ajuste de Comprimento de Passo)
SA-MODE	<i>Self-adaptive Multi-objective Optimization Differential Evolution</i> (Evolução Diferencial Auto-adaptativa para Otimização Multi-objetivo)
SORM	<i>Second Order Reliability Method</i> (Método de Confiabilidade de Segunda Ordem)

LISTA DE SÍMBOLOS

Símbolos Romanos

d	Direção de busca
ε	Medida de erro
f_x	Função densidade de probabilidade da variável aleatória X
F_x	Função de distribuição acumulada da variável aleatória X
\mathcal{F}	Função-objetivo modificada
g	Função de desempenho no espaço das variáveis de decisão
G	Função de desempenho no espaço normal padrão
\mathcal{LN}	Distribuição Log-normal
\mathcal{N}	Distribuição Normal
p_f	Probabilidade de falha
u^*	Ponto mais provável de falha

Letras Gregas

β	Índice de confiabilidade
δ	Nível de ruído
δ_η	Extensão do tamanho de passo
η	Limiar de robustez
μ	Média de uma variável aleatória
ϕ	Função densidade de probabilidade da distribuição normal padrão
Φ	Função de distribuição acumulada da distribuição normal padrão
σ	Desvio-padrão de uma variável aleatória
τ	Tamanho de passo

SUMÁRIO

	INTRODUÇÃO	18
1	CONCEITOS FUNDAMENTAIS DE OTIMIZAÇÃO	24
1.1	Formulação do Problema de Otimização Multi-objetivo	26
1.1.1	<u>Ótimo de Pareto</u>	28
2	METODOLOGIAS DE OTIMIZAÇÃO MULTI-OBJETIVO	36
2.1	Métodos <i>a Posteriori</i>	37
2.1.1	<u>Método da Soma Ponderada</u>	37
2.1.2	<u>Método da Negociação</u>	39
2.1.3	<u>Método de Otimização de Compromisso</u>	41
2.2	Métodos <i>a Priori</i>	42
2.2.1	<u>Método de Otimização Hierárquica</u>	42
2.2.2	<u>Método da Programação de Metas</u>	43
2.3	Métodos Progressivos	44
2.3.1	<u>Método de Fandel</u>	44
2.3.2	<u>Método STEP</u>	46
2.4	Métodos Baseados em Metaheurísticas	47
2.4.1	<u>Algoritmo Genético de Ordenamento Não-dominado</u>	49
2.4.1.1	Construção de Frentes de Pareto Ordenadas	49
2.4.1.2	Cálculo das Distâncias de Aglomeração	50
2.4.1.3	Procedimento Principal	52
2.4.2	<u>Evolução Diferencial Auto-adaptativa para Otimização Multi-objetivo</u>	55
2.4.2.1	Fundamentação do Algoritmo	55
2.4.2.2	Auto-adaptação dos Parâmetros de Controle Evolutivos	57
3	OTIMIZAÇÃO BASEADA EM ROBUSTEZ	62
3.1	Método de Taguchi	62
3.2	Método de Superfície de Resposta para Análise de Robustez	63
3.3	Otimização Robusta e Algoritmos Genéticos	66
3.3.1	<u>Técnicas de Análise de Robustez com Abordagem Genética</u>	69
3.3.1.1	Aspectos Essenciais do Teorema Fundamental dos Algoritmos Genéticos	69
3.3.1.2	Média Efetiva	72
3.3.1.3	Diferença Normalizada	74
3.3.1.4	Técnica de Análise de Robustez com Penalização	75
3.3.1.5	Um Primeiro Exemplo	76
4	OTIMIZAÇÃO BASEADA EM CONFIABILIDADE	79
4.1	Noções de Análise de Confiabilidade	80
4.1.1	<u>Noção Preliminar de Métodos Probabilísticos</u>	80

4.1.2	<u>Conceitos Fundamentais</u>	82
4.2	Métodos de Análise de Confiabilidade	84
4.2.1	<u>Método Local</u>	85
4.2.1.1	Método de Superfície de Resposta	85
4.2.2	<u>Métodos de Amostragem</u>	86
4.2.2.1	Método de Simulação de Monte Carlo	86
4.2.2.2	Amostragem de Importância	87
4.2.2.3	Amostragem por Hipercubo Latino	88
4.2.3	<u>Métodos Globais</u>	89
5	MÉTODOS DE CONFIABILIDADE DE PRIMEIRA E SEGUNDA ORDENS	91
5.1	Método de Confiabilidade de Primeira Ordem	91
5.1.1	<u>Mean Value First-order Second-moment</u>	91
5.1.2	<u>Advanced First-order Second Moment</u>	93
5.1.2.1	Transformação de Variáveis	94
5.1.2.2	Simplificação do Integrand e dos Limites de Integração	95
5.1.2.3	Obtenção do Ponto Mais Provável de Falha	97
5.1.2.4	Algoritmo do Método de Rackwitz	99
5.1.2.5	Descrição do Método de Hasofer-Lind-Rackwitz-Fiessler	100
5.1.2.6	Aplicação do FORM em Variáveis Aleatórias com Outros Tipos de Distribuição	102
5.2	Método de Confiabilidade de Segunda Ordem	104
5.3	Análise de Confiabilidade Inversa	109
5.3.1	<u>Visão Geral de Métodos para Análise de Confiabilidade Inversa</u>	112
6	MÉTODO DE TAMANHO DE PASSO DE SEGUNDA ORDEM ADAPTATIVO	117
6.1	Motivação	117
6.2	Cálculo do Tamanho de Passo Ótimo	121
6.3	Algoritmo do Método de Tamanho de Passo de Segunda Ordem Adaptativo	125
6.4	Resultados	126
6.4.1	<u>Uma Primeira Comparação Entre SLA e ASOSL</u>	127
6.4.2	<u>Análise Comparativa de Desempenho</u>	128
6.4.2.1	Problema 1—Função Convexa	130
6.4.2.2	Problema 2—Função Côncava	131
6.4.2.3	Problema 3—Função Altamente Não-linear	132
6.4.3	<u>Análise de Sensibilidade do ASOSL</u>	133
6.4.4	<u>Problema de Engenharia</u>	136
6.4.4.1	Sistema Dinâmico com Dois Graus de Liberdade	136

6.4.4.2	Comparação de Custo Computacional	138
6.4.4.3	Análise de Sensibilidade dos Parâmetros de Controle	138
7	OTIMIZAÇÃO MULTI-OBJETIVO BASEADA EM ROBUSTEZ E CON-	
	FIABILIDADE	144
7.1	Definição do Problema de Otimização com Incertezas	144
7.2	Avaliação de Incertezas na Otimização Multi-objetivo	145
7.2.1	<u>Custo Computacional</u>	147
7.3	Avaliação da Formulação Proposta	147
7.3.1	<u>Problema Matemático de <i>Benchmark</i></u>	149
7.3.2	<u>Problemas de Engenharia Química</u>	154
7.3.2.1	Projeto de Rede de Trocadores de Calor	154
7.3.2.2	Projeto de Rede de Reatores	159
7.3.2.3	Problema de Mistura de Catalisadores	166
7.3.2.4	Custo Anualizado Total de um Trocador de Calor de Casco e Tubo	171
	CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS	182
	REFERÊNCIAS	185
	APÊNDICE A – Técnica de <i>Backtracking</i>	202
	APÊNDICE B – Fundamentos da Teoria de Probabilidades	203

INTRODUÇÃO

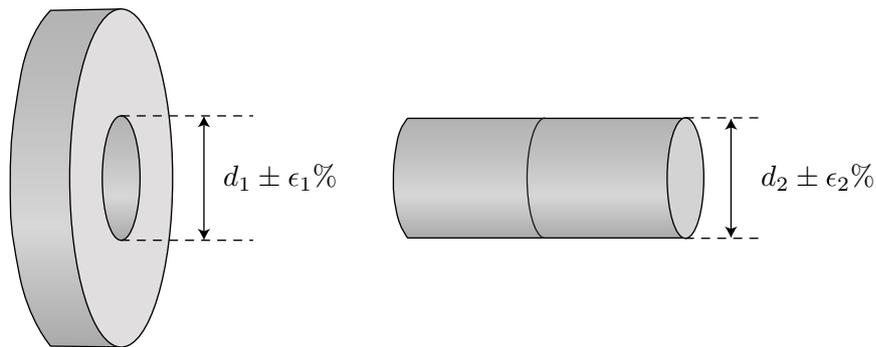
De forma crescente, a produção de novas tecnologias e criação de soluções de engenharia têm sido influenciadas pelas simulações computacionais. A formulação matemática de problemas físicos é essencial ao desenvolvimento teórico de modelos computacionais, uma vez que possibilita a análise de diversos fatores que podem influenciar no resultado final do produto, antes da sua efetiva produção e implantação. Através desses modelos, as simulações computacionais são capazes de reproduzir diversas combinações de situações, sob condições e fatores específicos, propiciando uma análise prévia de possíveis adversidades e reduzindo o tempo necessário para o projeto. Na formulação matemática, considera-se uma série de modelos e cálculos aproximados, devido à inviabilidade de se computar inúmeros problemas de interesse científico com relativa exatidão, fornecendo uma maneira de se processar os dados que são adequados para o uso de técnicas de otimização.

A otimização de projetos tornou-se uma ferramenta fundamental em diversos setores e desenvolvimento de produtos, através da tentativa constante de potencializar o desempenho de determinado sistema ou produto, visando a redução de custos. Para que isso ocorra, os modelos matemáticos devem ser coerentes com a realidade, descrevendo com exatidão suficiente as características físicas do problema. Resultados adequados também dependem da formulação apropriada dos problemas de otimização, que podem ter grande influência na tomada de decisões. Através da obtenção de resultados ótimos, que sejam fundamentalmente coerentes, torna-se possível aumentar o desempenho de sistemas e procedimentos, favoráveis à redução de custos.

Na prática, diversos valores considerados em modelos matemáticos estão sujeitos a variações e incertezas. Além das possíveis inconsistências existentes durante a modelagem computacional, o resultado do problema de otimização sofre alteração por conta do descuido em se desconsiderar o real efeito de determinados parâmetros, a fim de se reduzir o esforço computacional através da atribuição de valores determinísticos (JURECKA, 2007). Sempre que o modelo envolve parâmetros estocásticos, e mesmo se as variações são pequenas, podem originar grandes flutuações nos resultados ótimos. Além disso, a variabilidade relacionada às tolerâncias de fabricação e imprecisão geométrica podem afetar o projeto na prática (CURSI; SAMPAIO, 2015). A Fig. 1 ilustra um exemplo de variação geométrica de peças.

A incorporação das incertezas deve ser levada em conta no processo de modelagem matemática de produtos ou processos, com o objetivo de se obter soluções viáveis, nas quais sejam considerados eventuais erros e variações nas condições de operação. Uma abordagem habitual consiste na aplicação de fatores de segurança, através da utilização de coeficientes multiplicativos capazes de modificar o resultado do problema de otimização, com a finalidade de satisfazer as restrições específicas do problema. Contudo, como apontado por Cursi

Figura 1 - Exemplo de variação geométrica no encaixe de peças.



Legenda: Em geral, é improvável que se alcance a precisão especificada na fabricação de peças. As incertezas podem ser ocasionadas pela tolerância da fabricação, limitação de equipamentos ou imprecisão. É imprescindível que tais parâmetros sejam considerados durante projetos em que os resultados possam sofrer variações desse tipo.

Fonte: O autor, 2018

e Sampaio (2015), esta técnica é limitada em relação à sua generalização, uma vez que os coeficientes são específicos de situações particulares, além do fato de que os valores são gerados empiricamente, fazendo com que sua determinação seja pouco confiável quando a quantidade de dados é insuficiente.

Uma maneira consistente de se considerar as incertezas inerentes a um dado modelo matemático é através da otimização robusta. De maneira geral, a robustez pode ser compreendida como o estado onde o desempenho do processo ou tecnologia são pouco sensíveis a fatores externos, que podem causar variabilidade nos valores dos parâmetros e, consequentemente, no resultado final. Nesse sentido, a engenharia robusta concentra-se na redução efetiva de custos através da escolha seletiva dos parâmetros ideias, a fim de maximizar a confiabilidade do projeto. Portanto, na otimização robusta de projeto, o foco principal é a minimização dos efeitos de variações e incertezas no desempenho do produto, enquanto o seu rendimento é otimizado (TAGUCHI et al., 1999).

Por outro lado, as incertezas relativas ao projeto podem ser mensuradas através da estimativa da probabilidade de falha de um dado processo. Diferentemente da otimização robusta, em que o foco é a minimização da sensibilidade relacionada às perturbações que atuam no sistema, a otimização de projetos baseada em confiabilidade é uma abordagem que leva em conta a estimativa da probabilidade de falha em um sistema de engenharia, através de uma metodologia que considera tais incertezas com base em modelos probabilísticos, a fim de garantir a satisfação das restrições probabilísticas nos níveis desejados e obter o melhor compromisso entre custo e segurança (KESHTEGAR et al., 2018).

Como ambas as abordagens se concentram nas incertezas das variáveis de projeto e de outros parâmetros, revela-se uma tendência coerente de se aplicar as técnicas de robustez e confiabilidade em conjunto para problemas de otimização de modelos matemáticos. Naturalmente, a aplicação de todos esses métodos de otimização em projetos de engenharia requer mais esforços computacionais se comparados aos modelos puramente determinísticos. Contudo, o aumento do custo computacional é motivado pela necessidade de produzir aproximações de alta fidelidade ao verdadeiro comportamento físico do modelo, através da obtenção de resultados viáveis no que tange à utilização das soluções ótimas obtidas em condições reais de funcionamento e sem a necessidade de eliminação da fonte de variabilidade.

Grande parte das aplicações no campo da otimização está relacionada a problemas de objetivo único. Apesar disso, a maioria dos problemas reais estão sujeitos a múltiplos objetivos, que podem estar relacionados entre si, como aponta Babu et al. (2005). A modelagem de problemas complexos utilizando apenas um objetivo pode ser uma tarefa impraticável, fazendo com que sejam introduzidos eventuais erros ao modelo proposto, quando se assume determinadas simplificações para atender aos requisitos da modelagem. Em contrapartida, no contexto dos problemas de otimização multi-objetivo, os modelos possuem um grau de liberdade adicional em relação àqueles de objetivo único, de tal forma que não há uma solução única para o problema a ser resolvido, mas um conjunto de soluções, conhecidas como soluções ótimas de Pareto (COLLETTE; SIARRY, 2004). Essa flexibilidade faz com que a etapa de modelagem matemática possa ser desempenhada com maior rigor em relação a descrição do problema físico.

Naturalmente, a robustez é conflitante com o desempenho do sistema, visto que o nível de exigência relacionada ao grau de sensibilidade às perturbações externas está diretamente ligado à deterioração das soluções ótimas em relação ao caso determinístico. Portanto, os algoritmos de otimização não devem apenas encontrar as soluções que maximizam o desempenho, mas também garantir robustez suficiente (GASPAR-CUNHA; COVAS, 2008). Como a robustez e o desempenho são frequentemente conflitantes, eles podem ser abordados como um problema de otimização multi-objetivo, tal que a solução ótima deve ser escolhida entre aquelas obtidas no conjunto ótimo de Pareto, de forma a balancear a relação estabelecida entre desempenho e robustez (JIN; SENDHOFF, 2003 apud GASPAR-CUNHA; COVAS, 2008).

Além da presença de múltiplos objetivos conflitantes entre si, certos problemas de engenharia, quando formulados como problemas de otimização, apresentam espaços de busca extensivamente grandes e altamente complexos. Especificamente no contexto dos problemas multi-objetivo, as soluções ótimas de Pareto dominam coletivamente qualquer outra solução viável no espaço objetivo. Em outras palavras, considera-se que o conjunto obtido é melhor do que quaisquer outras soluções viáveis (MIETTINEN, 1998; COELLO COELLO et al., 2007). Nesse sentido, os algoritmos evolucionários provaram-se ser ótimas

ferramentas para a resolução de problemas deste tipo, com grande quantidade de aplicações conhecidas na literatura, uma vez que possuem diversas características que são desejáveis para evitar os obstáculos que tais aplicações podem produzir, como ampla capacidade exploratória do domínio considerado e diversidade da população, embora possuam simples mecanismos de funcionamento em relação às técnicas de otimização clássicas (ZITZLER et al., 2004).

Embora as abordagens relacionadas a robustez e confiabilidade já tenham sido empregadas em diversas aplicações na literatura, a incorporação de ambas em conjunto ainda é algo a ser explorado, tendo em vista sua grande aplicabilidade e pequena quantidade de publicações relacionadas ao assunto. Nesta tese, o tema em questão é abordado profundamente, através do emprego das técnicas na solução de problemas de engenharia, estudo das suas particularidades e procedimentos para aprimoramento das metodologias.

Os objetivos específicos desta tese são: (i) propor um algoritmo de análise de confiabilidade inversa, baseado no *steepest descent* e com tamanho de passo adaptativo, empregando informações implícitas sobre derivadas de segunda ordem, mas sem a necessidade de calcular explicitamente a matriz Hessiana e; (ii) apresentar uma formulação multi-objetivo para problemas de otimização levando em conta a robustez e a confiabilidade do conjunto de soluções e empregar o método proposto de análise de confiabilidade inversa na solução de problemas de *benchmark* e de engenharia para validação das propostas.

A organização deste trabalho está definida como apresentado a seguir. No Cap. 1, é definido um problema de otimização multi-objetivo e as características de um ótimo de Pareto. No capítulo seguinte, são apresentadas metodologias de otimização multi-objetivo. As metodologias determinísticas são divididas em três segmentos, que enquadram os métodos *a priori*, *a posteriori* e progressivos. Neste mesmo capítulos, também são apresentados os métodos baseados em metaheurísticas.

No Cap. 3, é apresentada a primeira das duas abordagens relacionadas com a avaliação de incertezas em problemas de otimização: análise de robustez. As diferentes maneiras de avaliação da sensibilidade das variáveis de projeto são apresentadas de forma progressiva em relação à evolução conceitual, desde o Método de Taguchi até as técnicas de análise de robustez com abordagem genética. Também é realizada uma revisão bibliográfica que embasa o progresso das técnicas e mostra diversas aplicações relacionadas.

A otimização baseada em confiabilidade é analisada no Cap. 4. Inicialmente são apresentados os fundamentos de análise de confiabilidade, que englobam as noções fundamentais acerca de variáveis aleatórias, medidas de dispersão e distribuições de probabilidade. Em seguida, são mostradas as noções de métodos probabilísticos capazes de aproximar a probabilidade de falha de um dado sistema. No mesmo capítulo, é feita uma revisão a respeito de métodos de análise de confiabilidade, a qual contempla alguns métodos de caráter local, global e de amostragem.

No capítulo seguinte, são abordados detalhadamente os métodos locais de confiabi-

lidade de primeira e segunda ordens. Apresenta-se todos os procedimentos necessários para aproximação do cálculo da probabilidade de falha e alguns algoritmos capazes de realizar essa estimativa. Também são discutidas maneiras de realizar a estimativa da probabilidade de falha, utilizando métodos locais, quando as variáveis aleatórias apresentam outros tipos de distribuição, senão a distribuição normal. Por fim, este capítulo estende-se à apresentação da abordagem inversa para análise de confiabilidade, também incluindo uma visão geral de alguns métodos neste contexto.

O Cap. 6 é dedicado à proposta de um método de análise de confiabilidade inversa, baseado no cálculo adaptativo de tamanho de passo de segunda ordem. Inicialmente, são apresentadas as motivações que embasam a elaboração desta nova metodologia e, em seguida, é desenvolvida a ideia para o cálculo do tamanho de passo do algoritmo e proposto o Método de Tamanho de Passo de Segunda Ordem Adaptativo. A efetividade da técnica é avaliada resolvendo-se alguns problemas de *benchmark* e um problema de engenharia da literatura. Todos os resultados são comparados com técnica de análise de confiabilidade recentes, no que diz respeito ao comportamento dos métodos, custo computacional e sensibilidade de seus parâmetros de controle.

No Cap. 7, é proposta uma formulação para problemas de otimização multi-objetivo capaz de considerar robustez e confiabilidade em conjunto. Neste contexto, é apresentado um esquema prático que detalha um procedimento para acoplamento da análise de incertezas durante o processo de otimização. A formulação proposta e o algoritmo para análise de confiabilidade inversa, estabelecido no capítulo anterior, são empregados na solução de diversos problemas de engenharia, a fim de verificar as particularidades dos resultados obtidos.

Comentários Preliminares Sobre Notação e Formatação

- Vetores e matrizes são escritos em **negrito**;
- Índices de vetores e matrizes são representados por elementos subscritos ou entre parênteses, precedendo o símbolo ao qual se refere;
- Elementos subscritos e sobrescritos precedendo símbolos matemáticos, quando representados em fonte romana (não itálicos), denotam um identificador complementar ao símbolo;
- Numerais e demais elementos sobrescritos, no formato matemático (em itálico), podem representar o expoente de determinado símbolo ou um identificador. Nestes casos, o significado é explicitado no texto;
- Elementos definidos no texto são mostrados em itálico ao serem introduzidos;

- Traços horizontais duplos são usados para especificar uma faixa de valores e traços triplos representam um separador textual;
- Termos em outras línguas são escritos em itálico;
- Apenas as equações referenciadas no texto são numeradas;
- Equações destacadas seguem a pontuação do texto;
- Unidades de medida são escritas em fonte romana.

1 CONCEITOS FUNDAMENTAIS DE OTIMIZAÇÃO

A otimização pode ser definida como a ciência de se determinar as melhores soluções para certos problemas frequentemente modelados a partir da análise do comportamento físico de algum sistema observado na engenharia. Esta área de pesquisa envolve o estudo de critérios de otimalidade para problemas diversos, a determinação de métodos algorítmicos de solução, o estudo da estrutura de tais métodos e a experimentação computacional, com técnicas tanto em condições simuladas quanto relacionadas a problemas reais.

Existe uma gama extremamente diversificada de aplicações práticas nesse ramo de pesquisa. Antes da década de 40, relativamente pouco se sabia sobre métodos para otimização numérica de funções de várias variáveis. Não há dúvida, entretanto, que o advento do computador foi primordial no desenvolvimento de técnicas de otimização e, de fato, em toda a análise numérica. A aplicabilidade dos métodos de otimização é generalizada, alcançando quase todas as atividades em que a informação numérica é processada, como a ciência, engenharia, matemática, economia, comércio e diversas outras (FLETCHER, 2013).

Para aplicar as técnicas numéricas da teoria de otimização a problemas de engenharia, em geral é necessário selecionar as variáveis do sistema que serão usadas para caracterizar ou identificar candidatos a solução e definir um modelo que expressará a maneira pela qual essas variáveis estão relacionadas. Em relação à primeira exigência, busca-se um conjunto de parâmetros de entrada capazes de caracterizar completamente o problema, do ponto de vista do projeto. Esses parâmetros, também chamados de variáveis, são considerados importantes, do ponto de vista do modelo computacional, se a variação de seus valores afeta significativamente a medida de desempenho do problema. Deve-se ter em mente que a complexidade de um problema de otimização cresce exponencialmente com o número de variáveis, portanto avaliar quais delas são realmente essenciais no modelo e adotar o menor número possível de variáveis pode viabilizar a otimização do problema (CAVAZZUTI, 2013).

Uma vez definidas as variáveis de projeto, o próximo passo na formulação do problema é construir o modelo que descreve a maneira como as variáveis estão relacionadas e podem influenciar os valores que são calculados. Na prática, a maioria dos estudos de otimização são realizados com a ajuda de modelos, que constituem uma representação matemática simplificada do sistema real analisado. Em muitos casos, os modelos são uma alternativa para a realização de estudos e testes, uma vez que a utilização do sistema real nessas tarefas pode ter um custo muito alto ou representar algum risco de operação, além da possibilidade de demandar um tempo inviável. Portanto, os modelos são normalmente usados no projeto de engenharia porque oferecem uma maneira mais eficiente de estudar os efeitos das mudanças das variáveis de projeto que podem afetar o desempenho de tal sistema (RAVINDRAN et al., 2006).

Em alguns problemas de otimização, as variáveis só fazem sentido se assumirem valores inteiros. Essa classe de problemas é chamada de programação inteira. O processo de obtenção de soluções ótimas de problemas desse tipo é conhecido como otimização discreta, a qual faz referência às possíveis soluções contidas em um conjunto finito de números. Para contornar essa condição, muitas técnicas propõem a resolução dos problemas, simulando a possibilidade de as variáveis serem capazes de assumir valores reais e, então, aproximando os resultados para o número inteiro mais próximo da solução obtida. Em contraste, a otimização contínua tem como objetivo obter soluções em um conjunto infinito de elementos. Nesse tipo de problema, a possível existência de suavidade das funções torna possível usar informações do objetivo e das restrições do problema em um determinado ponto, para deduzir informações sobre o seu comportamento em outros pontos da vizinhança.

Além dessa especificação, os problemas de otimização contínua também podem ser classificados de acordo com a natureza da função-objetivo (linear, não-linear, convexa e outras), o número de variáveis, a suavidade das funções (diferenciáveis ou não), entre outras. Outro aspecto muito importante é a existência de restrições no modelo. Na maioria das vezes, a sua presença pode afetar a solução ótima do problema. No caso mais simples, as restrições são obtidas definindo-se limites superiores e inferiores para cada variável ou, em casos mais complexos, podem ser estabelecidas usando-se equações ou inequações. Uma estratégia muito utilizada para o tratamento de restrições explícitas é a aplicação de técnicas de penalização, as quais penalizam os valores das variáveis que não satisfazem as condições impostas pelas restrições.

No caso dos procedimentos de otimização, a busca por soluções pode ser feita de maneira local ou global. No primeiro caso, o objetivo é a obtenção de um ponto em que o valor da função-objetivo seja menor do que em todos os outros pontos viáveis em sua vizinhança. Isso não significa que está sendo obtida a melhor solução possível, ou seja, a solução global, a qual pode ser mais complexa de ser computada, uma vez que todas as possíveis soluções do domínio devem ser consideradas, a fim de garantir que o resultado obtido se trata do ponto extremo. Uma classe de problemas muito estudados são os convexos, em que uma de suas principais características é o fato de que todas as soluções locais também são soluções globais (NOCEDAL; WRIGHT, 2006).

Modelos matemáticos também podem estar sujeitos a avaliação de diversos objetivos, para a obtenção de uma resposta adequada. Problemas desse tipo são inseridos em um ramo da ciência muito explorado, chamado de otimização multi-objetivo. Não é só em relação ao número de objetivos que essa classe de problemas se diferencia da otimização de um único objetivo. No caso das soluções relativas aos problemas de objetivo único, a intenção é calcular uma única solução que otimize a função-objetivo como um todo. A mesma afirmação não é válida para os problemas de otimização multi-objetivo. O aspecto conflitante dos objetivos insere diversas nuances ao tratamento desse tipo de modelo. A seguir, o problema geral de otimização multi-objetivo será formulado e, adicionalmente, serão apre-

sentadas suas principais características, a fim de fornecer sustentação para o restante do desenvolvimento desta tese.

Antes de expor os aspectos dos problemas de otimização multi-objetivo, porém, é imprescindível que se destaque uma propriedade muito importante acerca dos métodos de otimização. Inevitavelmente, surge o questionamento a respeito da existência do melhor método a ser aplicado para resolver um problema de otimização ordinário. Pensar em um método universal demandaria certo esforço para comprovar a sua capacidade de resolver qualquer problema de um certo tipo, sob quaisquer condições (como, por exemplo, infinitas possibilidades quanto às estimativas iniciais) e, principalmente, assegurar que o seu rendimento e eficácia sejam superiores a qualquer outra técnica da mesma espécie. A respeito desse tema, Wolpert e Macready (1997) enunciaram os Teoremas *No Free Lunch* para otimização.

Esses teoremas estabelecem que, se um algoritmo tem um bom desempenho em uma determinada classe de problemas, então, necessariamente, o seu desempenho é degradado no conjunto de todos os problemas restantes. Tais teoremas resultam em uma interpretação geométrica do que significa um algoritmo ser adequado para um problema de otimização. Em particular, se um algoritmo tem um desempenho melhor do que uma pesquisa aleatória em alguma classe de problemas, então deve ter um desempenho pior do que a pesquisa aleatória nos problemas restantes. Assim, as comparações que relatam o desempenho de um algoritmo específico com uma configuração de parâmetros dada em alguns problemas selecionados são de utilidade limitada. Embora tais resultados indiquem o comportamento em uma estreita faixa de problemas considerados, deve-se ter muito cuidado ao tentar generalizar esses resultados para outros problemas.

1.1 Formulação do Problema de Otimização Multi-objetivo

Seja $P \subset \mathbb{R}^n$ o hiperparalelepípedo em \mathbb{R}^n formado por restrições inferiores e superiores aos valores de $\mathbf{x} \in \mathbb{R}^n$, tal que $\mathbf{x} = (x(1), \dots, x(n))^T$. Portanto, $\mathbf{x}_{\text{inf}} \leq \mathbf{x} \leq \mathbf{x}_{\text{sup}}$, onde estas desigualdades são entendidas coordenada a coordenada, isto é, $x_{\text{inf}}(i) \leq x(i) \leq x_{\text{sup}}(i)$, para todo $i = 1, \dots, n$. Por sua vez, assuma que f , a chamada *função multi-objetivo*, é um função vetorial com valores em \mathbb{R}^m e definida em P , isto é, $\mathbf{f} : P \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, com $m \geq 2$, onde $\mathbf{f} = (f_1, \dots, f_m)^T$ é o *vetor de objetivos* e $f_k : P \rightarrow \mathbb{R}$, para $k = 1, \dots, m$.

Será definido mais adiante o problema de otimização multi-objetivo envolvendo a função \mathbf{f} . Para já, o enfoque será restrito a problemas de minimização. O domínio P da função \mathbf{f} é também chamado de *espaço das variáveis de decisão*. Por sua vez, o vetor $\mathbf{x} \in P$ é denominado *vetor de decisão* e suas entradas são chamadas de *variáveis de decisão*. Algumas vezes no decorrer do texto, o vetor \mathbf{x} será chamado de solução considerando que, ainda que não resolva nenhum problema específico, é entre os elementos ou pontos do espaço das

variáveis de decisão que eventualmente se encontrará a solução do problema de otimização multi-objetivo.

Além disso, normalmente, as soluções têm que satisfazer restrições adicionais, que escritas na forma vetorial, se expressam como $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ e $\mathbf{h}(\mathbf{x}) = \mathbf{0}$, onde $\mathbf{g} : P \rightarrow \mathbb{R}^p$ e $\mathbf{h} : P \rightarrow \mathbb{R}^q$, ou seja, $g_i(\mathbf{x}) \leq 0$, para $i = 1, \dots, p$, e $h_j(\mathbf{x}) = 0$, para $j = 1, \dots, q$. O conjunto das soluções que satisfazem tais restrições, denotado por Ω , é chamado de *conjunto viável* ou *espaço de busca*, e seus pontos são denominados como *soluções viáveis*. Observe-se, novamente, que os elementos de Ω são apenas candidatos a solução do problema de otimização multi-objetivo.

Um outro conjunto relevante na questão de otimização multi-objetivo que sequer é mencionado, por desnecessário, na otimização clássica (com uma única função-objetivo), é a imagem da função multi-objetivo quando considerado o espaço viável, chamado de *espaço objetivo* e definido como $V = \text{Im}(f|_{\Omega}) \doteq \{\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^m, \forall \mathbf{x} \in \Omega\}$. Conhecidas estas definições, um problema de otimização multi-objetivo é definido como

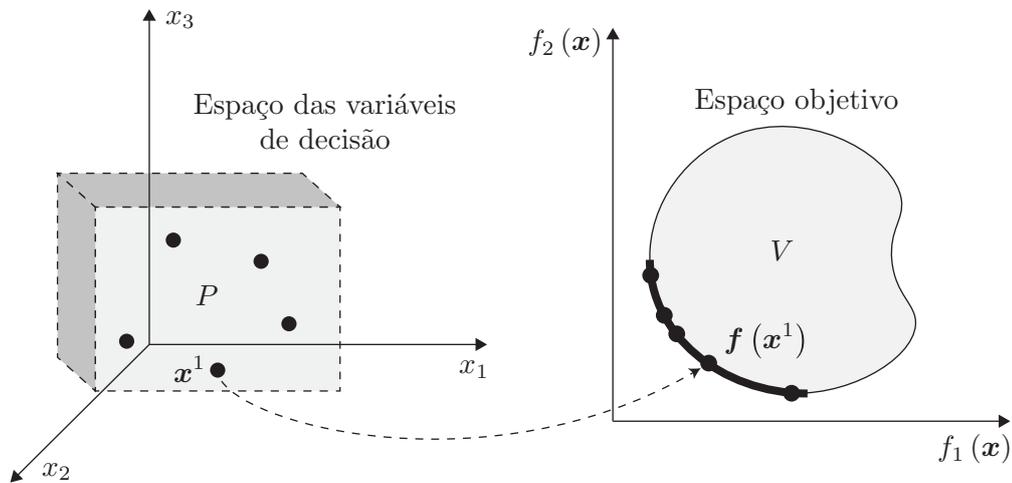
$$\begin{aligned} \min \mathbf{f}(\mathbf{x}) &= (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \\ \text{Sujeito a } g_i(\mathbf{x}) &\leq 0 \\ h_j(\mathbf{x}) &= 0 \\ \mathbf{x}_{\text{inf}} &\leq \mathbf{x} \leq \mathbf{x}_{\text{sup}} \end{aligned} \quad , \quad (1)$$

para $i = 1, \dots, p$ e $j = 1, \dots, q$. A Fig. 2 apresenta essa formulação de maneira ilustrativa, com o conjunto ótimo de Pareto representado pela curva em negrito, que será detalhado a seguir.

Por conta do conflito existente entre os objetivos, não existe um único ponto capaz de otimizar todas as funções simultaneamente. Em vez disso, as melhores soluções que podem ser obtidas são denominadas soluções ótimas de Pareto (as quais formam a curva de Pareto). Como os pontos pertencentes à solução ótima de Pareto dominam coletivamente qualquer outro ponto viável no espaço de busca, todas essas soluções são consideradas melhores do que quaisquer outras, como será definido a seguir (DEB; GUPTA, 2006).

Portanto, a otimização multi-objetivo consiste em encontrar um conjunto de pontos que represente o melhor equilíbrio em relação à tarefa de minimizar todos os objetivos simultaneamente, isto é, uma coleção de soluções que relacionem os objetivos, os quais são conflitantes entre si, na maioria dos casos. De acordo com Babu et al. (2005), isso significa que a melhoria de um objetivo pode levar à deterioração de outro e vice versa. Esse fato levanta a questão sobre como os diferentes objetivos devem ser combinados para produzir uma solução viável e também a demanda de procurar uma solução final—em um sentido a ser descrito mais à frente—para o problema a ser resolvido (pós-processamento do conjunto de pontos obtido). A seguir, serão apresentados os detalhes relacionados ao conjunto ótimo de Pareto e as definições relacionadas a essa teoria.

Figura 2 - Relação entre o espaço das variáveis de decisão e o espaço objetivo em um problema de otimização, onde a função multi-objetivo $f : P \subset \mathbb{R}^3 \rightarrow \mathbb{R}^2$.



Legenda: No espaço das variáveis de decisão, o cubo ilustra os limites inferior (\mathbf{x}_{inf}) e superior (\mathbf{x}_{sup}) do vetor de decisão. No espaço objetivo, é possível ver as soluções correspondentes àquelas que aparecem no interior do cubo e que compõem a curva de Pareto, representada pelo segmento em negrito.

Fonte: O autor, 2019.

1.1.1 Ótimo de Pareto

O conceito de otimalidade de Pareto faz referência a Vilfredo Pareto (1848–1923), um economista e sociólogo franco-italiano, que estabeleceu esse conceito no campo da economia, baseado em uma multiplicidade de objetivos, isto é, no permanente conflito de interesses e antagonismos na vida social. Esse campo especial de otimização também é chamado de otimização de Pareto (ESCHENAUER et al., 1990). A teoria desenvolvida foi estendida para diversas outras áreas além da economia, em virtude da expressa ocorrência de conflitos de grandezas, evidentes em vários processos naturais e tecnológicos. O conceito de otimalidade de Pareto será desenvolvido a seguir, em forma de exemplo, antes que sejam apresentadas as definições matemáticas.

Para isso, considere um problema clássico da engenharia, onde deseja-se otimizar a relação entre os custos de mão de obra e de matéria-prima utilizada em uma construção. Claramente, trata-se de um problema contendo dois objetivos que devem ser resolvidos simultaneamente, a fim de se obter o menor custo de mão de obra possível, desempenhando a menor aplicação de matéria-prima praticável. Nesse contexto, as funções-objetivo que descrevem os custos não são importantes. Apenas assumamos que o valor da mão de obra é calculado por $f_1(\mathbf{x})$ e a função que determina a quantidade de recursos empregados no processo é representada por $f_2(\mathbf{x})$.

A solução ideal desse problema seria a obtenção de uma solução viável com custo nulo, tanto de mão de obra quanto de matéria-prima. Contudo, espera-se que o modelo matemático utilizado para computar essas grandezas conte com restrições relacionadas às variáveis de decisão, de forma que o vetor de decisão assumira valores dentro de uma faixa que seja possível contabilizar resultados diferentes dessa solução impraticável. Apesar disso, conhecer a solução ideal significa compreender o caso ideal, que dispõe de valores de f mais próximos da origem quanto for possível, no espaço objetivo.

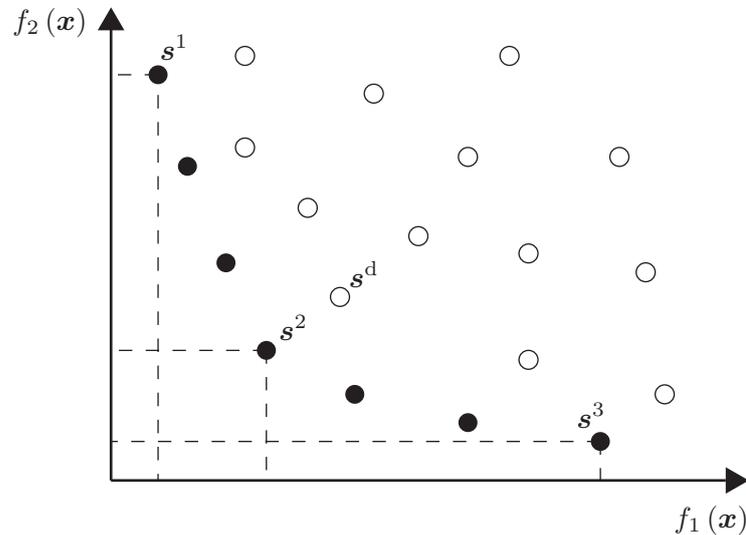
Conceitualmente, o problema apresenta uma dependência explícita entre os objetivos. Independentemente dos aspectos particulares do modelo, esse tipo de aplicação, em geral, requer maior custo de mão de obra quando são utilizados determinados materiais. Como exemplo, assumamos um vetor de decisão x composto por entradas que representem o custo de mão de obra exigido por funcionários com diferentes qualificações e a quantidade de matéria-prima de diferentes tipos utilizada no serviço. Nesse exemplo, o manuseio de um determinado material requer o empenho de um ou mais funcionários com certo grau de qualificação e os diferentes tipos de matéria-prima têm valores distintos. Portanto, o problema apresenta um conflito entre os objetivos, uma vez que a utilização de certos materiais (e a quantidade necessária) afeta o valor da mão de obra empregada.

A Fig. 3 apresenta um conjunto de possíveis valores ótimos para o problema proposto. Os pontos ilustrados pelos círculos preenchidos são chamados de valores dominantes e compõem a curva de Pareto, isto é, são os valores ótimos do problema de otimização que relaciona custo de mão de obra e de matéria-prima. Esses resultados são caracterizados por não haver nenhum outro que apresente valores melhores. Portanto, qualquer outra solução que seja obtida no conjunto viável é chamada de solução dominada. Esse conjunto, cujos valores correspondentes no espaço objetivo são complementares à curva de Pareto, é representado pelos pontos não preenchidos na Fig. 3. A seguir, esses conceitos serão definidos matematicamente.

Antes, entretanto, é conveniente analisar os aspectos do conjunto ótimo de soluções. O conflito entre os objetivos na otimização multi-objetivo revela uma característica fundamental e diferente da otimização mono-objetivo, que diz respeito a tomada de decisões. Todos os valores ótimos que compõem o conjunto ótimo de Pareto podem ser escolhidos, pois demonstram compromisso viável entre os objetivos. Entretanto, não há um critério absoluto para se definir qual é o resultado ideal, dentre os que estão contidos no conjunto ótimo. Ao invés disso, os interesses específicos (condição adicional) da aplicação devem ser primordiais para se adotar um resultado. Na realidade, a solução que será tomada como ideal depende de interesses particulares.

Em geral, os valores ótimos diferenciam-se em relação ao nível de compromisso com cada objetivo. Na Fig. 3, é possível notar que s^1 e s^3 são pontos mínimos que apresentam compromissos extremos e distintos, uma vez que o primeiro prioriza o custo de mão de obra e a segunda privilegia o custo de matéria-prima. Por sua vez, a solução s^2 mostra

Figura 3 - Exemplo de otimização multi-objetivo de um problema genérico, que relaciona custo de mão de obra e matéria-prima.



Fonte: O autor, 2019.

um compromisso intermediário. Isso não significa que esse balanço represente a melhor opção. Na realidade, essa decisão está vinculada a outros critérios inerentes a interesses particulares. Portanto, a subjetividade faz parte do processo de otimização multi-objetivo, na etapa de pós-processamento das soluções.

Todos esses conceitos são definidos rigorosamente e serão apresentados a seguir. Inicialmente, considere a noção de dominância. Esse conceito tem a finalidade de classificar as soluções, definindo uma maneira de compará-las no contexto das funções multi-objetivo. Uma solução é dita dominante em relação a outra, se não for pior em nenhum dos objetivos e se for estritamente melhor em pelo menos um dos objetivos. Como o conceito de dominância proporciona uma maneira de comparar soluções com múltiplos objetivos, a maioria dos métodos de otimização multi-objetivo faz uso desse conceito para procurar por soluções não-dominadas.

Definição 1.1.1. Uma solução \mathbf{x}^1 é dita dominante em relação a outra solução \mathbf{x}^2 , representado por $\mathbf{x}^1 < \mathbf{x}^2$, se ambas as condições abaixo forem satisfeitas:

1. A solução \mathbf{x}^1 não é pior do que \mathbf{x}^2 em todos os objetivos, isto é, $f_k(\mathbf{x}^1) \leq f_k(\mathbf{x}^2)$, para todo $k = 1, \dots, m$.
2. A solução \mathbf{x}^1 é estritamente melhor do que \mathbf{x}^2 em, pelo menos, um objetivo, ou seja, $f_k(\mathbf{x}^1) < f_k(\mathbf{x}^2)$ para algum $k \in \{1, \dots, m\}$.

Se alguma dessas condições é violada, diz-se que a solução \mathbf{x}^1 não domina a solução \mathbf{x}^2 . Como exemplo, pode-se reavaliar as soluções referentes ao problema da Fig. 3. Para isto,

considere o conjunto de soluções ótimas \mathbf{x}^i , tal que os seus respectivos valores ótimos são denotados por $\mathbf{s}^i = \mathbf{f}(\mathbf{x}^i)$. Tomando como base a solução \mathbf{x}^2 , correspondente ao mínimo \mathbf{s}^2 e comparando-a com a solução dominada \mathbf{x}^d , relacionada ao valor \mathbf{s}^d , pode-se constatar que, de fato, \mathbf{x}^2 não é pior do que \mathbf{x}^d em todos os objetivos, o que satisfaz a primeira condição da definição. Além disso, \mathbf{x}^2 é estritamente melhor do que \mathbf{x}^d em, pelo menos, um objetivo, assegurando que a segunda condição seja satisfeita e, conseqüentemente, a definição de dominância como um todo. O mesmo não ocorre se dois pontos presentes no conjunto ótimo de Pareto forem analisados. Note que a comparação entre \mathbf{x}^2 e qualquer uma das duas soluções relativas aos valores na vizinhança de \mathbf{s}^2 viola o primeiro critério de dominância. Portanto, não são dominantes entre si e fazem parte do conjunto ótimo.

Existem três possibilidades resultantes da verificação de dominância entre duas soluções: (i) a solução \mathbf{x}^1 pode dominar \mathbf{x}^2 ; (ii) a solução \mathbf{x}^1 pode ser dominada por \mathbf{x}^2 ; (iii) pode não haver relação de dominância entre \mathbf{x}^1 e \mathbf{x}^2 . Além disso, a relação de dominância não é uma relação de equivalência, uma vez que não satisfaz às três propriedades previstas, como detalhado abaixo (CORMEN et al., 1990 apud DEB, 2001):

- Reflexividade: a segunda condição da Def. 1.1.1 impede que a relação de dominância seja reflexiva, visto que qualquer solução \mathbf{x}^1 nunca é dominada por ela mesma;
- Simetria: a relação de dominância não é simétrica, uma vez que $\mathbf{x}^1 < \mathbf{x}^2$ não implica que $\mathbf{x}^2 < \mathbf{x}^1$. De fato, se \mathbf{x}^1 domina \mathbf{x}^2 , então \mathbf{x}^2 não domina \mathbf{x}^1 , o que impede que a relação de dominância atenda às condições da propriedade de simetria;
- Transitividade: é verdade que se $\mathbf{x}^1 < \mathbf{x}^2$ e $\mathbf{x}^2 < \mathbf{x}^3$, então $\mathbf{x}^1 < \mathbf{x}^3$. Portanto, a dominância é uma relação transitiva.

Há outra propriedade interessante que a relação de dominância possui. Se a solução \mathbf{x}^1 não dominar a solução \mathbf{x}^2 , isso não implica que \mathbf{x}^2 domine \mathbf{x}^1 . Uma vez conhecido o conceito de dominância, fica clara a definição do conjunto ótimo de Pareto.

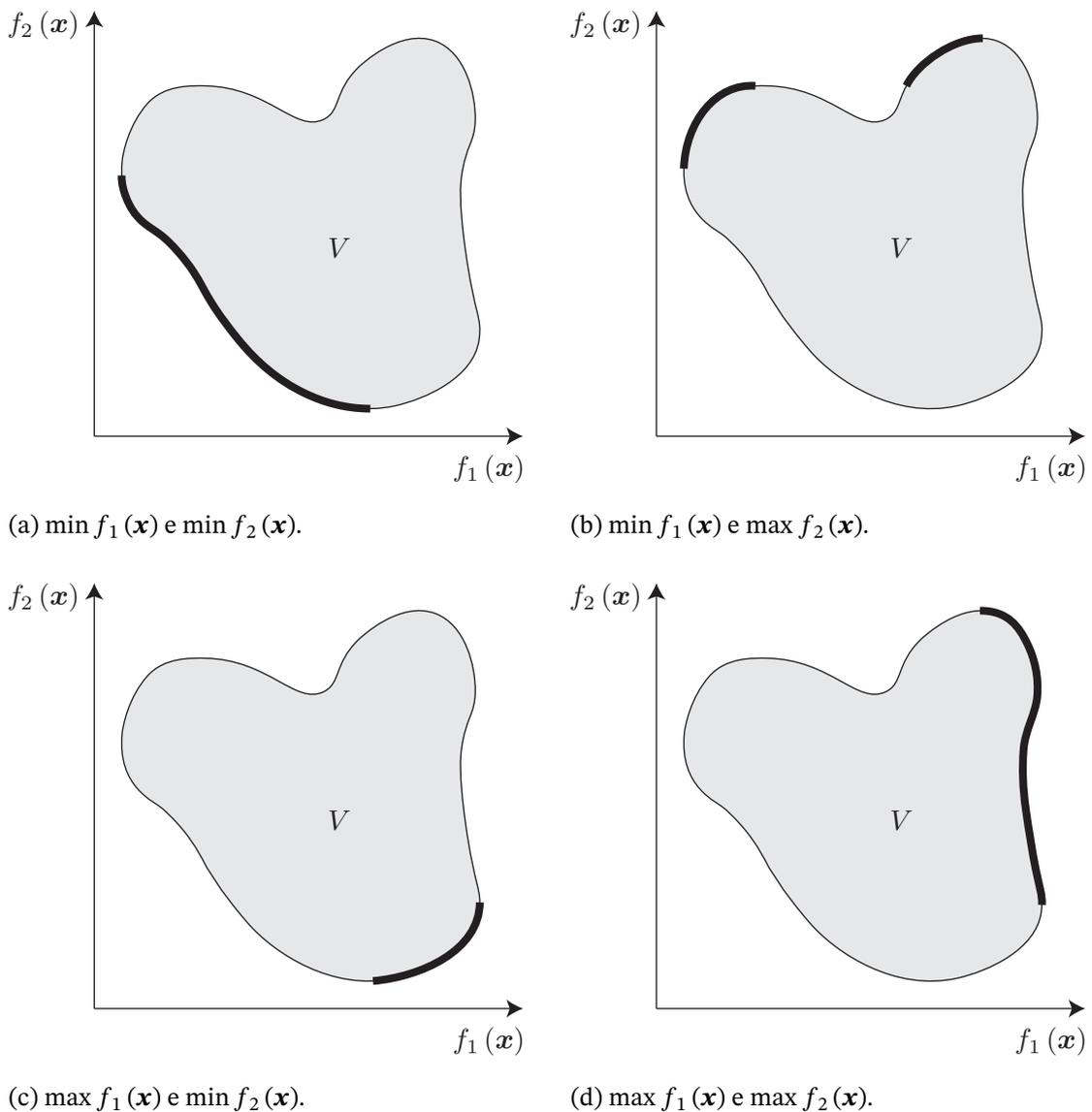
Definição 1.1.2. Um vetor viável $\mathbf{x}^* \in \Omega$ é dito ser um *minimizador de Pareto* (ou *otimizador de Pareto*) se não existir um outro vetor viável $\mathbf{x} \in \Omega$, tal que $f_k(\mathbf{x}) \leq f_k(\mathbf{x}^*)$ para todo $k = 1, \dots, m$ e $f_k(\mathbf{x}) < f_k(\mathbf{x}^*)$ para algum índice $k \in \{1, \dots, m\}$.

Definição 1.1.3. Um vetor de objetivos $\mathbf{z}^* \in V$ é dito ser um *mínimo de Pareto* (ou *ótimo de Pareto*) se não existir um outro vetor de objetivos $\mathbf{z} \in V$ tal que $z_i \leq z_i^*$ para todo $k = 1, \dots, m$ e $z_i < z_i^*$ para algum índice $k \in \{1, \dots, m\}$. Equivalentemente, \mathbf{z}^* é um mínimo de Pareto se o vetor viável correspondente for um minimizador de Pareto.

Geralmente, há um número infinito de valores ótimos de Pareto, os quais constituem o chamado *conjunto de Pareto* ou *frente de Pareto*, isto é, a curva formada pela junção dos valores ótimos do problema de otimização multi-objetivo. Esse conjunto pode ser não-conexo, ou seja, as soluções podem estar dispostas sobre um segmento que não apresenta

continuidade. Os diferentes tipos de conjuntos de Pareto, em um problema de dois objetivos, podem ser vistos na Fig. 4. As soluções apresentadas dependem da finalidade de cada objetivo. Como mencionado anteriormente, o enfoque será restrito aos problemas de minimização. Apesar disso, o princípio da dualidade, no contexto dos problemas de otimização, define que um problema de minimização pode ser convertido em um de maximização bastando que a minimização seja realizada sobre o objetivo com valor negativo, como apontado por Deb (2001). Um exemplo de conjunto não-conexo pode ser visto na Fig. 4b.

Figura 4 - Tipos de conjuntos de Pareto, dependendo da finalidade de cada um dos objetivos.



Fonte: Adaptado de Deb (2001).

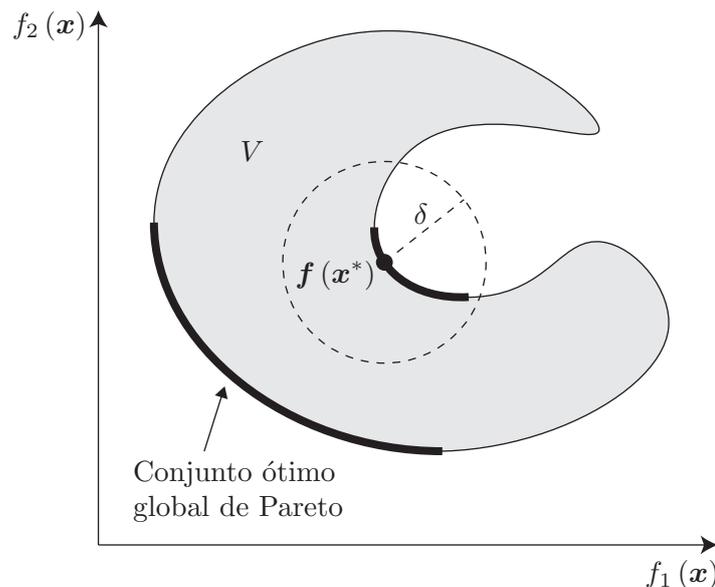
O conceito de otimalidade apresentado na Def. 1.1.2 diz respeito aos ótimos de Pareto globais. Contudo, possíveis restrições presentes no espaço de busca podem limitar o conjunto de soluções viáveis. Nesse caso, os resultados obtidos podem se tratar de soluções

ótimas locais, caracterizando o conjunto de minimizadores locais de Pareto, como definido abaixo.

Definição 1.1.4. Um vetor de solução viável \mathbf{x}^* é dito ser um *minimizador local de Pareto* se existe um $\delta > 0$, com $\delta \in \mathbb{R}$, tal que não exista um vetor \mathbf{x} que domine \mathbf{x}^* , com \mathbf{x} tal que $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m \cap B(\mathbf{f}(\mathbf{x}^*), \delta)$, onde $B(\mathbf{f}(\mathbf{x}^*), \delta)$ representa uma bola aberta, de raio δ e centrada em $\mathbf{f}(\mathbf{x}^*)$, isto é, $B(\mathbf{f}(\mathbf{x}^*), \delta) = \{\mathbf{y} \in \mathbb{R}^m \mid \|\mathbf{y} - \mathbf{f}(\mathbf{x}^*)\| < \delta\}$.

Na prática, somente soluções ótimas locais de Pareto são obtidas computacionalmente, a menos que algum requisito adicional, como convexidade, seja cumprido. A Fig. 5 ilustra o conjunto de valores ótimos locais, no interior da bola aberta dada por $B(\mathbf{f}(\mathbf{x}^*), \delta)$. Nesse exemplo, fica claro que o conjunto ótimo local de Pareto não é um subconjunto do conjunto dos valores ótimos globais. Este fato ocorre, exatamente, pois o exemplo ilustrado não corresponde a um problema de otimização multi-objetivo convexo.

Figura 5 - Conjunto de valores ótimos locais de Pareto, no interior de $B(\mathbf{f}(\mathbf{x}^*), \delta)$.



Fonte: O autor, 2019.

A convexidade de um problema de otimização multi-objetivo é uma propriedade muito importante, pois ela garante que, uma vez computados os minimizadores de Pareto, o conjunto de soluções ótimas obtido corresponde ao conjunto ótimo global de Pareto (MITTINEN, 1998). Além disso, existem muitos algoritmos capazes de lidar bem com problemas convexas na otimização multi-objetivo, mas enfrentam dificuldades na resolução dos problemas não-convexos, como poderá ser visto no Cap. 2. Por isso, é importante que se conheça as definições relacionadas a convexidade, apresentadas a seguir.

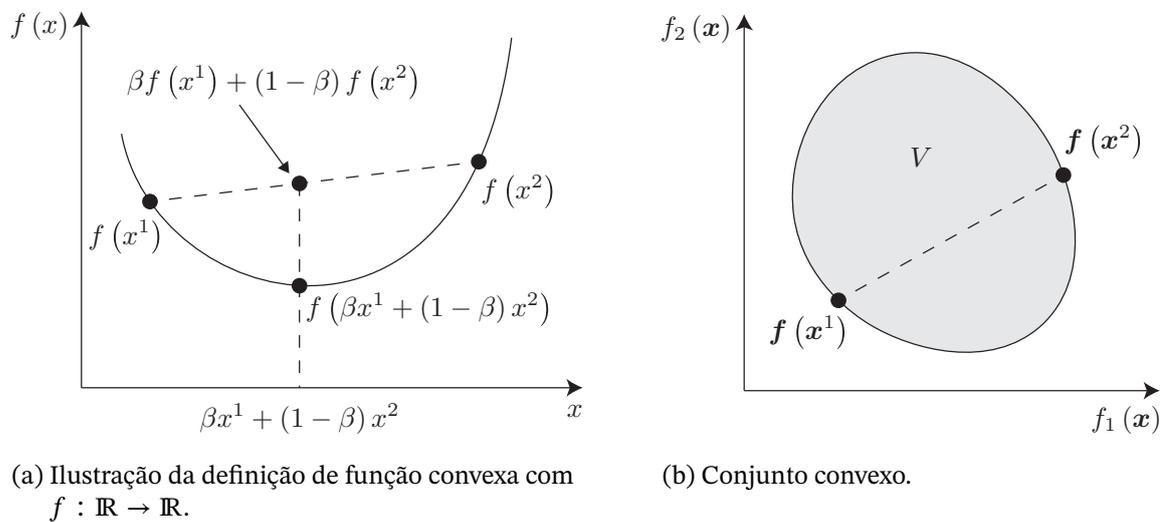
Definição 1.1.5. Uma função $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ é dita convexa se, para todo $\mathbf{x}^1, \mathbf{x}^2 \in \mathbb{R}^n$, é

válido que $f_k(\beta \mathbf{x}^1 + (1 - \beta) \mathbf{x}^2) \leq \beta f_k(\mathbf{x}^1) + (1 - \beta) f_k(\mathbf{x}^2)$, para todo $0 \leq \beta \leq 1$.

Definição 1.1.6. Um conjunto $S \subset \mathbb{R}^n$ é dito convexo se $\mathbf{x}^1, \mathbf{x}^2 \in S$ implica que $\beta \mathbf{x}^1 + (1 - \beta) \mathbf{x}^2 \in S$, para todo $0 \leq \beta \leq 1$.

Em outras palavras, um conjunto S é convexo se, dados dois pontos distintos neste conjunto, o segmento que os liga também está no conjunto S (COLLETTE; SIARRY, 2004). Por sua vez, um problema de otimização multi-objetivo é convexo se todas as funções-objetivo e a região viável forem convexas (MIETTINEN, 1998). A Fig. 6 mostra dois casos distintos de funções convexas.

Figura 6 - Exemplos de funções-objetivo convexas.



(a) Ilustração da definição de função convexa com $f : \mathbb{R} \rightarrow \mathbb{R}$.

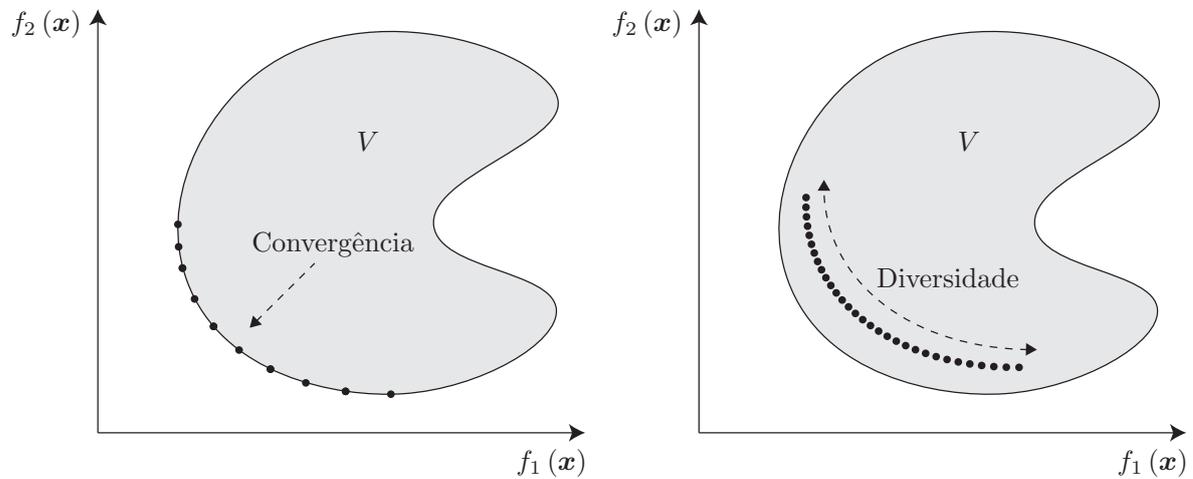
(b) Conjunto convexo.

Fonte: O autor, 2019.

Na presença de múltiplas soluções ótimas de Pareto, é difícil eleger um resultado em relação aos demais, sem qualquer informação adicional sobre o problema. Na ausência de tais informações, todas as soluções ótimas de Pareto são igualmente importantes. Portanto, é essencial encontrar tantas soluções ótimas de Pareto quanto possível. Essa necessidade engloba duas exigências fundamentais no processo de otimização multi-objetivo: (i) encontrar um conjunto de valores ótimos o mais próximo possível da fronteira de Pareto; (ii) encontrar um conjunto de valores ótimos tão diversificadas quanto possível.

O primeiro requisito é obrigatório em qualquer tarefa de otimização. Por outro lado, a exigência de diversidade de resultados é totalmente específica para otimização multi-objetivo. Além de serem suficientemente próximas da curva de Pareto, as soluções também devem estar esparsamente dispostas na frente de Pareto, como mostra o exemplo da Fig. 7. Somente com um conjunto diversificado de soluções, pode-se ter a certeza da qualidade dos valores ótimos computados, possibilitando a escolha adequada de uma solução na etapa de pós-processamento dos resultados.

Figura 7 - Exemplo da relação de convergência e diversidade de soluções em um problema de $\min f_1(\mathbf{x})$ e $\min f_2(\mathbf{x})$.



(a) Conjunto de resultados sobre a frente de Pareto, mas com diversidade insatisfatória.

(b) Conjunto de resultados com diversidade adequada, mas afastados da frente de Pareto.

Fonte: O autor, 2019.

Em vista destes conceitos, no próximo capítulo é feito uma revisão dos principais métodos capazes de obter soluções de problemas de otimização multi-objetivo. Os aspectos gerais de cada um deles serão apresentados, possibilitando a avaliação dos seus pontos positivos e negativos. De antemão, é possível antecipar que os métodos serão subdivididos em duas grandes vertentes: determinísticos e evolucionários. Naqueles da primeira categoria, é praticamente unânime a utilização do conceito de escalarização do problema, discutida no referido capítulo. Essa prática, uma vez que resolve problemas isolados, não propicia a diversidade de soluções. Por outro lado, o emprego de populações de indivíduos, que as técnicas evolucionárias utilizam, possibilitam a ocorrência de resultados ótimos diversificados, uma vez que as soluções podem ser comparadas entre si.

2 METODOLOGIAS DE OTIMIZAÇÃO MULTI-OBJETIVO

A obtenção de valores ótimos de Pareto desempenha um papel crucial na otimização multi-objetivo e, matematicamente, o problema é considerado resolvido quando o conjunto ótimo de Pareto é encontrado, como visto no Cap. 1. Contudo, é essencial que no processo de otimização o conjunto ótimo de Pareto seja único. Portanto, o ordenamento das soluções ótimas de Pareto possibilita a correta tomada de decisões, de maneira coerente, uma vez que as soluções calculadas durante o processo de otimização são, teoricamente, as melhores possíveis. Nesse contexto, são apresentados os métodos de otimização multi-objetivo, que têm a capacidade de obter o conjunto ótimo de Pareto.

Dentre as abordagens clássicas, uma estratégia muito utilizada é a chamada escalarição do problema, que significa converter um problema de otimização multi-objetivo em um único—ou em uma família de problemas de objetivo único. Essa abordagem permite a aplicação das técnicas e conceitos de otimização mono-objetivo na obtenção de uma família de soluções do problema multi-objetivo original. Contudo, devido à dependência em certos parâmetros, o uso desta estratégia pode ser limitado (MIETTINEN, 1998).

Naturalmente, com o advento das técnicas evolucionárias de otimização, diversas estratégias para resolução de problemas multi-objetivo foram desenvolvidas. Em vista das limitações pertinentes às técnicas clássicas de otimização multi-objetivo, uma das grandes vantagens dos métodos evolucionários é a utilização de uma população de soluções a cada iteração. Esse recurso, por si só, confere a esta classe de métodos uma enorme vantagem na solução de problemas de otimização de objetivos múltiplos. Tal aspecto propicia o aumento da diversidade de soluções, o que concede ao algoritmo a capacidade de não demandar que o problema de otimização multi-objetivo seja convertido em um problema equivalente, no contexto mono-objetivo. Consequentemente, os resultados também não estarão condicionados à escolha de parâmetros responsáveis pela transformação do problema (DEB, 2001).

Em relação à classificação dos métodos clássicos, existem algumas tendências na literatura, que levam em conta critérios distintos. Evidentemente, não existe uma classificação absoluta, uma vez que um dado método pode se encaixar em mais de uma categoria, dependendo da interpretação. Uma das categorizações mais utilizadas, veja-se Fonseca e Fleming (1995), Zitzler (1999), van Veldhuizen e Lamont (2000), e Collette e Siarry (2004), considera três gêneros de métodos distintos: *a priori*, *a posteriori* e progressivos. Sinteticamente, eles diferem uns dos outros no que diz respeito ao instante da realização do procedimento de escolha das melhores soluções (comumente conhecido, nessa área de pesquisa, como *trade-off*). Nesta divisão, não estão incluídos os métodos de otimização multi-objetivo que utilizam metaheurísticas.

A seguir, serão apresentados vários métodos clássicos, considerando as três classes mencionadas anteriormente, para otimização multi-objetivo. Além disso, também serão

mostrados dois métodos que utilizam a abordagem de metaheurísticas. Deve-se ter em mente que a quantidade de métodos existentes é muito grande. A proposta dessa revisão não é cobrir todos os métodos presentes na literatura, mas introduzir várias filosofias e formas de abordar a solução de problemas de otimização multi-objetivo, com o intuito de expor a diversidade de maneiras que um problema desse tipo pode ser resolvido.

2.1 Métodos *a Posteriori*

2.1.1 Método da Soma Ponderada

O Método da Soma Ponderada (*Weighted Sum Method*), cujos trabalhos pioneiros são os de Gass e Saaty (1955) e Zadeh (1963), dentre os registros da literatura a que se tem acesso, é o mais simples e provavelmente o mais utilizado, dentre os métodos clássicos. Como o nome sugere, a proposta do método é escalarizar um conjunto de objetivos em um único, pré-multiplicando cada um desses objetivos por um peso fornecido, minimizando a soma ponderada dos objetivos. Através desse procedimento, o problema de otimização multi-objetivo é convertido em um problema mono-objetivo, para o qual existe uma vasta gama de métodos para resolução (MIETTINEN, 1998; DEB, 2001).

Matematicamente, o funcionamento do método é como segue: considere os coeficientes de ponderação $w_i \in \mathbb{R}$, tal que $w_i \geq 0$ para todo $i = 1, \dots, k$, sendo k o número total de objetivos do problema original. Em geral, supõe-se que os pesos estejam normalizados, isto é

$$\sum_{i=1}^k w_i = 1 . \quad (2)$$

Sob estas condições, define-se um novo problema de otimização, escalarizado a partir do problema de otimização multi-objetivo original, da forma $\min_{\mathbf{x} \in \Omega} \bar{f}(\mathbf{x})$, com

$$\bar{f}(\mathbf{x}) = \sum_{i=1}^k w_i f_i(\mathbf{x}) . \quad (3)$$

Nesta formulação, os pesos não refletem a relativa importância das funções-objetivo (COELLO COELLO, 1999 apud LOBATO; STEFFEN JR, 2017). Caso haja necessidade de se ponderar cada objetivo, através de um critério arbitrário, é possível estabelecer seus níveis de influência através de um procedimento de normalização de funções. Para tanto, considera-

se a seguinte modificação na Eq. (3):

$$\bar{f}(\mathbf{x}) = \sum_{i=1}^k w_i c_i f_i(\mathbf{x}) ,$$

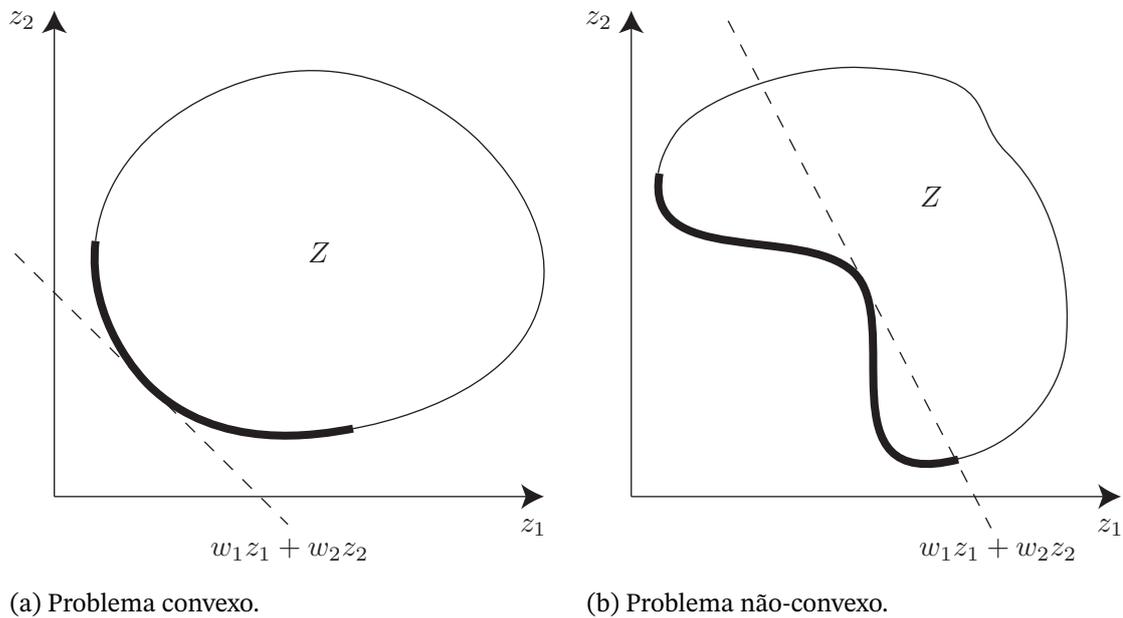
onde c_i são constantes que normalizam as funções a serem otimizadas. Dentre algumas possibilidades, esses parâmetros podem assumir os valores $c_i = 1/f_i^{\text{opt}}$, tal que f_i^{opt} denota o valor ótimo da função $f_i(\mathbf{x})$ minimizada individualmente, na presença de todas as restrições originais (OSYCZKA, 1984).

Embora a ideia seja simples, ela introduz uma questão não tão trivial: quais valores dos pesos devem ser usados? Obviamente, não há uma resposta única para essa pergunta. Para efeitos de argumentação, considera-se o caso em que a função-objetivo é composta por dois objetivos, isto é, $m = 2$. Neste caso, a tarefa é encontrar uma linha de contorno com o valor mínimo da função. Isso acontece com a reta que é tangencial ao espaço objetivo e também está no canto inferior esquerdo desse espaço, como pode ser visto na Fig. 8. Se um vetor de peso diferente for usado, a inclinação da reta de contorno seria diferente e o procedimento acima, em geral, resultaria em uma solução ótima diferente. Para um problema de otimização multi-objetivo, várias soluções ótimas de Pareto podem ser encontradas resolvendo-se o problema correspondente com múltiplos vetores de peso positivo, progressivamente, cada vez encontrando uma solução ótima de Pareto (DEB, 2001).

Portanto, a solução geral do Método da Soma Ponderada é sempre uma solução ótima de Pareto se os coeficientes de ponderação são todos positivos e se a solução é única, sem qualquer imposição adicional. Embora essa afirmação tenha garantia matemática (mais detalhes podem ser vistos em Miettinen (1998)), há uma grande desvantagem, baseada no fato de que o método não é capaz de obter todas as soluções ótimas de Pareto para problemas não-convexos, dependendo do método de otimização utilizado. Essa afirmação fica clara através da análise da Fig. 8. Seja Z o espaço objetivo, definido na Seção 1.1. No primeiro caso, onde aborda-se um problema convexo, é fácil perceber que, através da variação de w_1 e w_2 , a reta tangente, representada pela linha tracejada, sofrerá uma variação na sua inclinação, tornando capaz a obtenção de todas as soluções que compõem o conjunto ótimo de Pareto (ilustrado pelo segmento em negrito), através da solução de múltiplos problemas de minimização independentes.

Em contrapartida, a Fig. 8b apresenta um caso em que a solução do método não é única, contrariando uma das condições necessárias para se obter o conjunto ótimo de Pareto. Além disso, mesmo com a variação dos coeficientes de ponderação, algumas soluções na região não-convexa do conjunto ótimo podem não ser alcançadas. Uma análise mais aprofundada dessas adversidades pode ser encontrada em Das e Dennis (1997). Adicionalmente, Deb (2001) e Collette e Siarry (2004) apresentam exemplos com resultados numéricos que ilustram com clareza cada um desses casos.

Figura 8 - Aplicação do Método da Soma Ponderada sob a perspectiva de dois tipos de problemas distintos.



Fonte: Adaptado de Miettinen (1998).

2.1.2 Método da Negociação

A fim de atenuar as dificuldades impostas pelo Método da Soma Ponderada, na tarefa de escolher valores para os pesos que melhor resultem em retas tangentes capazes de obter pontos mínimos em problemas não-convexos, Haimes et al. (1971) sugeriram o Método da Negociação (*ϵ -Constraint Method*). A sua concepção é baseada em um conceito muito utilizado na área de otimização, onde todos os objetivos de interesse, exceto um, são transformados em restrições do problema, e o objetivo mais importante é otimizado com relação às restrições impostas. Também denominado de Método ϵ -Restrito, a técnica estabelece que se deve definir um limite superior ϵ_j , para cada uma das funções a serem otimizadas e que serão convertidas em restrições do problema, exceto aquela que será, efetivamente, tomada como a função objetivo do problema de otimização.

Usando esta estratégia, o problema original é convertido em um problema de otimização mono-objetivo, tornando capaz a aplicação de uma técnica de minimização de objetivo único. Em termos práticos, seja $\bar{f}(\mathbf{x}) = f_p(\mathbf{x})$, a abordagem é definida como

$$\min_{\mathbf{x} \in \Omega} \bar{f}(\mathbf{x})$$

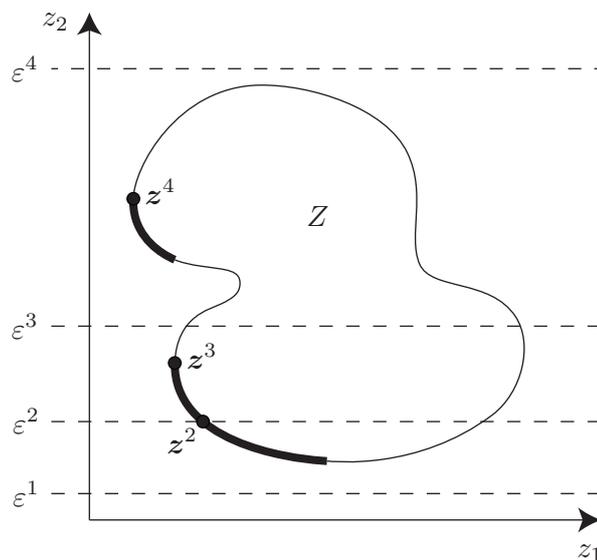
Sujeito a $f_j(\mathbf{x}) \leq \epsilon_j$,

tal que $j = 1, \dots, k$ e $j \neq p$, sendo k o número total de objetivos do problema original.

Basicamente, o conjunto ótimo de Pareto é obtido através da resolução reiterada do problema modificado, variando-se os valores de cada um dos limites superiores das restrições a cada execução. Dessa forma, é garantida a obtenção das soluções ótimas de Pareto para problemas não-convexos, desde que as soluções que constituem o conjunto sejam únicas.

Portanto, uma das vantagens do método é a capacidade de obtenção de diferentes soluções ótimas de Pareto, usando valores de ε distintos. Em termos de necessidade de definição de parâmetros, o algoritmo se assemelha ao Método da Soma Ponderada, uma vez que a localização da solução ótima depende de cada limite superior estipulado para os objetivos tomados como restrições. Essa dependência fica clara ao se analisar a Fig. 9, em que diferentes limites superiores são fixados para a função-objetivo f_2 , enquanto f_1 deve ser minimizado, e o conjunto ótimo de Pareto é ilustrado pelo segmento em negrito.

Figura 9 - Aplicação do Método da Negociação para diferentes valores de ε .



Fonte: Adaptado de Miettinen (1998).

Analisando cada uma das restrições, percebe-se que ε^1 foi escolhida de maneira inadequada, uma vez que a função foi limitada fora do intervalo em que se encontram as soluções ótimas, fazendo com que o procedimento de otimização correspondente não obtenha a solução esperada. Por sua vez, as restrições ε^3 e ε^4 não representam restrições efetivas para o conjunto de soluções ótimas. Com isso, as soluções obtidas são únicas, mas correspondem aos pontos extremos z^3 e z^4 , respectivamente. Por fim, a restrição ε^2 delimita o conjunto ótimo de Pareto, fazendo com que o procedimento de otimização produza a solução que é relativa ao ponto z^2 . Em vista disso, é notória a possibilidade de esforço computacional desperdiçado em circunstâncias onde o método de otimização seja executado de forma que os limites superiores definidos não produzam soluções viáveis.

2.1.3 Método de Otimização de Compromisso

O Método de Otimização de Compromisso (*Method of Weighted Metrics*) apresenta uma abordagem alternativa ao Método de Soma Ponderada dos objetivos, apresentado na Seção 2.1.1, em relação ao recurso de se combinar múltiplos objetivos em um único para tratar de problemas de otimização multi-objetivo. Considere os coeficientes de ponderação $w_i \geq 0$, para $i = 1, \dots, k$, sendo k a quantidade total de objetivos presentes no problema de otimização. Nessa abordagem, a necessidade de normalização dos coeficientes também se faz necessária, como mostrado na Eq. (2). Existem duas variantes bem difundidas do Método de Otimização de Compromisso, que se diferenciam entre si quanto às métricas usadas na sua idealização. Na primeira delas, o problema L_p ponderado é definido como $\min_{\mathbf{x} \in \Omega} \bar{f}(\mathbf{x})$, tal que

$$\bar{f}(\mathbf{x}) = \left(\sum_{i=1}^k w_i |f_i(\mathbf{x}) - z_i^*|^p \right)^{1/p}, \quad (4)$$

onde o parâmetro p pode assumir qualquer valor, com $p \geq 1$.

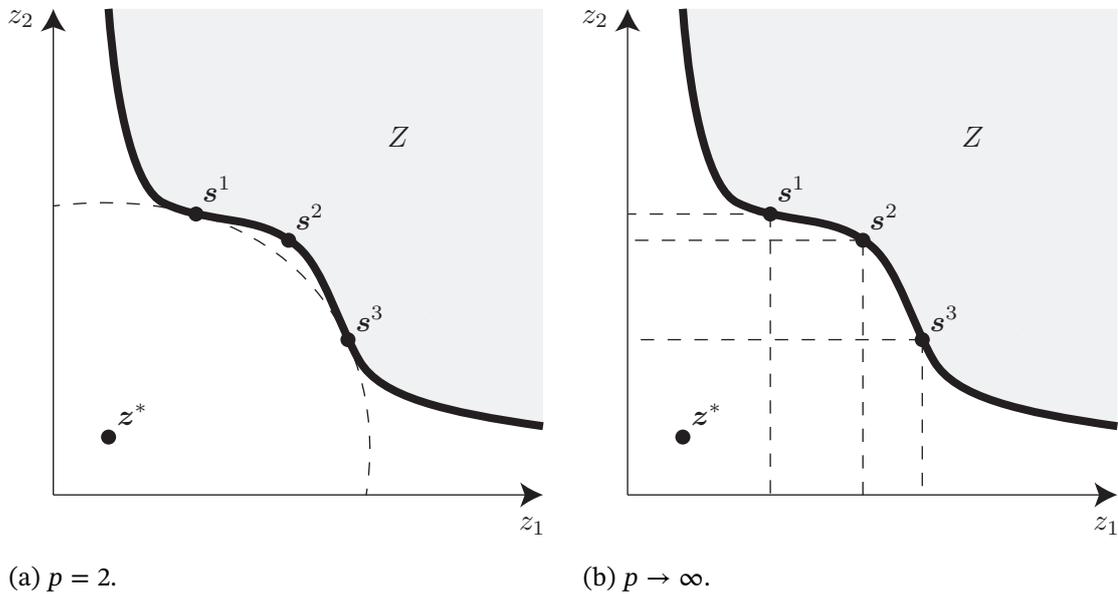
Quando o valor de p torna-se muito grande, o problema L_p ponderado é reduzido ao problema de minimização do desvio máximo. Nessa variante, conhecida como problema de Tchebycheff ponderado, utiliza-se a métrica L_∞ para derivar a formulação em que

$$\bar{f}(\mathbf{x}) = \max_{i=1, \dots, k} (w_i |f_i(\mathbf{x}) - z_i^*|),$$

a qual foi enunciada por Bowman (1976). Por outro lado, quando $p = 1$, o problema reduz-se àquele do Método da Soma Ponderada, exceto pela constante z^* , que pode ser obtida através da otimização independente de cada função-objetivo. Ainda, no caso em que $p = 2$, tem-se o Método de Mínimos Quadrados (MIETTINEN, 1998).

A Fig. 10 apresenta um esquema conceitual de aplicação do método, em um problema genérico, considerando dois casos, a saber, quando $p = 2$ e $p \rightarrow \infty$. Como pode ser percebido na Fig. 10a, a solução \mathbf{s}^2 do conjunto ótimo de Pareto, novamente representado pelo segmento em negrito, não pôde ser obtida, uma vez que o método é limitado quanto ao cálculo de soluções em regiões não-convexas, quando escolhe-se um valor de p reduzido. Por sua vez, à medida que o valor do parâmetro p aumenta, a estratégia torna-se capaz de alcançar regiões não-convexas, uma vez que o problema é convertido no cálculo do desvio máximo, como pode ser visto na Fig. 10b. Contudo, a desvantagem dessa escolha é que o problema se torna não-diferenciável (PARDALOS et al., 2017).

Figura 10 - Método de Otimização de Compromisso aplicado a um problema genérico, com diferentes valores do parâmetro p .



Fonte: Adaptado de Deb (2001).

2.2 Métodos a Priori

2.2.1 Método de Otimização Hierárquica

A ideia do Método de Otimização Hierárquica (*Lexicographic Ordering Method*) é simples e intuitiva, uma vez que consiste na organização das funções-objetivo em ordem de importância, as quais são otimizadas sequencialmente, iniciando por aquela tomada como mais relevante (COELLO COELLO, 2000). Assuma que os índices $i = 1, \dots, k$ indicam a ordem de relevância das funções-objetivo. Portanto, o primeiro problema mono-objetivo, equivalente à minimização da função mais importante, é definido como $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \Omega} \bar{f}(\mathbf{x})$, onde $\bar{f}(\mathbf{x}) = f_1(\mathbf{x})$.

Uma vez tendo calculado o ponto de mínimo ou minimizador de f_1 , com valor mínimo denotado por $f_1^* = f_1(\mathbf{x}^*)$, o próximo passo do método diz respeito à obtenção da solução referente ao segundo objetivo mais importante, tendo como restrição do novo problema mono-objetivo o valor de f_1^* . Portanto, seja $\bar{f}(\mathbf{x}) = f_2(\mathbf{x})$, o novo problema é definido como

$$\begin{aligned} & \min \bar{f}(\mathbf{x}) \\ & \text{Sujeito a } f_1(\mathbf{x}) = f_1^* . \end{aligned}$$

Esse procedimento iterativo é repetido até que sejam resolvidos os k problemas de mini-

mização, considerando $k - 1$ restrições relativas às soluções obtidas previamente. Então, a solução obtida na última etapa é tomada como o resultado ótimo do problema de otimização multi-objetivo original.

Incontestavelmente, esta técnica é uma das mais simples para a resolução de problemas desse tipo. No âmbito computacional, o procedimento de ordenamento de objetivos é muito simples de ser implementado e as soluções são obtidas gradualmente. Contudo, esse método tem uma desvantagem evidente: a escolha da sequência de objetivos pode ser uma tarefa arbitrária. Além disso, duas execuções distintas do Método de Otimização Hierárquico com sequências distintas de funções-objetivo não necessariamente produzem a mesma solução, visto que os problemas subsequentes terão restrições distintas em cada execução (COLLETTE; SIARRY, 2004).

2.2.2 Método da Programação de Metas

Como o próprio nome sugere, o conceito principal do Método da Programação de Metas (*Goal Programming Method*) é encontrar soluções que alcancem um alvo pré-definido para uma ou mais funções-objetivo. Se não forem obtidas soluções que atinjam essas metas em todos os objetivos, a tarefa passa a ser encontrar soluções que minimizem os desvios dos alvos (DEB, 2001). As noções de programação de metas foram originalmente introduzidas por Charnes et al. (1955), mas o termo “programação de metas” foi inserido por Charnes e Cooper (1961). Trata-se de um dos primeiros métodos expressamente criados para otimização multi-objetivo.

O compromisso do Método da Programação de Metas é minimizar os desvios positivo (δ_i^+) e negativo (δ_i^-) da função-objetivo, com respeito às metas impostas para cada um dos objetivos, definidas da forma $f_i(\mathbf{x}) \leq z_i$ (no caso de problemas de minimização), para $i = 1, \dots, k$, tal que k é a quantidade total de objetivos do problema de otimização. As variáveis de desvio podem ter sinais positivos ou negativos, dependendo do problema, de forma que satisfaçam $\delta_i = z_i - f_i(\mathbf{x})$. Então, as metas podem ser representadas, em relação a ambos os desvios, como $f_i(\mathbf{x}) + \delta_i^- - \delta_i^+ = z_i$. Tem-se, assim, um problema de otimização multi-objetivo no qual pode-se minimizar as variáveis de desvio (MIETTINEN, 1998).

Considerando esses conceitos, na estratégia clássica para otimização de problemas multi-objetivo utilizando a noção de programação de metas, também chamada de abordagem ponderada, deve-se minimizar a soma dos desvios ponderados. Matematicamente, essa formulação é definida através da incorporação dos desvios positivo e negativo à Eq. (4), com $p = 1$. Para isso, considere o desvio positivo como $\delta_i^+ = \max(0, f_i(\mathbf{x}) - z_i)$. Analogamente, o desvio negativo pode ser representado como $\delta_i^- = \max(0, z_i - f_i(\mathbf{x}))$. Neste

caso, a função a ser minimizada é dada por

$$\bar{f}(\mathbf{x}) = \sum_{i=1}^k w_i^- \delta_i^- + w_i^+ \delta_i^+$$

e o Problema de Programação de Metas Ponderado é definido da forma

$$\min \bar{f}(\mathbf{x})$$

$$\text{Sujeito a } f_i(\mathbf{x}) + \delta_i^- - \delta_i^+ = z_i .$$

para todo $i = 1, \dots, k$. Aqui, são atribuídos pesos diferentes para cada um dos desvios, denotados como w^- e w^+ , mantendo-se a correspondência de sinais. Assim como nos demais métodos, essa estratégia pode apresentar dificuldades na obtenção de soluções ótimas localizadas nas concavidades de problemas não-convexos. Para uma análise mais apurada e exemplos numéricos, consulte as obras de Miettinen (1998), Coello Coello (2000) e Deb (2001).

2.3 Métodos Progressivos

2.3.1 Método de Fandel

Como visto anteriormente, uma das limitações do Método da Soma Ponderada referia-se à escolha adequada dos valores dos pesos utilizados na ponderação de cada objetivo. Visando a necessidade de uma escolha mais pragmática, o Método de Fandel (*Fandel Method*) foi idealizado com o objetivo de auxiliar nessa tarefa. Neste contexto, considere a formulação fundamental do Método da Soma Ponderada, representada pela Eq. (3), assim como as condições estabelecidas na Seção 2.1.1. Diferentemente da proposta original, quando a variação dos valores de w permitiam a obtenção gradativa das soluções que constituíam o conjunto ótimo de Pareto, no Método de Fandel pressupõe-se que os coeficientes de ponderação sejam desconhecidos. A partir desse fato, a estratégia descrita abaixo obtém possíveis valores mais adequados para os pesos do que aqueles tomados arbitrariamente (ESCHERNAUER et al., 1990; COLLETTE; SIARRY, 2004).

Inicialmente, computa-se o minimizador referente a cada objetivo independentemente (considerando as possíveis restrições do problema original), os quais são representados por $\hat{\mathbf{x}}^i$, para $i = 1, \dots, k$, tal que k representa o número total de objetivos do problema de otimização. Esses resultados devem ser arranjados em uma estrutura, chamada de vetor de funções-objetivo ideal e denotado por $\hat{\mathbf{f}} = (\hat{f}_1, \dots, \hat{f}_k)$. Neste caso, $\hat{f}_i = f_i(\hat{\mathbf{x}}^i)$. Em seguida, é possível definir uma matriz quadrada \mathbf{B} , de ordem k , composta pelo cruzamento

de cada $\hat{\mathbf{x}}^i$ com os objetivos do problema. Portanto,

$$\mathbf{B} = \begin{pmatrix} f_1(\hat{\mathbf{x}}^1) & \cdots & f_i(\hat{\mathbf{x}}^1) & \cdots & f_k(\hat{\mathbf{x}}^1) \\ \vdots & \ddots & \vdots & & \vdots \\ f_1(\hat{\mathbf{x}}^i) & \cdots & f_i(\hat{\mathbf{x}}^i) & \cdots & f_k(\hat{\mathbf{x}}^i) \\ \vdots & & \vdots & \ddots & \vdots \\ f_1(\hat{\mathbf{x}}^k) & \cdots & f_i(\hat{\mathbf{x}}^k) & \cdots & f_k(\hat{\mathbf{x}}^k) \end{pmatrix}. \quad (5)$$

Através dessa disposição, fica claro que o vetor de funções-objetivo ideal compõe a diagonal principal da matriz \mathbf{B} .

Imediatamente, calcula-se o vetor de médias dos objetivos (\mathbf{f}^m), considerando todas as soluções individualmente computadas, por meio de

$$f_j^m = \frac{1}{k} \sum_{i=1}^k B_{i,j},$$

para $j = 1, \dots, k$. Então pode ser definido um novo conjunto de problemas de otimização mono-objetivo, que assume a forma

$$\tilde{\mathbf{x}}^i = \begin{cases} \arg \min f_i(\mathbf{x}) \\ \text{Sujeito a } f_i(\mathbf{x}) \leq f_i^m \end{cases}, \quad (6)$$

para $i = 1, \dots, k$. Em seguida, constrói-se uma nova matriz, agora denotada como $\bar{\mathbf{B}}$ mas análoga àquela da Eq. (5), isto é, composta pela combinação de cada objetivo com $\tilde{\mathbf{x}}^i$, os quais são obtidos pela solução dos k problemas de otimização independentes da Eq. (6). Então, define-se um hiperplano dado por $\bar{\mathbf{B}}\mathbf{a} = \mathbf{c}$, onde $\mathbf{a}, \mathbf{c} \in \mathbb{R}^k$ são as incógnitas do sistema de equações. Neste caso, o vetor $\mathbf{c} = (c_1, \dots, c_1)^T$ e \mathbf{a} deve ser normalizado, o que garante que o sistema a ser resolvido não tenha nenhum grau de liberdade, como apontado por Eschenauer et al. (1990).

O objetivo é encontrar um valor para c_1 tal que o hiperplano $\bar{\mathbf{B}}\mathbf{a} = \mathbf{c}$ seja tangencial ao espaço de valores da função-objetivo. Este problema de otimização linear leva a uma solução em que o vetor \mathbf{a} corresponde aos valores desejados dos coeficientes de ponderação \mathbf{w} para o Método da Soma Ponderada, que pode ser aplicado em seguida, possibilitando a escalarização do problema de otimização multi-objetivo. De maneira geral, a principal desvantagem deste método, que é comum a muitos métodos progressivos, é a capacidade de se obter apenas uma solução a cada execução (COLLETTE; SIARRY, 2004).

2.3.2 Método STEP

O Método STEP, originalmente projetado por Benayoun et al. (1971) para a maximização de problemas de otimização multi-objetivo lineares, pode ser estendido para problemas não-lineares. Inicialmente, admite-se o seguinte problema:

$$\begin{aligned} & \min \alpha \\ & \text{Sujeito a } \mathbf{w}(\mathbf{f}(\mathbf{x}) - \bar{\mathbf{y}})^T \leq \alpha \\ & \mathbf{f}(\mathbf{x}) \leq \bar{\mathbf{y}}, \end{aligned}$$

onde $\alpha \in \mathbb{R}$, $\bar{\mathbf{y}}$ equivale ao vetor que contém os limites superiores utilizados para restringir o espaço de busca e \mathbf{w} são os pesos.

Com esta formulação, a determinação dos pesos é novamente necessária para a resolução do problema proposto. Para realizar essa tarefa, inicialmente é preciso calcular a matriz \mathbf{B} , assim como na Eq. (5), por meio do mesmo procedimento descrito na seção correspondente. Apesar disso, no Método STEP os coeficientes de ponderação são computados respeitando uma relação que leva em conta os valores extremos dos resultados relacionados às soluções dos problemas de otimização de forma independente, isto é,

$$w_i = \frac{v_j}{\sum_{i=1}^k v_i}, \quad (7)$$

tal que

$$v_j = \frac{\max_{i=1,\dots,k} \{B_{i,j}\} - B_{j,j}}{B_{j,j} \min_{i=1,\dots,k} \{B_{i,j}\}}, \quad (8)$$

para $i = 1, \dots, k$ e $j = 1, \dots, k$. Nessa metodologia, os valores dos pesos são todos positivos e têm que ser normalizados, como nos demais métodos que contam com técnica de ponderação (COLLETTE; SIARRY, 2004).

Através dessa forma de cálculo e atribuição de pesos, o método tende a definir ponderações mais elevadas às funções-objetivo que apresentem maiores diferenças entre os valores mínimo e máximo das linhas de \mathbf{B} . Após a definição dos pesos, é calculada a solução inicial ($\tilde{\mathbf{f}}$) do problema utilizando o Método da Soma Ponderada. Esse resultado é comparado ao vetor de funções-objetivo ideal, semelhante àquele definido na Seção 2.3.1. Se a comparação não for satisfatória em relação a algum dos valores, é possível relaxar a restrição correspondente, fazendo

$$\bar{y}_j = \tilde{f}_j + \Delta f_j,$$

para um dado $j \in \{1, \dots, k\}$, onde Δf_j é um incremento arbitrário. Portanto, seja $\bar{f}(\mathbf{x}) = \mathbf{w}(\mathbf{f}(\mathbf{x}) - \bar{\mathbf{y}})^T$. O problema final a ser resolvido, após os ajustes das restrições, é definido como

$$\begin{aligned} & \min \bar{f}(\mathbf{x}) \\ & \text{Sujeito a } \mathbf{f}(\mathbf{x}) \leq \bar{\mathbf{y}}, \end{aligned}$$

cujos pesos são calculados pela Eq. (7).

Apesar de ser notável a forma de cálculo dos pesos, sem que haja a necessidade de escolhas arbitrárias, o método apresenta um grande limitante em relação à sua aplicação em alguns tipos de problemas: as funções-objetivo não podem resultar em zero, pois esse fato ocasionaria uma singularidade na Eq. (8). Além disso, o método também carece de uma estratégia de obtenção de soluções simultâneas, sendo que cada valor ótimo do conjunto de Pareto deve ser computado individualmente, como no Método de Fandel, apresentado anteriormente (COLLETTE; SIARRY, 2004).

2.4 Métodos Baseados em Metaheurísticas

Como pôde ser visto na breve descrição de cada um dos métodos clássicos apresentados até aqui, a proposta de se converter o problema de otimização multi-objetivo em um problema mono-objetivo foi sempre a base para a elaboração de diferentes estratégias. Esse processo de escalarização resulta em um conjunto de problemas de objetivo único, que podem ser resolvidos através de métodos de otimização utilizados para esse propósito, os quais existem em profusão na literatura (uma boa referência é a obra de Bazaraa et al. (2006)). Nesses casos, espera-se que a solução ótima obtida seja um ponto pertencente ao conjunto ótimo de Pareto. Com isso, teoricamente o conjunto ótimo de Pareto completo é obtido através da solução reiterada de problemas mono-objetivo, com a variação do parâmetro utilizado na escalarização do problema. Isso significa que, para a obtenção de n soluções distintas, é necessário que o procedimento de conversão do problema multi-objetivo e otimização do esquema mono-objetivo correspondente seja realizado, pelo menos, n vezes.

Apesar disso, ainda existe o fato de que algumas metodologias têm limitações quanto ao cálculo de soluções ótimas localizadas em regiões não-convexas da curva de Pareto, dentre as quais pode-se citar o Método da Soma Ponderada, o Método da Negociação e o Método de Otimização de Compromisso, sendo que neste último a limitação decorre da escolha do parâmetro p . Nestas circunstâncias, não importa o valor dos parâmetros utilizados no processo de escalarização, pois os métodos nunca conseguirão alcançar essas regiões e obter todas as soluções do conjunto ótimo de Pareto, independentemente da quantidade de execuções.

Ainda no que diz respeito às limitações dos métodos clássicos, alguns também necessitam de certo conhecimento prévio do comportamento da função-objetivo que auxilie na tomada de decisões quanto à escolha de parâmetros. Tendo como exemplo o Método da Negociação, a escolha de um conjunto de valores para ε , dentro de um intervalo predefinido e com elementos igualmente espaçados entre si, não garante a obtenção de todas as soluções de um dado problema, uma vez que os valores podem estar limitando uma região da função-objetivo na qual não há soluções ótimas. Ademais, um vetor de parâmetros uniformemente espaçados não garante a obtenção de soluções ótimas com disposição semelhante, em virtude da não-linearidade de alguns problemas.

Apesar dessas deficiências, é evidente que os métodos clássicos de otimização multi-objetivo têm diversas vantagens. Possivelmente, a maior delas é em relação aos teoremas que garantem a convergência de várias destas metodologias, sob condições específicas. Por mais que não tenham sido discutidos nesta tese, por conta do escopo que se destina esta revisão, a concepção da maioria dos métodos descritos neste capítulo é fundamentada por tais teoremas. Em vista dessas características, esses métodos são altamente dependentes das técnicas escolhidas para resolver os problemas de otimização mono-objetivo associados e computar as verdadeiras soluções ótimas de Pareto (DEB, 2001).

Nesse contexto, uma das tarefas essenciais de uma técnica de otimização para problemas multi-objetivo é a capacidade de obtenção de tantas soluções ótimas quanto for possível, as quais devem compor uma discretização do conjunto ótimo de Pareto, através de uma única execução do algoritmo. Outra tarefa fundamental, em vista dos métodos clássicos, é buscar soluções ótimas de Pareto com uma boa diversidade de valores das funções-objetivo. Essas duas características—quantidade e diversidade—compõem necessidades cruciais para métodos de otimização multi-objetivo e são a base para elaboração de uma classe muito importante: os métodos de otimização baseados em metaheurísticas.

As metaheurísticas são uma classe de algoritmos caracterizados pela auto-adaptação e utilizados para encontrar soluções de problemas através de técnicas que mimetizam os processos inteligentes e comportamentos observados na natureza. Entre todos os paradigmas computacionais inspirados na natureza, a computação evolutiva é o mais influente. Trata-se de uma abstração computacional para obter as melhores soluções possíveis em um enorme espaço de solução baseado no princípio de sobrevivência do indivíduo mais apto. Algoritmos evolutivos são uma classe de técnicas efetivas de otimização global para muitos problemas complexos de engenharia (DU; SWAMY, 2016; SIARRY, 2016; LOBATO; STEFFEN JR, 2017). A seguir são apresentadas duas metaheurísticas para otimização de problemas multi-objetivo, baseados nos conceitos da computação evolutiva. Métodos deste tipo serão empregados na resolução de problemas de otimização multi-objetivo com incerteza.

2.4.1 Algoritmo Genético de Ordenamento Não-dominado

Assim como na maioria dos métodos de mesmo gênero, o processo evolutivo do Algoritmo Genético de Ordenamento Não-dominado (NSGA-II) é baseado no conceito de dominância, abordado no Cap. 1. Na maioria dos aspectos, este algoritmo não tem muita semelhança com o NSGA original, proposto por Srinivas e Deb (1994), mas os autores utilizaram o nome NSGA-II para destacar sua gênese e conceituação (DEB, 2001). A técnica foi proposta por Deb et al. (2000) e, mais tarde, o seu prestígio foi consolidado por Deb et al. (2002). A evolução da técnica se deu, principalmente, motivada por dois aspectos fundamentais, presentes no algoritmo predecessor: alta complexidade computacional no procedimento de ordenamento não-dominado e falta de elitismo. Este segundo aspecto diz respeito à evolução dos indivíduos mais aptos da população. O elitismo é introduzido no algoritmo, de maneira sumária, comparando-se os descendentes com as soluções utilizadas para originar tais indivíduos e mantendo as melhores soluções, com mais capacidade de se desenvolverem no decorrer das gerações.

Haja vista as limitações já apontadas anteriormente, em relação ao NSGA original, uma outra deficiência dessa técnica era a necessidade de definição do parâmetro de compartilhamento σ_{share} , uma vez que o algoritmo utilizava a abordagem de função de compartilhamento (para mais detalhes, consulte Srinivas e Deb (1994)). No método NSGA-II, o algoritmo não requer nenhuma definição de parâmetros desse tipo, uma vez que faz uso da abordagem de comparação da multidão (com tradução livre, a partir do termo *crowded-comparison approach*) para definir a diversidade da população. Em suma, o algoritmo NSGA-II é composto por dois procedimentos principais, que se alternam durante o processo de otimização. Ambos são detalhados a seguir.

2.4.1.1 Construção de Frentes de Pareto Ordenadas

Inicialmente, o algoritmo realiza o processo de ordenamento das frentes de Pareto não-dominadas, começando pelo cálculo de duas grandezas importantes, a saber, o contador de dominância, denotado por n_p , que representa a quantidade de soluções que dominam uma solução \mathbf{p} , e o conjunto de soluções \mathbf{S}_p , que são dominadas por \mathbf{p} . Na segunda parte do processo de ordenamento, o método realiza a tentativa de desassociar do conjunto de soluções dominantes, aquelas que farão parte da segunda frente de Pareto. Para isso, toma-se iterativamente cada solução \mathbf{p} com contador de dominância $n_p = 0$ e, para cada um desses pontos, uma solução \mathbf{q} contida no conjunto \mathbf{S}_p é visitada, sendo que n_p é subtraído de uma unidade. Dessa forma, caso haja algum membro de \mathbf{S}_p no qual o contador de dominância seja igual a zero, naturalmente essa solução pertence à segunda frente de Pareto não-dominada e, portanto, esse ponto é inserido no conjunto \mathbf{Q} . Então, esse processo itera-

tivo é efetuado até que todas as frentes sejam devidamente identificadas. O Alg. 1 mostra uma possibilidade de implementação para a técnica descrita.

Algoritmo 1 - Ordenamento das frentes de Pareto não-dominadas.

FUNÇÃO NONDOMINATEDSORT (P)

```

para cada  $p \in P$ 
   $S_p \leftarrow \emptyset$ 
   $n_p \leftarrow 0$ 
  para cada  $q \in P$ 
    se  $p < q$ 
       $S_p \leftarrow S_p \cup q$       {Adiciona  $q$  ao conjunto de soluções dominadas por  $p$ .}
    senão
       $n_p \leftarrow n_p + 1$       {Incrementa o contador de dominância de  $p$ .}
    fim se
  fim para cada
  se  $n_p = 0$ 
     $p_{\text{rank}} \leftarrow 1$ 
     $\mathcal{F}(1) \leftarrow \mathcal{F}(1) \cup p$       { $p$  pertence à primeira frente.}
  fim se
fim para cada
 $i \leftarrow 1$ 
enquanto  $\mathcal{F}(i) \neq \emptyset$ 
   $Q \leftarrow \emptyset$       {Armazena os membros da próxima frente.}
  para cada  $p \in \mathcal{F}(i)$ 
    para cada  $q \in S_p$ 
       $n_q \leftarrow n_q - 1$ 
      se  $n_q = 0$ 
         $q_{\text{rank}} \leftarrow i + 1$ 
         $Q \leftarrow Q \cup q$       { $q$  pertence à próxima frente.}
      fim se
    fim para cada
  fim para cada
   $i \leftarrow i + 1$ 
   $\mathcal{F}(i) \leftarrow Q$ 
fim enquanto
retornar  $\mathcal{F}$ 
FIM FUNÇÃO

```

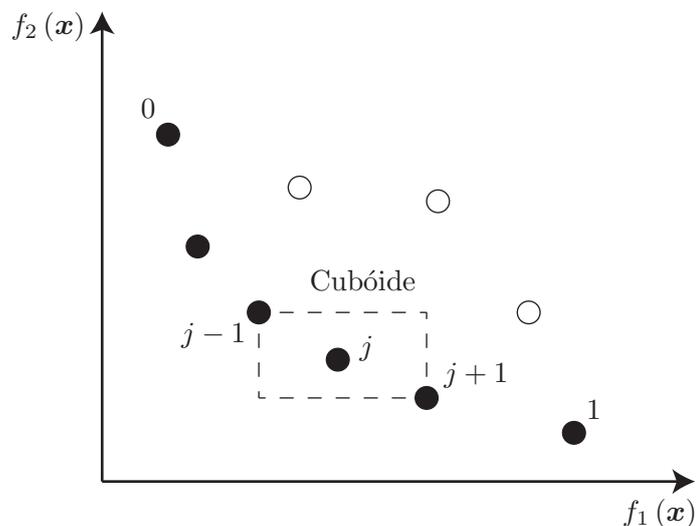
2.4.1.2 Cálculo das Distâncias de Aglomeração

Na etapa seguinte do algoritmo, onde avalia-se as distâncias de aglomeração dos indivíduos, o objetivo é preservar a diversidade da população. Na computação evolutiva, o conceito de diversidade da população é de suma importância pois, uma população de indi-

víduos com razoável grau de diversidade tende a gerar descendentes com maior aptidão a se desenvolverem ao longo do tempo. No âmbito das metaheurísticas e, mais especificamente, no sentido computacional, essa compreensão da diversidade e manutenção de uma espécie pode ser conceitualmente convertida na possibilidade de aumento das chances de obtenção de soluções ótimas de um problema de otimização, através da expansão da capacidade de exploração do domínio viável.

Para apresentar esse conceito, primeiro é preciso conhecer a definição de estimativa da densidade da população ao redor de uma solução. Para calcular essa grandeza, deve-se computar a distância média de dois pontos em ambos os lados de uma dada solução, ao longo de cada um dos objetivos. Essa quantidade serve como uma estimativa do perímetro do cuboide formado usando os vizinhos mais próximos como vértices. Na Fig. 11, essa medida em relação à j -ésima solução, situada na frente de Pareto cujos pontos estão marcados com círculos sólidos, é o comprimento médio lateral do cuboide (mostrado com uma caixa tracejada).

Figura 11 - Esquema geométrico do procedimento de aglomeração de soluções para problema com dois objetivos.



Legenda: Pontos marcados em círculos preenchidos são soluções da mesma frente não-dominada.

Fonte: Adaptado de Deb et al. (2002).

Sob a condição de ordenamento dos indivíduos da população em ordem crescente, relativamente ao valor da função-objetivo, o cômputo da distância de aglomeração se refere à diferença normalizada absoluta dos valores das funções de duas soluções adjacentes. Esse cálculo é relativo a todos os objetivos do problema. O valor global da distância de aglomeração é calculado como a soma dos valores de distância individuais correspondentes a

cada objetivo, de tal forma que cada um deles deve ser normalizado antes de se calcular o valor final. Uma solução com um valor menor dessa medida de distância é, de certo modo, mais rodeada por outras soluções. O procedimento computacional para o cálculo da distância de aglomeração, através da abordagem de comparação da multidão, é mostrado no Alg. 2. No pseudocódigo apresentado, \mathcal{S} refere-se às soluções da geração corrente e a medida calculada pelo método é armazenada no vetor \mathcal{S}_d . Para cada função-objetivo, as soluções extremas (aquelas com o menor e maior valores) são atribuídas um valor de distância infinito. Além disso, o termo $\mathcal{S}(i, j)$ refere-se ao valor do i -ésimo objetivo com respeito ao j -ésimo indivíduo.

Algoritmo 2 - Distância de aglomeração de soluções.

PROCEDIMENTO CROWDINGDISTANCE (\mathcal{S})

```

 $n_s \leftarrow \text{size}(\mathcal{S})$                                 {Número de soluções.}
 $\mathcal{S}_d \leftarrow \emptyset$                             {Vetor de distâncias.}
para  $i$  de 1 até  $k$                                 {Para cada um dos  $k$  objetivos.}
     $\mathcal{S} \leftarrow \text{sort}(\mathcal{S}, f_i)$                 {Ordenamento em relação ao valor dos objetivos.}
     $\mathcal{S}_d(1) \leftarrow \infty$ 
     $\mathcal{S}_d(n_s) \leftarrow \infty$ 
    para  $j$  de 2 até  $n_s - 1$ 
         $\mathcal{S}_d(j) \leftarrow \mathcal{S}_d(j) + \frac{\mathcal{S}(i, j+1) - \mathcal{S}(i, j-1)}{f_i^{\max} - f_i^{\min}}$ 
    fim para cada
fim para cada
FIM PROCEDIMENTO

```

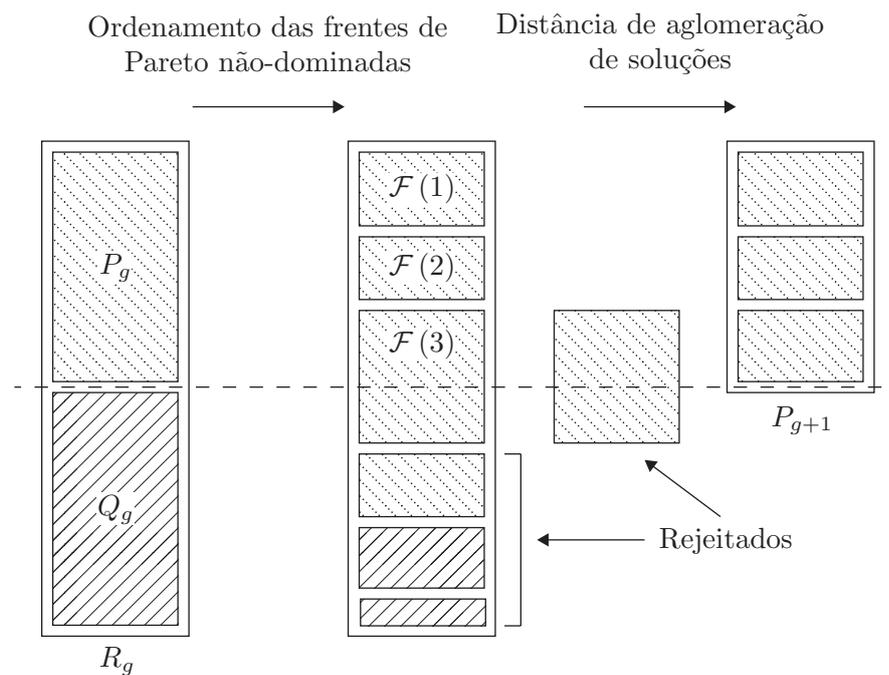
2.4.1.3 Procedimento Principal

No procedimento principal da metaheurística, deve-se criar uma população \mathbf{P} , que servirá como parâmetro de entrada do algoritmo, além do número máximo de gerações, simbolizado por g_{\max} . A população inicial é utilizada para gerar uma população de descendentes, denotada por \mathbf{Q} , a qual será combinada com a população originária, gerando um novo conjunto, representado por \mathbf{R} , com dimensão $2N$. O próximo passo é a realização do ordenamento dessa nova estrutura, através da função NONDOMINATEDSORT, como descrito no Alg. 1. Como todos os indivíduos estão inseridos no conjunto \mathbf{R} , garante-se o elitismo da técnica. Em seguida, verifica-se a quantidade de indivíduos situados em cada uma das frentes, representadas por $\mathcal{F}(i)$. Se o tamanho de $\mathcal{F}(1)$ for menor que N , definitivamente escolhe-se todos os membros do conjunto $\mathcal{F}(1)$ para a nova população, representada por \mathbf{P}_{g+1} . Os membros restantes da nova população são escolhidos de frentes não-dominadas subsequentes na ordem de sua classificação. Assim, as soluções do conjunto $\mathcal{F}(2)$ são es-

colhidas a seguir, seguidas pelas soluções do conjunto $\mathcal{F}(3)$ e assim por diante.

A fim de que a nova população tenha exatamente N membros, os indivíduos da última frente ordenada de \mathcal{F} são escolhidos para compor a população da geração seguinte, por meio dos critérios definidos no processo de ordenamento, com base no operador \prec_n . Chamado de operador de comparação da multidão, as soluções são ordenadas com base na seguinte regra: entre duas soluções com diferentes graus de posicionamento (*rank*), deve-se dar preferência à solução com a classificação mais baixa. Caso contrário, se ambas as soluções pertencerem à mesma frente, então deve-se selecionar a solução que está localizada em uma região menos populosa, cuja indicação é obtida pela execução do procedimento CROWDINGDISTANCE, mostrado no Alg. 2. Então, os indivíduos selecionados passam à próxima geração, se juntando àqueles que já compõem P_{g+1} . O procedimento principal do NSGA-II está descrito no Alg. 3 e a Fig. 12 apresenta um esquema da evolução das frentes de Pareto, que reforça a compreensão do método.

Figura 12 - Funcionamento do processo evolutivo do NSGA-II.



Fonte: Adaptado de Deb et al. (2002).

Os indivíduos descendentes daqueles escolhidos para compor a nova geração são modificados, por meio de operadores genéticos, representados no Alg. 3 pela função MAKENEWPOP. Nessa sub-rotina, estão compreendidas três abordagens muito utilizadas nos algoritmos evolutivos, conhecidas como mutação, cruzamento e seleção, para modificação de indivíduos através de conceitos genéticos. No método NSGA-II, utiliza-se os conceitos de mutação *bit-a-bit* e cruzamento de ponto único para gerar descendentes a partir dos mem-

Algoritmo 3 - Laço de repetição principal do NSGA-II.

PROCEDIMENTO NSGAI (P, g_{\max})

para g **de** 1 **até** g_{\max}

$R_g \leftarrow P_g \cup Q_g$

{Combinação entre genitores e descendentes.}

$P_{g+1} \leftarrow \emptyset$

$\mathcal{F} \leftarrow \text{NONDOMINATEDSORT}(R_g)$

$i \leftarrow 1$

enquanto $\text{size}(P_{g+1}) + \text{size}(\mathcal{F}(i)) \leq N$

$\text{CROWDINGDISTANCE}(\mathcal{F}(i))$

$P_{g+1} \leftarrow P_{g+1} \cup \mathcal{F}(i)$

$i \leftarrow i + 1$

fim enquanto

$\text{sort}(\mathcal{F}(i), <_n)$

{Composição em ordem decrescente.}

para j **de** 1 **até** $N - \text{size}(P_{g+1})$

$P_{g+1} \leftarrow P_{g+1} \cup \mathcal{F}(i, j)$

fim para

$Q_{g+1} \leftarrow \text{MAKENEWPOP}(P_{g+1})$

{Execução dos operadores genéticos.}

fim para

FIM PROCEDIMENTO

bro mais aptos da geração anterior. Os melhores descendentes permanecem na população e são submetidos ao procedimento descrito anteriormente, por meio da seleção binária pela técnica de “torneio”. Essas rotinas não são descritas aqui, por se tratar de estratégias muito difundidas no meio das metaheurísticas e amplamente utilizadas em grande parte dos métodos dessa classe. Entretanto, os seus detalhes podem ser vistos em Deb e Agrawal (1995) e Beyer e Deb (2001).

Na técnica anterior, o NSGA, a ordem de complexidade do algoritmo, relatada pelos autores, era de $\mathcal{O}(kN^3)$, sendo k o número de objetivos e N a quantidade de indivíduos da população. Isto o torna computacionalmente custoso para grandes tamanhos populacionais. Essa alta complexidade surge devido à necessidade de processamento envolvida no procedimento de ordenamento das frentes de Pareto não-dominadas, em todas as gerações. Por sua vez, no NSGA-II, esse mesmo procedimento foi reduzido, em relação à ordem de complexidade, para $\mathcal{O}(k(2N)^2)$. Além disso, os autores mostram que o processo de cálculo da distância de aglomeração das soluções tem complexidade de $\mathcal{O}(k(2N)\log_2(2N))$ e o ordenamento utilizando o operador $<_n$ tem ordem de complexidade de $\mathcal{O}(2N\log_2(2N))$, no pior caso. Portanto, a complexidade geral do NSGA-II é de $\mathcal{O}(kN^2)$.

Uma das desvantagens do método, apontada por Deb (2001), diz respeito à utilização da abordagem de comparação da multidão. Mesmo que essa técnica tenha incorporado vantagens importantes ao método, em relação ao NSGA, a sua utilização faz com que o algoritmo perca a propriedade de convergência. Isso se dá pois, nas últimas gerações, o método pode ser capaz de substituir membros da primeira frente não-dominada de Pareto,

por indivíduos dominados, fazendo com que a solução se deteriore.

2.4.2 Evolução Diferencial Auto-adaptativa para Otimização Multi-objetivo

Diversas técnicas de otimização existentes na literatura consideram valores constantes para alguns de seus parâmetros de controle durante o processo evolutivo. Naturalmente, o comportamento dos indivíduos em um meio tende a ser adaptativo, em relação às necessidades impostas para sobrevivência. Além disso, esse comportamento pode ser diferente com o decorrer das gerações, evidenciando a tentativa de aumento da taxa de manutenção e evolução da população. Evidentemente, essa ideia deve se estender às metaheurísticas, de maneira a prover mais flexibilidade e aumento de desempenho na solução de problemas de otimização cada vez mais complexos.

Dentre as diversas metaheurísticas sugeridas para solução de problemas de otimização, a Evolução Diferencial (DE) é, provavelmente, uma das mais conceituadas. Proposta por Storn e Price (1997), a técnica foi originalmente elaborada para a solução de problemas mono-objetivo. Com o passar dos anos, a literatura especializada pôde presenciar uma profusão de aplicações da técnica em problemas dos mais variados tipos. Naturalmente, com a notória potencialidade da técnica, foram propostas diversas variantes da DE, tanto com o intuito de torná-la adaptativa em relação aos seus parâmetros de controle, quanto motivadas pela necessidade de estender a sua formulação ao contexto dos problemas de otimização multi-objetivo.

Foi realizada por Lobato e Steffen Jr (2017) uma revisão bibliográfica a respeito das técnicas de otimização multi-objetivo, baseadas nos conceitos da DE. Na mesma obra, os autores propuseram uma técnica de otimização considerando parâmetros auto-adaptativos, intitulada Evolução Diferencial Auto-adaptativa para Otimização Multi-objetivo (SA-MODE). O algoritmo foi proposto, a fim de estender as funcionalidades da DE canônica, objetivando a solução de problemas de otimização multi-objetivo através da incorporação de três operadores ao algoritmo original, os quais são: os mecanismos de classificação hierárquica, o cálculo da distância de aglomeração e a exploração dos candidatos potenciais à solução da vizinhança.

2.4.2.1 Fundamentação do Algoritmo

A metaheurística tem como base o procedimento de obtenção das curvas de Pareto realizado pelo algoritmo MODE, proposto pelos mesmos autores, mas sem o emprego de técnicas de variação de parâmetros. A estrutura fundamental do MODE é composta pelos procedimentos computacionais descritos a seguir. Inicialmente, uma população de N indi-

vídus é gerada aleatoriamente no domínio de busca viável. Dessas soluções iniciais, são eliminadas aquelas dominadas, ordenando a população em F_k frentes não-dominadas de Pareto, através da técnica mostrada no Alg. 1. Após esse procedimento de ordenamento, são selecionadas três soluções ao acaso. Desse conjunto de indivíduos, gera-se um descendente que irá compor a população. Reiteradamente, esse processo de seleção de três soluções e geração de uma nova é executado até que sejam produzidos N novos pontos.

De posse da nova população, o MODE emprega a Estratégia de Evolução da Vizinhaça (HU et al., 2005 apud LOBATO; STEFFEN JR, 2017). O propósito desta prática é criar vizinhos ao redor de cada uma das $2N$ potenciais soluções, com a expectativa de que pequenas perturbações nestes candidatos possam produzir vizinhos mais aptos, que sejam dominantes em relação àqueles da população original. O esquema de geração de novos pontos segue a relação

$$\mathcal{X}(\mathbf{x}) = \left(\mathbf{x} - \frac{1}{2} \mathbf{D}_k^{(g)}, \mathbf{x} + \frac{1}{2} \mathbf{D}_k^{(g)} \right),$$

onde

$$\mathbf{D}_k^{(g)} = \frac{k}{R} (\mathbf{L}_{\text{sup}} - \mathbf{L}_{\text{inf}}).$$

Nessa estratégia, k representa a frente de Pareto correspondente, R é o número de pseudo-frentes definidas pelo usuário, \mathbf{L}_{sup} e \mathbf{L}_{inf} simbolizam os limites superior e inferior de busca das variáveis, respectivamente. Portanto, $\mathbf{D}_k^{(g)} \in \mathbb{R}^n$, sendo n a dimensão do problema, e varia de acordo com a geração g , uma vez que são obtidas diferentes frentes de Pareto—em quantidade de conjuntos e em número de indivíduos—a cada iteração do algoritmo.

O número de indivíduos em cada pseudo-frente de Pareto pode ser definido proporcionalmente a

$$\eta_k = r\eta_{k-1},$$

com $k = 2, \dots, R$, tal que η_k denota o número de indivíduos presentes na k -ésima frente de Pareto e o parâmetro de redução $r \in \mathbb{R}$, com $0 < r < 1$. A cada geração, o valor de η_k pode ser estimado, considerando uma população de N indivíduos, através de

$$\eta_k = \frac{1-r}{1-r^R} r^{k-1} N.$$

Como $r < 1$, o número de indivíduos na primeira pseudo-frente de Pareto é o maior. Portanto, cada pseudo-frente subsequente tem um número de candidatos exponencialmente menor. Com isso, um valor pequeno atribuído a r resulta em mais potenciais soluções na primeira pseudo-frente e, por conseguinte, enfatiza a pesquisa local. Por outro lado, tomar um valor de r superior resulta em mais candidatos na última pseudo-frente e, assim,

tendência a uma busca global (HU et al., 2005 apud LOBATO; STEFFEN JR, 2017).

Os vizinhos gerados, que compõem um conjunto à parte da população atual da geração, são classificados de acordo com o critério de dominância. Os indivíduos da população de vizinhos, classificados como não-dominados, são agrupados aos demais para compor a população geral. Esse novo conjunto de potenciais soluções é, novamente, submetido ao mesmo procedimento de classificação, mencionado anteriormente. Se a quantidade de pontos que satisfazem tais critérios for maior do que o número máximo de indivíduos que devem compor a população, sendo tal parâmetro definido pelo usuário, o conjunto é truncado conforme a técnica de distância de aglomeração de soluções, similar ao que é realizado no método NSGA-II, detalhada no Alg. 2. De maneira geral, a metaheurística MODE realiza a obtenção da curva de Pareto através da execução reiterada do procedimento descrito. Os detalhes do algoritmo MODE e resultados comparativos podem ser vistos em Lobato e Steffen Jr (2011). O fluxograma conceitual do método é apresentado na Fig. 13, com os passos principais da metaheurística.

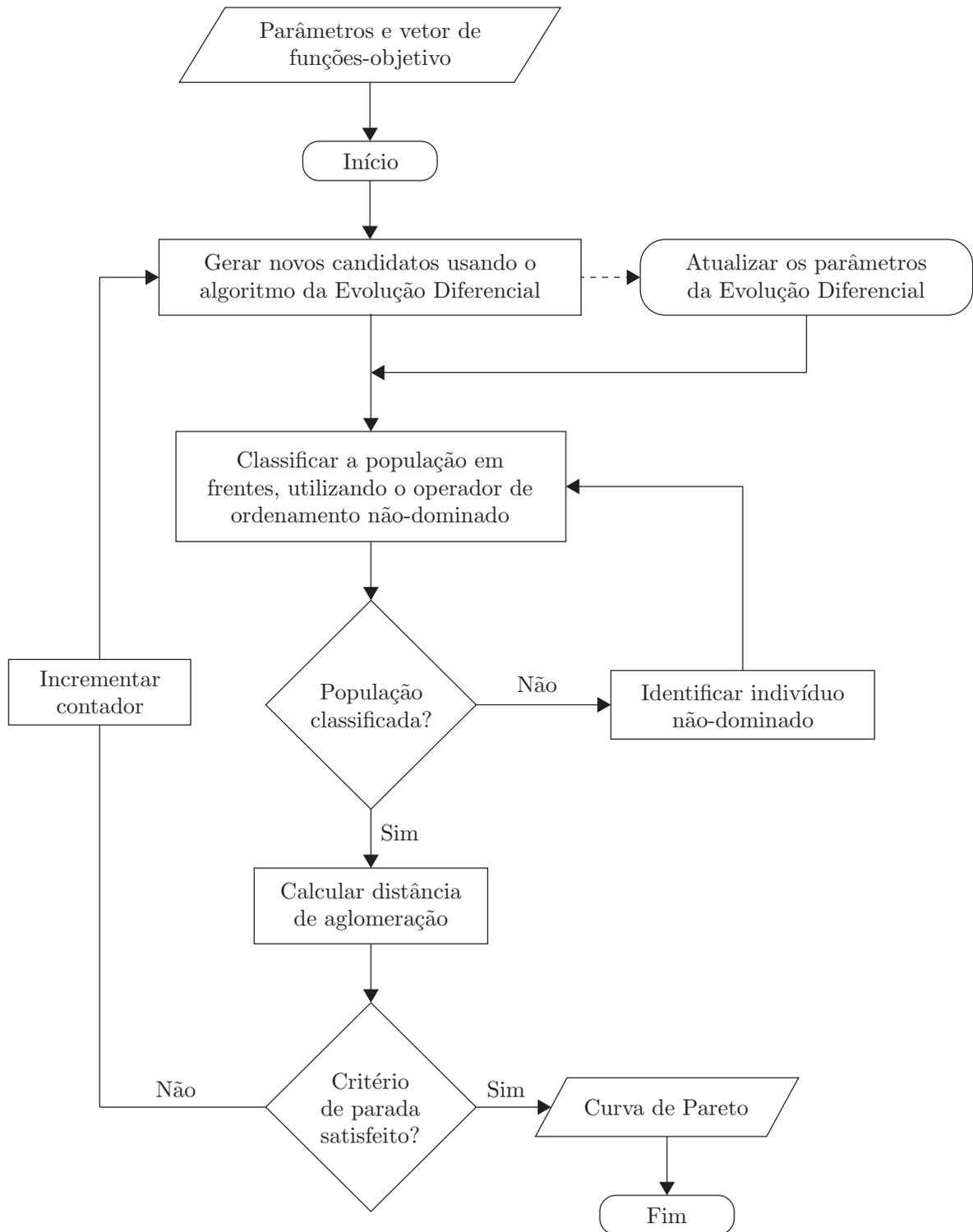
2.4.2.2 Auto-adaptação dos Parâmetros de Controle Evolutivos

Como dito anteriormente, o algoritmo SA-MODE tem como base o processo evolutivo idealizado na metaheurística MODE, mas conta com o atributo de adaptação dos parâmetros oriundos da DE. Especificamente, a cada geração da técnica, são atualizados os parâmetros de tamanho da população (N), fator de amplificação (F) e probabilidade de cruzamento (CR). A atualização desses valores ocorre de maneira dinâmica, através de operadores capazes de avaliar a variação da população e quantificar o processo de convergência do algoritmo. Essa característica tende a propiciar aceleração de convergência (e evitar convergência prematura), fazendo com que o método alterne entre um comportamento mais exploratório ou com tendência de refinamento de soluções, de acordo com a evolução dos indivíduos.

No contexto dos algoritmos de otimização evolutivos, pode ser benéfico ampliar o tamanho da população nas primeiras gerações, quando a diversidade de indivíduos é elevada, e reduzi-la quando a população se torna homogênea. No método SA-MODE, o tamanho da população é atualizado de acordo com a técnica proposta por Cavalini Jr et al. (2016), inicialmente idealizada para o contexto da otimização mono-objetivo. Nessa estratégia, a atualização ocorre com base no parâmetro de taxa de convergência, calculado por $\lambda = f^m / f^w$, onde o numerador e denominador dessa razão significam a média e o pior valor da função-objetivo, respectivamente. Através da avaliação deste parâmetro, é possível mensurar a homogeneidade da população durante o processo evolutivo, uma vez que, conforme $\lambda \rightarrow 1$, presume-se que a homogeneidade da população esteja aumentando.

A fim de expandir o conceito do parâmetro de homogeneidade para os problemas

Figura 13 - Fluxograma do processo iterativo simplificado do SA-MODE.



Fonte: Adaptado de Lobato e Steffen Jr (2017).

multi-objetivo, considere um problema de minimizar dois objetivos simultaneamente. Neste caso, o valor do parâmetro é obtido pelo cálculo aproximado da área sob a curva discretizada, composta pelos valores da população no espaço objetivo em uma dada geração. No

início do processo evolutivo, a tendência é que o valor da área seja elevado. Com o decorrer das iterações, as soluções ótimas calculadas, e avaliadas por cada um dos objetivos, devem se posicionar sobre a curva de Pareto. Levando em conta a propriedade de que as soluções devem ser dominantes ao fim de g gerações, a área abaixo da curva é a menor possível desde aquela correspondente à população inicial. Portanto, o comportamento esperado é que, no decorrer das gerações, a área calculada aproxime-se de um limiar, à medida que o processo iterativo se aproxima do fim, com decaimento proporcional à distância dos objetivos em relação à curva de Pareto.

Desta forma, a estimativa do fator de homogeneidade que considera a área abaixo da curva descrita pelas soluções segue a relação

$$\Lambda^{(g)} = \frac{A^{(g)}}{\frac{1}{M} \sum_{j=g-(M-1)}^g |A^{(j)}|},$$

onde g representa a geração atual, A simboliza o valor aproximado da área sob a curva correspondente a g , e M é o valor médio das áreas computadas nas últimas gerações. Por meio desta estratégia, torna-se possível estabelecer o número de indivíduos que irão compor a população, a cada geração. O cálculo dessa grandeza é através de

$$N^{(g)} = \text{int} \left(\Lambda^{(g)} N_{\min} + (1 - \Lambda^{(g)}) N_{\max} \right). \quad (9)$$

Os termos N_{\min} e N_{\max} representam o mínimo e máximo valores que o tamanho da população pode assumir, respectivamente. Por sua vez, o operador int aproxima o valor do seu argumento para o número inteiro mais próximo.

No caso da atualização do fator de amplificação e da taxa de cruzamento, o procedimento adotado também foi estudado por Cavalini Jr et al. (2016) no contexto mono-objetivo e adaptado para este método. De acordo com Zaharie (2003), a qual foi a idealizadora da estratégia, como os operadores de mutação e cruzamento atuam independentemente em todos as componentes de um vetor que representa um indivíduo da população, é suficiente restringir a análise a uma população composta por escalares, isto é, $\mathbf{x} = (x_1, \dots, x_N)^T$. Sob esta circunstância, a diversidade da população pode ser computada por meio de

$$\text{Var}(\mathbf{x}) = \mathbf{x}^T \mathbf{x} - \left(\sum_{i=1}^N \frac{x_i}{N} \right)^2.$$

Após a aplicação dos operadores de mutação e cruzamento em \mathbf{x} , o valor esperado da variância da população é dado por

$$E(\text{Var}(\mathbf{x})) = \left(1 + 2F^2CR - \frac{2CR}{N} + \frac{CR^2}{N} \right) \text{Var}(\mathbf{x}). \quad (10)$$

Esse resultado é apresentado em forma de teorema por Zaharie (2003) e fundamentado nas devidas demonstrações. No mesmo trabalho, a autora conclui que a convergência prematura da DE pode ser prevenida se os operadores de mutação e cruzamento provocarem aumento na variância populacional. A informação acerca da tendência da variância populacional pode ser obtida através de

$$\varrho = \frac{\text{Var}(\mathbf{x}^{(g+1)})}{\text{Var}(\mathbf{x}^{(g)})}, \quad (11)$$

onde g denota a geração da população correspondente.

Portanto, se a razão da variância das populações de duas gerações subsequentes resultar em $\varrho > 1$, as situações de convergência prematura podem ser evitadas. Por outro lado, se $\varrho < 1$, a convergência do método tende a ser acelerada. Através de experimentos numéricos, foi constatado pela autora que o processo de seleção, usualmente, diminui a variância da população. Considerando esse fato, a escolha de parâmetros que garantam que a Eq. (11) assumir valores positivos é uma condição suficiente para prevenir a redução acentuada da diversidade da população e, conseqüentemente, a convergência prematura do método.

Considere o cálculo do ponto crítico

$$1 + 2F^2CR - \frac{2CR}{N} + \frac{CR^2}{N} = 0. \quad (12)$$

Como dito anteriormente, o lado direito da Eq. (10) resulta da aplicação dos operadores genéticos sobre a população \mathbf{x} . Através da inserção de uma constante de proporcionalidade, denotada por c , a fim de definir uma relação entre ambos os lados desta igualdade, é possível estabelecer que

$$c \equiv \varrho \frac{\text{Var}(\mathbf{x}^{(g+1)})}{\text{Var}(\mathbf{x}^{(g)})} \quad (13)$$

Em vista disso, se a Eq. (12) for resolvida em relação a F com o lado direito igual à constante definida pela Eq. (13), pode-se obter um valor que propicie a convergência de maneira não prematura do método, por meio da imposição de $\varrho > 1$. Para que isso ocorra,

$$F \equiv \begin{cases} \sqrt{\frac{1}{N}} \sqrt{\frac{Y}{2CR}}, & \text{se } Y \geq 0 \\ F_{\min}, & \text{se } Y < 0 \end{cases}, \quad (14)$$

onde $Y = N(c - 1) + CR(2 - CR)$ e F_{\min} representa o valor mínimo que o fator de amplificação pode assumir.

Uma condição suficiente para o aumento da variância da população por meio da mutação é através da escolha de $F > \sqrt{1/N}$. Portanto, é razoável tomar o limite inferior

$F_{\min} = \sqrt{1/N}$. Por sua vez, o limite superior do fator de amplificação pode ser escolhido conforme sugestão proposta por Storn e Price (1997), onde $F_{\max} = 2$. No que diz respeito ao parâmetro da probabilidade de cruzamento, o mesmo procedimento deve ser adotado, mas resolvendo-se a Eq. (12) em relação a CR , com o lado direito igual a c . Assim, obtém-se a solução

$$CR \equiv \begin{cases} 1 - F^2N + \sqrt{(F^2N - 1)^2 - N(1 - c)}, & \text{se } c \geq 0 \\ CR_{\min}, & \text{se } c < 0 \end{cases}, \quad (15)$$

com $CR_{\min} \in (0,01, 1)$.

Resgatando o princípio desta construção, o procedimento está condicionado aos problemas de otimização mono-objetivo. A fim de que esta ideia seja convertida e aplicada no contexto multi-objetivo, deve-se admitir uma condição simples: no desenvolvimento apresentado, a população \mathbf{x} deve ser entendida como o vetor de objetivos, calculados a cada geração. Essa hipótese é razoável, devido à necessidade de se garantir a diversidade de soluções no espaço de objetivos. Portanto, a cada geração o número de indivíduos da população deve ser atualizado respeitando a relação definida pela Eq. (9) e, conseqüentemente, o fator de amplificação da etapa de mutação e o parâmetro de probabilidade de cruzamento devem ser escolhidos de acordo com as Eqs. (14) e (15), respectivamente.

No fluxograma do método SA-MODE, ilustrado na Fig. 13, o critério de parada do algoritmo foi definido de maneira genérica. De fato, existem algumas possibilidades quanto à adoção de critérios de parada para algoritmos evolucionários no contexto multi-objetivo. Dentre os mais utilizados, está o número máximo de gerações, que pode ser aplicado à técnica sem qualquer restrição. Além disso, a linha tracejada e o quadro arredondado do processo de atualização dinâmica dos parâmetros da DE representam um processo opcional. Isso significa que, caso se opte por não efetuar esta sub-rotina, o algoritmo reduz-se à metodologia MODE, apresentada anteriormente.

3 OTIMIZAÇÃO BASEADA EM ROBUSTEZ

3.1 Método de Taguchi

De acordo com extensa pesquisa realizada por Park et al. (2006), o pioneiro no estudo da robustez de projetos foi Genechi Taguchi, introduzindo os conceitos relacionados ao assunto através do estabelecimento do conhecido método de Taguchi (TAGUCHI, 1986). Engenheiro e estatístico, Taguchi teve notória atuação no desenvolvimento de metodologias para aplicação de conceitos estatísticos no aprimoramento de processos de produção e da qualidade de produtos industrializados. O método de Taguchi contribuiu muito para a melhoria da qualidade de vários projetos em empresas renomadas, como AT&T, Ford e Xerox, e o seu legado permanece até os dias atuais, servindo como base para diversas pesquisas e aplicações.

No método de Taguchi, considera-se as variações de desempenho devido a fatores externos de ruído sobre as variáveis de controle do projeto. As variações não podem ser controladas, mas podem ser conhecidas pelo projetista e descritas por funções de densidade de probabilidade. No contexto do método de Taguchi, o controle da qualidade está relacionado às perdas decorrentes de variação excessiva no desempenho funcional. A consistência é alcançada quando o desempenho está próximo da meta com menor variação (ROY, 1990). Portanto, existem dois objetivos na execução de um projeto robusto: minimizar a variabilidade produzida pelos fatores de ruído e tornar o valor médio das variáveis de controle próximo do valor alvo preestabelecido (PARK et al., 2006).

Para tornar os produtos e processos insensíveis à influência de ruídos, Taguchi incorpora uma maneira formal de incluir tais fatores no projeto, através da abordagem de planejamento de experimento (YATES, 1964; BOX, 1980). As funções-objetivo devem levar em conta dois tipos de parâmetros de entrada, a saber, os parâmetros de controle (\mathbf{x}) e fatores de ruído (ξ), que representam a fonte de variação no sistema, no processo de avaliação do nível de variação das variáveis de controle. Nesta abordagem, os parâmetros de controle (denominado *array* interno) são sistematicamente modificados. Para cada um dos valores de \mathbf{x} , o fator de ruído também varia (*array* externo), gerando assim um *array* cruzado através da combinação dos valores de ambos (BEYER; SENDHOFF, 2007; JURECKA, 2007).

Utilizando este procedimento de amostragem, Taguchi propôs uma relação de sinal-ruído, que permite mensurar o desvio do objetivo avaliado em relação às metas estipuladas para a variação de desempenho do sistema. A avaliação sistemática, através da abordagem de planejamento de experimento, permite a determinação de um índice, de tal forma que quanto maior for o seu valor, menor será a variação do produto em torno do valor definido como meta (ROY, 1990). Seja κ o número de observações e \hat{y} o objetivo definido como meta.

O desvio médio, com valor fixado para um dado parâmetro de controle, é calculado através de

$$D(\mathbf{x}, \boldsymbol{\xi}) = \frac{1}{\kappa} \sum_{i=1}^{\kappa} (y(\mathbf{x}, \xi_i) - \hat{y})^2 .$$

Pela variação de \mathbf{x} , obtém-se um conjunto de dados que descrevem o desvio quadrático médio para cada valor do *array* interno. Então, a relação sinal-ruído é calculada com

$$R(\mathbf{x}, \boldsymbol{\xi}) = -10 \log_{10} (D(\mathbf{x}, \boldsymbol{\xi})) , \quad (16)$$

que deve ser maximizada em relação a \mathbf{x} e é utilizada na escala logarítmica por conveniência da linearidade e influência da possível utilização de elevada quantidade de dados. Mais detalhes podem ser encontrados em Taguchi (1987).

Dentre diversas aplicações existentes na literatura, pode-se destacar o trabalho de Park et al. (1994), que adotaram a metodologia proposta por Taguchi em um projeto estrutural, durante a etapa de pós-processamento de otimização do projeto. Mais tarde, Hwang et al. (2001) também apresentaram um estudo da otimização robusta de um espelho retrovisor de automóvel para redução de vibração. Por sua vez, Ku et al. (1998) propuseram uma forma modificada de relação sinal-ruído, de tal forma a estender o procedimento para problemas de otimização com restrições e na presença de funções de objetivo único e múltiplo. De forma abrangente, Taguchi et al. (1999) descreveram diversos estudos de caso de sucesso, nos quais as aplicações estão relacionadas a problemas relatados por grandes empresas de diferentes ramos da indústria.

Do ponto de vista da eficiência de otimização, a abordagem de Taguchi apresenta custo computacional elevado, que depende da quantidade de elementos que constituem o *array* cruzado, composto pela combinação dos parâmetros de controle e fatores de ruído. Se considerarmos um *array* n -dimensional, serão necessários, pelo menos, dois pontos por dimensão, o que significa que serão realizadas 2^n avaliações. Contudo, se forem consideradas todas as possíveis combinações entre \mathbf{x} e $\boldsymbol{\xi}$, a relação sinal-ruído da Eq. (16) necessitará de $\kappa 2^n$ avaliações para ser estimada (BEYER; SENDHOFF, 2007; TROSSET, 1996). As vantagens e desvantagens da metodologia proposta por Genechi Taguchi foram sintetizadas por Lochner (1991).

3.2 Método de Superfície de Resposta para Análise de Robustez

Em muitos casos, o modelo matemático do sistema também não permite o desacoplamento dos fatores de controle que afetam a média e a variância na estimação da relação sinal-ruído, fazendo com que as variáveis de controle e parâmetros de ruído sejam difícil-

mente identificados. Compreender e modelar o sistema envolve entender quais variáveis impactam a média do sistema e quais impactam a variância do sistema (algumas variáveis afetam ambos). Nas pesquisas de Box et al. (1988) e Nair e Pregibon (1988), são discutidas metodologias que permitem determinar os fatores que podem afetar as medidas estatísticas de um sistema. Em vista dos modelos matemáticos cada vez mais complexos, naturalmente são propostos métodos mais eficazes, capazes de estimar resultados ótimos robustos, com custo computacional mais baixo. Neste contexto, o conceito da metodologia de superfície de resposta envolve a experimentação sequencial, a fim de construir um modelo apropriado, que permita compreender o sistema de engenharia a ser modelado (BOX; WILSON, 1951 apud ROBINSON et al., 2003).

Baseado nesta concepção, Vining e Myers (1990) propuseram uma abordagem de superfície de resposta a modelos robustos, ampliando os conceitos de resposta dupla estabelecidos por Myers e Carter (1973). O problema de resposta dupla envolve a determinação do conjunto de variáveis de controle \mathbf{x}^* que otimiza uma resposta primária \hat{y}_p , a qual está sujeita a $\hat{y}_s = \theta$, onde θ é algum valor aceitável para a resposta secundária. Para tal, suponha o modelo quadrático que dá origem a uma função de resposta primária, da forma

$$\hat{y}_p(\mathbf{x}) = b_0 + \mathbf{x}^T \mathbf{b} + \mathbf{x}^T \mathbf{B} \mathbf{x} ,$$

onde \mathbf{x} é um vetor de variáveis independentes e \hat{y}_p é a resposta primária estimada, calculada através de dados obtidos experimentalmente, desconsiderando-se as prováveis restrições do problema. Por sua vez, os elementos de \mathbf{b} e \mathbf{B} representam os estimadores de mínimos quadrados.

A resposta secundária (\hat{y}_s) pode ser obtida através do mesmo experimento, com o uso de regressão linear múltipla e levando-se em conta as restrições do problema, como por exemplo os limites de custo do projeto. A função de resposta secundária é definida por

$$\hat{y}_s(\mathbf{x}) = c_0 + \mathbf{x}^T \mathbf{c} + \mathbf{x}^T \mathbf{C} \mathbf{x} ,$$

de tal forma que os seus parâmetros são análogos ao modelo de resposta primária. Conhecendo os modelos ajustados de resposta, Myers e Carter (1973) propuseram a utilização de multiplicadores de Lagrange para encontrar os valores de \mathbf{x} que otimizam \hat{y}_p , sujeito a $\hat{y}_s = \theta$.

Portanto, tomando λ_θ como o multiplicador de Lagrange associado às restrições de \hat{y}_p e considerando uma restrição adicional, definida como $\mathbf{x}^T \mathbf{x} \leq \rho^2$, tal que ρ representa o raio da hipersfera que limita as variáveis de controle, o problema de otimização é definido como

$$L(\mathbf{x}, \theta, \rho) = \hat{y}_p - \lambda_\theta (\hat{y}_s - \theta) - \lambda_\rho (\mathbf{x}^T \mathbf{x} - \rho) ,$$

com λ_ρ denotando o multiplicado de Lagrange associado à restrição adicional. A obtenção das soluções ótimas para \mathbf{x} decorrem de $\partial L/\partial \mathbf{x} = 0$, para uma combinação de valores arbitrários de λ_θ e λ_ρ , de tal maneira que \mathbf{x} estará definido em um lugar geométrico determinado pelo raio ρ .

Diversas formulações, baseadas nos conceitos de superfície de resposta a modelos robustos, foram sugeridas na sequência, a fim de aperfeiçoar as ideias apresentadas. O uso de programação não-linear para obtenção de ótimos robustos foi sugerido por Castillo e Montgomery (1993), propondo a utilização de um algoritmo de gradiente reduzido generalizado para resolver problemas que continham restrições de desigualdade. Mais tarde, os conceitos foram expandidos por Myers e Montgomery (1995).

No mesmo contexto, o método Simplex (NELDER; MEAD, 1965) foi empregado por Copeland e Nelson (1996), a fim de realizar a otimização direta dos objetivos do problema. Em outra ocasião, Lin e Tu (1995) indicaram que a abordagem de Vining e Myers (1990) nem sempre garante soluções ótimas globais devido às limitações da otimização no caso de problemas que contêm restrições de igualdade, e sugeriram o uso do critério de erro quadrático médio para otimização simultânea da média e variância dos modelos de superfície de resposta. Posteriormente, Kim e Lin (1998) também propuseram uma nova formulação de programação matemática para o problema de dupla resposta, baseada na metodologia de otimização *fuzzy*.

Na tentativa de tornar a metodologia mais generalista, Castillo et al. (1997) forneceram uma implementação de um algoritmo para otimização global do problema de resposta dupla. Este método de otimização é o mais completo, de acordo com Ankenman e Dean (2003), pois garante a obtenção do ótimo global (apesar de alguns “casos degenerados”, descritos pelos autores). Contudo, os demais métodos podem ter desempenhos satisfatórios, na maioria dos problemas, além da possibilidade de serem mais simples quanto a implementação computacional.

A literatura apresenta inestimável quantidade de aplicações distintas, que utilizam a abordagem de superfície de resposta. Nota-se uma extensa diversidade, haja vista a aplicabilidade da metodologia, dentre as quais pode-se citar: sistema de irrigação movido a energia solar (CHEN et al., 1996), otimização robusta de torque no projeto de motor (GAO et al., 2002), motor supercondutor de alta temperatura (KIM et al., 2002), gestão de inventário (DELLINO et al., 2010), otimização de um processo de destilação extrativa (BIGLARI et al., 2010), remoção fotocatalítica de poluentes ambientais utilizando nanocatalisadores (KHATAEE et al., 2011) e tratamento de águas residuais de reciclagem de papel (BIRJANDI et al., 2013).

Uma breve pesquisa é capaz de revelar a abrangência do assunto e a profusão de aplicações dos métodos e suas variantes, em problemas dos mais variados setores. Em vista da grande quantidade de publicações, Hill e Hunter (1966) apresentaram uma extensa pesquisa bibliográfica acerca das aplicações da metodologia de superfície de resposta, especi-

almente aquelas relativas à engenharia química, a qual concentrava a grande maioria do emprego da técnica até aquele momento. Com a crescente evolução das técnicas e propostas de métodos mais eficazes, além da contínua aplicação dos conceitos da metodologia de superfície de resposta em áreas mais diversificadas, Myers et al. (1989) apresentaram uma síntese dos desenvolvimentos mais significativos do período em que a sua pesquisa abrange. Mais recentemente, o assunto foi novamente revisado através dos trabalhos publicados por Robinson et al. (2003) e Myers et al. (2004).

Apesar da grande variedade de aplicações e diversidade de técnicas que foram propostas para aperfeiçoar os conceitos propostos por Myers e Carter (1973), a estratégia apresenta limitações evidentes. No caso de ajustes a modelos quadráticos, a determinação dos parâmetros requer n^2 avaliações da função, um valor elevado quando consideradas simulações complexas, o que pode impactar excessivamente o custo computacional, mesmo que sejam utilizadas técnicas sofisticadas de amostragem, como hipercubo latino e planejamento fatorial fracionário. Ademais, as incertezas inerentes aos dados utilizados nos ajustes dos parâmetros do modelo estão relacionadas a um problema particular de robustez (BEYER; SENDHOFF, 2007).

3.3 Otimização Robusta e Algoritmos Genéticos

Em muitos problemas práticos, as funções-objetivo são perturbadas por ruídos aleatórios, e algoritmos genéticos (GA) (GOLDBERG, 1989a) têm sido amplamente propostos como uma ferramenta de otimização eficaz para lidar com problemas deste tipo. Algumas das abordagens mais recentes acerca de otimização robusta surgiram paralelamente ao avanço das técnicas relacionadas aos algoritmos genéticos. Como se sabe, o conceito da otimização evolucionária é inspirado na teoria da seleção natural de Charles Darwin. Dessa forma, os algoritmos projetados de acordo com esse paradigma devem obedecer a propriedades similares. Dentre os conceitos admitidos pela teoria, espera-se que indivíduos com maior capacidade de suportar alterações no meio em que estão dispostos, sejam capazes de evoluir e gerar descendentes, em detrimento àqueles mais sensíveis a tais variações. Conceitualmente, essa noção é empregada nos algoritmos genéticos e foi estendida para a avaliação de problemas sob efeito de ruído.

Sob esta perspectiva, pode-se afirmar que as características fenotípicas de um indivíduo, determinadas pelo código genético nos cromossomos que, por sua vez, estão sujeitos a mutações causadas por fatores externos, podem ser determinantes na sua sobrevivência e capacidade de reprodução de novos indivíduos. Portanto, aqueles menos sensíveis a tais perturbações, isto é, que possuem características fenotípicas favoráveis ao desenvolvimento, são ditos robustos e tendem a gerar descendentes igualmente propícios a evolução (TSUTSUI et al., 1996; TSUTSUI; GHOSH, 1997; BRANKE, 1998; BRANKE, 2000).

No âmbito da computação, tais aspectos estão associados às soluções de um problema e sua relação com a função-objetivo. Espera-se que não apenas a solução do problema seja de boa qualidade, mas também os indivíduos na sua vizinhança. Matematicamente, esse fato deve ser interpretado de tal forma que, soluções capazes de sofrer perturbações arbitrárias sem variação excessiva do valor da função-objetivo tendem a ser soluções robustas. A princípio, para a avaliação de problemas de otimização com a presença de ruídos, considerava-se um ruído aditivo com distribuição normal (HAMMEL; BÄCK, 1994), tal que a função-objetivo $f : M \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e

$$F(\mathbf{x}) = f(\mathbf{x}) + \delta ,$$

onde $\delta \in \mathbb{R}$ é uma variável aleatória, com $\delta \sim \mathcal{N}(0, \sigma_\delta^2)$ e M é o espaço de busca.

Grande parte das pesquisas relacionadas a esta abordagem foram motivadas pelos estudos realizados por Holland (1973) e Jong (1975). A capacidade de resolução de problemas de otimização com a presença de ruído foi explorada por Fitzpatrick e Grefenstette (1988), que analisaram o esforço despendido na avaliação de cada estrutura, em relação ao número total de estruturas, durante uma determinada iteração do algoritmo genético, constatando a possibilidade de resultados mais eficientes a partir de avaliações menos precisas. Posteriormente, Aizawa e Wah (1994) desenvolveram novos métodos para ajustar parâmetros de configuração de algoritmos genéticos que operam em ambientes com ruído, e Bäck e Hammel (1994) demonstraram que perturbações aos valores da função-objetivo não influenciam o desempenho das estratégias evolucionárias, desde que o nível de ruído seja pequeno em comparação com o valor da função.

Dentre as pesquisas que consideram as incertezas do sistema como perturbações aplicadas diretamente ao valor da função-objetivo, também pode-se destacar o trabalho de Hammel e Bäck (1994), que discutiram a influência de alguns parâmetros no processo de convergência das estratégias evolucionárias e alternativas para melhoria das propriedades de convergência em problemas desse tipo. Mais tarde, pôde-se perceber que a avaliação do efeito do ruído sendo adicionado à saída do sistema não se tratava da maneira mais adequada de representar as modificações das quais o código genético estaria sujeito, e que determinavam o fenótipo de um dado indivíduo, uma vez que a perturbação deveria ocorrer durante a transformação do genótipo. Portanto, Tsutsui et al. (1996) propuseram um esquema de busca por soluções robustas, de forma que o ruído fosse adicionado às variáveis de decisão, que representam o fenótipo na metáfora biológica. Para emular esse tipo de característica genética, seja $\mathbf{x} \in \mathbb{R}^n$ o vetor das variáveis de decisão e $\boldsymbol{\delta} \in \mathbb{R}^n$ o vetor de ruídos, as soluções robustas podem ser avaliadas como

$$F(\mathbf{x}, \boldsymbol{\delta}) = f(\mathbf{x} + \boldsymbol{\delta}) . \tag{17}$$

Nota-se que a adição de ruído, neste caso, pode ter semelhança com o operador ge-

nético de mutação dos algoritmos evolucionários. Contudo, a proposta não é exatamente a mesma, uma vez que a adição de ruído não afeta os indivíduos da população, fazendo com que as perturbações da Eq. (17) sejam efetivas apenas na etapa de seleção (TSUTSUI et al., 1996; TSUTSUI; GHOSH, 1997). Através desta abordagem, também se torna possível quantificar o ruído relacionado a cada uma das variáveis de decisão, tornando o modelo matemático mais flexível e capaz de descrever o problema físico com maior exatidão.

Na ocasião, Tsutsui et al. (1996) também apresentaram um novo método computacional, que estende a aplicação dos algoritmos genéticos a domínios que exigem detecção de soluções robustas. Os autores utilizaram a abordagem da Eq. (17) para demonstrar a sensibilidade das soluções através de testes empíricos. Mais tarde, Tsutsui e Ghosh (1997) consolidaram a pesquisa prévia, demonstrando a efetividade da técnica na obtenção de múltiplas soluções robustas. A solução de problemas multimodais também foi objeto de estudo de Branke (1998), que sugeriu modificações ao GA e comparou os resultados com o algoritmo canônico. Em vista desses resultados, posteriormente Branke (2000) examinou uma série de estratégias adicionais para tornar a busca por soluções robustas mais eficiente e para compreender a influência de diferentes parâmetros nos algoritmos evolucionários.

Pouco havia sido explorado, até então, em termos de soluções robustas relacionadas a problemas multi-objetivo. Em uma das primeiras pesquisas que considerou essa possibilidade, as técnicas existentes para a exploração do espaço de busca com base no critério de dominância de Pareto foram estendidas para o caso em que um ou mais objetivos de uma função estão sujeitos a incertezas (TEICH, 2001 apud DEB; GUPTA, 2005). Seguindo essa tendência, o procedimento probabilístico de classificação de soluções mais promissoras de algoritmos evolucionários foi reformulado, a fim de considerar incertezas e problemas que contém ruídos (HUGHES, 2001 apud DEB; GUPTA, 2005). Com esses avanços, o conflito de escolhas entre desempenho e robustez foi identificado na forma da frente de Pareto e foram sugeridos dois métodos para estimativa de robustez no contexto multi-objetivo (JIN; SENDHOFF, 2003 apud DEB; GUPTA, 2005).

Apesar de tudo, nenhum desses estudos considerou a sensibilidade das variáveis envolvidas no projeto, devido a incertezas em um cenário multi-objetivo, como apontam Deb e Gupta (2005). Motivados por esse indício, os mesmos autores ampliaram as abordagens apresentadas por Tsutsui e Ghosh (1997) e Branke (1998), visando o cálculo de soluções robustas no contexto multi-objetivo e, em uma pesquisa seguinte, também realizada por Deb e Gupta (2006), os conceitos definidos anteriormente passaram a englobar os problemas de otimização com restrições e foram aplicados na solução de um problema de soldagem de viga.

3.3.1 Técnicas de Análise de Robustez com Abordagem Genética

Nesta seção, o objetivo é dispor detalhes pertinentes à aplicação das metodologias de avaliação de robustez em problemas de otimização multi-objetivo, com foco na abordagem apresentada na Eq. (17). Vale ressaltar que existe uma gama de possibilidades no que diz respeito a metodologias de avaliação de robustez, de tal forma que as incertezas na computação evolucionária também abrangem diversas categorias. Na pesquisa realizada por Jin e Branke (2005), onde as incertezas foram categorizadas em quatro grandes grupos, os autores apresentam uma visão geral das aplicações relacionadas a robustez, até aquele momento, sob uma perspectiva unificada. Nesta tese, a intenção é detalhar apenas as metodologias relacionadas às aplicações apresentadas aqui.

3.3.1.1 Aspectos Essenciais do Teorema Fundamental dos Algoritmos Genéticos

Antes de apresentar os métodos para a análise de robustez, é preciso abordar alguns detalhes acerca dos algoritmos genéticos e introduzir o Teorema Fundamental dos Algoritmos Genéticos, que serve como base para as metodologias que serão mostradas a seguir. A teoria foi formulada por Holland (1975) e difundida por Goldberg (1989a), e tem como objetivo formalizar o funcionamento dos algoritmos genéticos de maneira genérica. De forma resumida, a teoria indica o seguinte: inicialmente, considere um esquema, estruturado como uma sequência de zeros, uns e/ou asteriscos, sendo este último chamado de meta-símbolo, uma vez que pode assumir qualquer um dos dois demais valores possíveis. A sequência $10*0*11$ representa um conjunto com quatro cromossomos distintos (também denominados como indivíduos da população), correspondentes à troca de asteriscos por zeros ou uns.

Também pode ser interessante conhecer o número de esquemas distintos que um determinado cromossomo pode representar, uma vez que podem ser recombinados. Uma vez considerado um alfabeto binário, cada cromossomo dado por uma sequência com k entradas ou genes pode representar 2^k esquemas distintos e, portanto, uma população contendo n membros corresponde a $n2^k$ esquemas diferentes. Se tomarmos um cromossomo específico do exemplo anterior, como 1010011 , será possível produzir $2^7 = 128$ combinações distintas de cromossomos. Obviamente, em uma população de n indivíduos, pode haver cromossomos semelhantes, que não necessitam ser processados tantas quanto forem as vezes da sua ocorrência.

O GA utiliza três operadores genéticos—mutação, *crossover* e seleção—na população, a fim de obter soluções ótimas. Para examinar o efeito destes três operadores, é necessário quantificar a noção do fato de que alguns esquemas podem ser mais específicos do que outros. Para esta tarefa, dois conceitos são essenciais: a ordem dos esquemas e o compri-

mento definidor. O primeiro, denotado por $o(\mathbf{H})$, exprime a quantidade de uns e zeros no esquema \mathbf{H} , isto é, o total de posições fixas. Já o comprimento definidor, que é representado por $\nu(\mathbf{H})$, corresponde à distância entre as posições do primeiro e último genes específicos do cromossomo, ou seja, a distância entre zeros e uns que ocupam as posições mais ao extremo. Como exemplo, pode-se verificar que $o(011*1**) = 4$ e $\nu(011*1**) = 5 - 1 = 4$.

A considerar estes elementos, a intenção é avaliar o efeito dos operadores genéticos sobre a população no decorrer das gerações. Especificamente, com o objetivo de esclarecer o efeito do operador de seleção nos esquemas, considere o valor médio de adaptação dos cromossomos (valor da função-objetivo) representados por um esquema \mathbf{H} , denotado por $f(\mathbf{H})$ e calculado através da média aritmética da adaptação dos cromossomos que o constituem. Além disso, seja $\bar{f}(\mathbf{A}(t))$ o valor da adaptação média de toda a população \mathbf{A} no instante t . Portanto, espera-se haver $m(\mathbf{H}, t + 1)$ representantes do esquema \mathbf{H} na população $\mathbf{A}(t + 1)$, cujo valor é estimado pelo número inteiro mais próximo resultante de

$$m(\mathbf{H}, t + 1) = m(\mathbf{H}, t) \frac{f(\mathbf{H})}{\bar{f}(\mathbf{A}(t))}, \quad (18)$$

tal que $m(\mathbf{H}, t)$ define o número de esquemas na geração atual. Em outras palavras, um esquema arbitrário desenvolve-se proporcionalmente em relação à taxa que associa a adaptação média dos seus cromossomos com a adaptação da população como um todo.

Por sua vez, no caso do *crossover*, é essencial que seja considerada a destruição de determinados esquemas, pela ação do operador. Para efetivar esta análise, considere um exemplo que envolva dois cromossomos selecionados para cruzamento, formados pelas sequências 101011 e 010101. Pressupondo que o ponto de corte do cruzamento, escolhido aleatoriamente, seja $s = 2$, as novas sequências, resultantes da operação de *crossover*, são 111011 e 000101. É fato que os cromossomos obtidos podem ser oriundos de diversas possibilidades de esquema, dentre as quais, cabe tomar 1110** e 0*0*01 como suposições, respectivamente. Admitindo que, em um dado problema de otimização, o nível de adaptação do esquema 0*0*01 seja superior ao de 1110**, mesmo assim pode haver a ocorrência de destruição do esquema mais apropriado para evolução.

Este fato ocorre em virtude da influência do comprimento definidor de cada esquema, que vale $\nu(1*0*01) = 5$ no primeiro caso, ao passo de que $\nu(1110**) = 3$. Através deste exemplo, verifica-se que apenas cortes em posições extremamente propícias evitariam a destruição do candidato mais favorável a evolução, como por exemplo um corte em $s = 1$, dependendo do problema avaliado. De forma mais geral, é possível estimar um limite inferior da probabilidade de sobrevivência após o cruzamento. Para tanto, considere p_c como a probabilidade de ocorrência de cruzamento. A probabilidade mínima de sobrevivência é

calculada através da expressão

$$p_{sc} \geq (1 - p_c) \frac{\nu(\mathbf{H})}{k - 1}, \quad (19)$$

onde k é o número de genes.

Por fim, a ação do operador de mutação é avaliada partindo da premissa de que um esquema \mathbf{H} sobrevive e avança para a próxima geração, uma vez que nenhum de seus genes é afetado pelo operador, que realiza uma alteração única em algum dos genes do esquema (através da alternância do respectivo valor absoluto), com probabilidade p_m . Uma vez que a mutação é independente para cada uma das posições fixas do esquema, pode-se considerar que a sobrevivência ocorre quando $p_{sm} = (1 - p_m)^{o(\mathbf{H})}$. Para valores reduzidos da probabilidade de ocorrência da mutação ($p_m \ll 1$), a chance de sobrevivência ao operador pode ser aproximada pela expressão

$$p_{sm} \approx 1 - o(\mathbf{H}) p_m. \quad (20)$$

Esta aproximação é baseada no fato de que a mutação causa um salto (aleatório) na localização das soluções geradas e, portanto, valores elevados podem reduzir a capacidade de exploração do GA, tornando praticamente aleatória a busca pelo ótimo (GOLDBERG, 1989b).

De maneira geral, a evolução de um dado esquema no decorrer das gerações de um GA deve satisfazer três princípios fundamentais: o esquema deve ter adaptação média superior à adaptação média da população, com comprimento definidor curto e ordem dos esquemas reduzida. Em decorrência destes fatos, foi enunciado o Teorema Fundamental dos Algoritmos Genéticos, que é capaz de estimar o número de representantes de um dado esquema que estarão presentes na população em um instante $t + 1$, relacionando os efeitos dos operadores de mutação, *crossover* e seleção. Combinando as Eqs. (18)–(20), o número de representantes do esquema \mathbf{H} na geração $t + 1$ é estimado como

$$m(\mathbf{H}, t + 1) \geq m(\mathbf{H}, t) \frac{f(\mathbf{H})}{\bar{f}(\mathbf{A}(t))} \left(1 - p_c \frac{\nu(\mathbf{H})}{k - 1} - o(\mathbf{H}) p_m \right). \quad (21)$$

Em suma, a concepção do teorema abrange conceitos relativamente simples, mas de grande valia para a teoria dos algoritmos genéticos e diversas técnicas evolucionárias que são derivadas desta formulação. O desenvolvimento mais detalhado, além de exemplos numéricos e aplicações podem ser vistas na obra de Goldberg (1989a). A seguir, são apresentados alguns métodos para análise de robustez, com ideias derivadas, direta e indiretamente, da teoria apresentada nesta seção.

3.3.1.2 Média Efetiva

Como citado anteriormente, Tsutsui e Ghosh (1997) apresentaram uma técnica que ampliou a aplicação do GA para problemas que requerem identificação de soluções robustas, com ruídos sendo adicionados aos valores dos parâmetros fenotípicos enquanto se avalia o valor funcional dos indivíduos. A proposta é derivada do Teorema Fundamental dos Algoritmos Genéticos, apresentado anteriormente e sintetizado pela Eq. (21). Seja \mathbf{X} um vetor de decisão com distribuição normal. Analisando o valor da adaptação média $\bar{f}(\mathbf{A}(t))$ de toda a população, quando o seu tamanho $N \rightarrow \infty$, a *fitness* média dos indivíduos de \mathbf{A} , em um instante t , é dada por

$$\bar{a}(t) = \int_{\mathbf{x}} f(\mathbf{x}) p(\mathbf{x}, t) d\mathbf{x} \approx \sum_{i=1}^N \frac{f(\mathbf{x}^i)}{N}, \quad (22)$$

onde $p(\mathbf{x}, t)$ representa a distribuição de probabilidade de \mathbf{X} na população. De maneira análoga, o valor da adaptação média dos representantes do esquema \mathbf{H} na população $\mathbf{A}(t)$ é calculado por

$$f(\mathbf{H}) = \int_{\mathbf{x}} f(\mathbf{x}) p(\mathbf{x}, \mathbf{H}, t) d\mathbf{x}, \quad (23)$$

tal que $p(\mathbf{x}, \mathbf{H}, t)$ fornece a distribuição de probabilidade de \mathbf{X} no esquema \mathbf{H} de uma população (TSUTSUI; GHOSH, 1997).

No caso da abordagem robusta, considere que o nível de adaptação é avaliado através de $f(\mathbf{x} + \boldsymbol{\delta})$, onde $\boldsymbol{\delta}$ representa o vetor de incertezas associadas a cada valor de \mathbf{X} , também com distribuição normal. Neste caso, o Teorema Fundamental dos Algoritmos Genéticos, representado pela Eq. (21), toma a forma

$$m(\mathbf{H}, t+1) \geq m(\mathbf{H}, t) \frac{f_r(\mathbf{H})}{\bar{f}_r(\mathbf{A}(t))} \left(1 - p_c \frac{\nu(\mathbf{H})}{k-1} - o(\mathbf{H}) p_m \right),$$

onde $f_r(\mathbf{H})$ e $\bar{f}_r(\mathbf{A}(t))$ denotam o valor da adaptação média dos cromossomos de um esquema \mathbf{H} e da população $\mathbf{A}(t)$, no contexto do problema de robustez, respectivamente. Assuma que \mathbf{X} e $\boldsymbol{\delta}$ são mutuamente independentes, com $q(\boldsymbol{\delta})$ sendo a função densidade de probabilidade do ruído $\boldsymbol{\delta}$. Nestas circunstâncias, a adaptação média da população no con-

texto de robustez é dada por

$$\begin{aligned}
 \bar{f}_r(\mathbf{A}(t)) &= \int_{\mathbf{x}} \int_{-\infty}^{\infty} f(\mathbf{x} + \boldsymbol{\delta}) p(\mathbf{x}, t) q(\boldsymbol{\delta}) d\boldsymbol{\delta} d\mathbf{x} \\
 &= \int_{\mathbf{x}} \left[\int_{-\infty}^{\infty} f(\mathbf{x} + \boldsymbol{\delta}) q(\boldsymbol{\delta}) d\boldsymbol{\delta} \right] p(\mathbf{x}, t) d\mathbf{x} \\
 &= \int_{\mathbf{x}} F(\mathbf{x}) p(\mathbf{x}, t) d\mathbf{x} \\
 &\approx \sum_{i=1}^N \frac{f(\mathbf{x}^i + \boldsymbol{\delta}^i)}{N}.
 \end{aligned} \tag{24}$$

Por sua vez, o valor da adaptação média dos cromossomos que representam o esquema \mathbf{H} , ainda no contexto do problema de robustez, é representado por

$$f_r(\mathbf{H}) = \int_{\mathbf{x}} F(\mathbf{x}) p(\mathbf{x}, \mathbf{H}, t) d\mathbf{x}. \tag{25}$$

Portanto, se forem comparadas as expressões que descrevem a adaptação média dos indivíduos da população na versão canônica do GA, definida pela Eq. (22), e a expressão da adaptação média da população no contexto de robustez, representada pela Eq. (24), nota-se que $F(\mathbf{x})$ corresponde a $f(\mathbf{x})$. Também é possível inferir sobre a equivalência dos mesmos termos, se tomarmos a expressão que representa a adaptação média dos representantes de um dado esquema \mathbf{H} , mostrada na Eq. (23), em comparação com aquela que descreve a adaptação média dos representantes de um esquema no problema de robustez, definida pela Eq. (25). A partir desta constatação, Tsutsui e Ghosh (1997) concluíram que, quando a quantidade de indivíduos $N \rightarrow \infty$, o número médio de representantes de cada esquema no problema de robustez depende de $F(\mathbf{x})$ ao invés de $f(\mathbf{x})$.

Portanto, a formulação capaz de avaliar a robustez em um problema de otimização, no qual $F(\mathbf{x})$ é equivalente a $f(\mathbf{x})$ avaliada em $\mathbf{x} + \boldsymbol{\delta}$, recebe o nome de média efetiva. Sendo assim, a definição do problema de otimização robusta, referente à abordagem apresentada na Eq. (17), é dada por

$$F(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{x} + \boldsymbol{\delta}) q(\boldsymbol{\delta}) d\boldsymbol{\delta}.$$

Em razão da simetria inerente à função densidade de probabilidade, a mesma análise realizada anteriormente vale quando $q(\boldsymbol{\delta}) = q(-\boldsymbol{\delta})$ na Eq. (24).

Na tentativa de ampliar o conceito de média efetiva para o contexto dos problemas de otimização multi-objetivo, Deb e Gupta (2005) definiram as soluções multi-objetivo robustas de tipo I, como segue: uma solução \mathbf{x}^* é chamada de *solução multi-objetivo robusta*

de tipo I, se for um conjunto ótimo de Pareto para o problema de otimização multi-objetivo:

$$\begin{aligned} \arg \min \quad & \mathbf{F}^{\text{eff}}(\mathbf{x}) = (f_1^{\text{eff}}(\mathbf{x}), \dots, f_m^{\text{eff}}(\mathbf{x}))^T, \\ \text{Sujeito a} \quad & \mathbf{x} \in \Omega \end{aligned}$$

definido com respeito à vizinhança δ , tal que $f_i^{\text{eff}}(\mathbf{x})$ é dada por

$$f_i^{\text{eff}}(\mathbf{x}) = \frac{1}{|\mathcal{B}_\delta(\mathbf{x})|} \int_{\mathbf{y} \in \mathcal{B}_\delta(\mathbf{x})} f_i(\mathbf{y}) d\mathbf{y}, \quad (26)$$

com $i = 1, \dots, m$. Nesta formulação, $\mathcal{B}_\delta(\mathbf{x})$ denota a vizinhança δ do candidato a solução $\mathbf{x} \in \mathbb{R}^n$ e $|\mathcal{B}_\delta(\mathbf{x})|$ representa o hipervolume da vizinhança.

Analisando a Eq. (26), fica claro que se trata da integração analítica da função-objetivo com o máximo possível de perturbação nas variáveis. Essa equação é aplicável aos problemas com solução conhecida da integral no espaço de busca. Para problemas nos quais o espaço de busca é mais complexo, a Eq. (26) pode ser aproximada pelo valor médio atingido pela função-objetivo, em virtude da complexidade da avaliação da integral, além do custo computacional elevado. Nesse sentido, a aproximação pode ser calculada por

$$f_i^{\text{eff}}(\mathbf{x}) \approx \frac{1}{M} \sum_{j=1}^M f_i(\mathbf{y}^j),$$

com M representando o número finito de pontos distribuídos ao redor de \mathbf{x} . Considerando o vetor de incertezas δ associadas às variáveis de decisão \mathbf{x} , as soluções distribuídas aleatoriamente—ou respeitando algum esquema estruturado—devem estar dispostas de tal forma que $\mathbf{y} \in [\mathbf{x} - \delta\mathbf{x}, \mathbf{x} + \delta\mathbf{x}]$, com $\delta_k \in \mathbb{R}_+$, para $k = 1, \dots, n$.

3.3.1.3 Diferença Normalizada

Fundamentada na estratégia de avaliação de robustez através do cálculo de média efetiva, Deb e Gupta (2005) propuseram outra formulação para o cálculo da robustez, por meio de uma abordagem mais pragmática, que considera a diferença normalizada entre os valores das funções original $F(\mathbf{x})$ e perturbada $\mathcal{F}(\mathbf{x})$, de forma a estabelecer um limiar capaz de quantificar as alterações causadas pelos ruídos no problema determinístico. Desta maneira, uma solução é considerada robusta se tal diferença normalizada satisfaz uma meta de robustez preestabelecida.

Matematicamente, uma solução \mathbf{x}^* é chamada de *solução multi-objetivo robusta de tipo II* se for um conjunto ótimo de soluções de Pareto para o problema de otimização multi-

objetivo:

$$\begin{aligned} \arg \min \quad & \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^T \\ \text{Sujeito a} \quad & \frac{\|\mathcal{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x})\|}{\|\mathbf{F}(\mathbf{x})\|} \leq \eta \\ & \mathbf{x} \in \Omega \end{aligned} .$$

Neste caso, a função perturbada $\mathcal{F}(\mathbf{x})$ pode ser tomada como a média efetiva, calculada através da Eq. (26), ou selecionando-se o pior valor dentre as M amostras avaliadas na vizinhança δ do vetor \mathbf{x} . Além disso, o parâmetro η de controle da robustez é diretamente proporcional ao nível de robustez desejado, uma vez que quanto maior for o valor definido para o limiar η , a frente de Pareto correspondente estará cada vez mais afastada do conjunto ótimo relacionado ao problema sem ruído. Obviamente, o problema original é recuperado quando se define $\eta = 0$, uma vez que essa condição só é satisfeita quando a função perturbada é igual à original.

3.3.1.4 Técnica de Análise de Robustez com Penalização

Uma das abordagens mais recentes que se tem conhecimento, para avaliação de robustez em problemas de otimização, foi proposta por Mirjalili e Lewis (2016). Similarmente à formulação apresentada na Eq. (17), o objetivo é analisar a robustez através da avaliação da vizinhança de uma dada solução. Para desempenhar esta tarefa, os autores propuseram uma abordagem de penalização da função-objetivo, que prevê uma adição ao seu valor que seja proporcional ao nível de flutuação dos indivíduos na vizinhança de um ponto arbitrário. Dessa forma, a avaliação da robustez de um dado problema de otimização através da abordagem de penalização é da forma

$$\mathbf{F}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{P}(\mathbf{x}) . \quad (27)$$

Em contraste com a otimização robusta apresentada na Eq. (17), nesse caso a função-objetivo permanece intacta e o seu valor aumenta com base na robustez dos candidatos a solução avaliados. No mesmo trabalho, os autores propuseram uma série de problemas de *benchmark* para análise de eficiência de algoritmos na otimização robusta e aplicaram a abordagem de penalização, obtendo resultados satisfatórios. A seguir, a função de penalização da Eq. (27) é detalhada, e também são mostradas mais informações acerca da aplicação da técnica de avaliação de robustez através da penalização da função-objetivo.

Nesta técnica, a função-objetivo é penalizada de acordo com o grau de robustez dos indivíduos localizados na vizinhança de um candidato a solução arbitrário. Com isso, o problema de otimização robusta, que assume um novo formato definido pela Eq. (27), após a

adição do termo de penalização ao problema determinístico, tende a propiciar valores mais elevados, que são proporcionais ao nível de flutuação das amostras aleatórias. Uma vez que o grau de robustez é calculado em relação a um dado ponto no domínio do problema, Mirjalili e Lewis (2016) propuseram uma função de penalização que calcula a diferença média entre um candidato a solução e os indivíduos distribuídos na sua vizinhança. Portanto, o problema determinístico é penalizado de acordo com

$$P(\mathbf{x}^i) = \frac{\frac{1}{M} \sum_{j=1}^H |\mathbf{f}(\mathbf{x}^j) - \mathbf{f}(\mathbf{x}^i)|}{|\mathbf{f}(\mathbf{x}^i)|}, \quad (28)$$

para $i = 1, \dots, k$, onde M representa o número de amostras aleatórias e \mathbf{x}^j denota a j -ésima amostra localizada na vizinhança do candidato a solução \mathbf{x}^i . Originalmente a técnica foi proposta no contexto mono-objetivo e Libotte et al. (2019a) constataram que a formulação pode ser facilmente estendida ao contexto multi-objetivo.

3.3.1.5 Um Primeiro Exemplo

Para consolidar o conceito de otimização robusta, tome como exemplo o problema de minimização unidimensional sem restrições, definido como $x^* = \arg \min F(x)$, onde

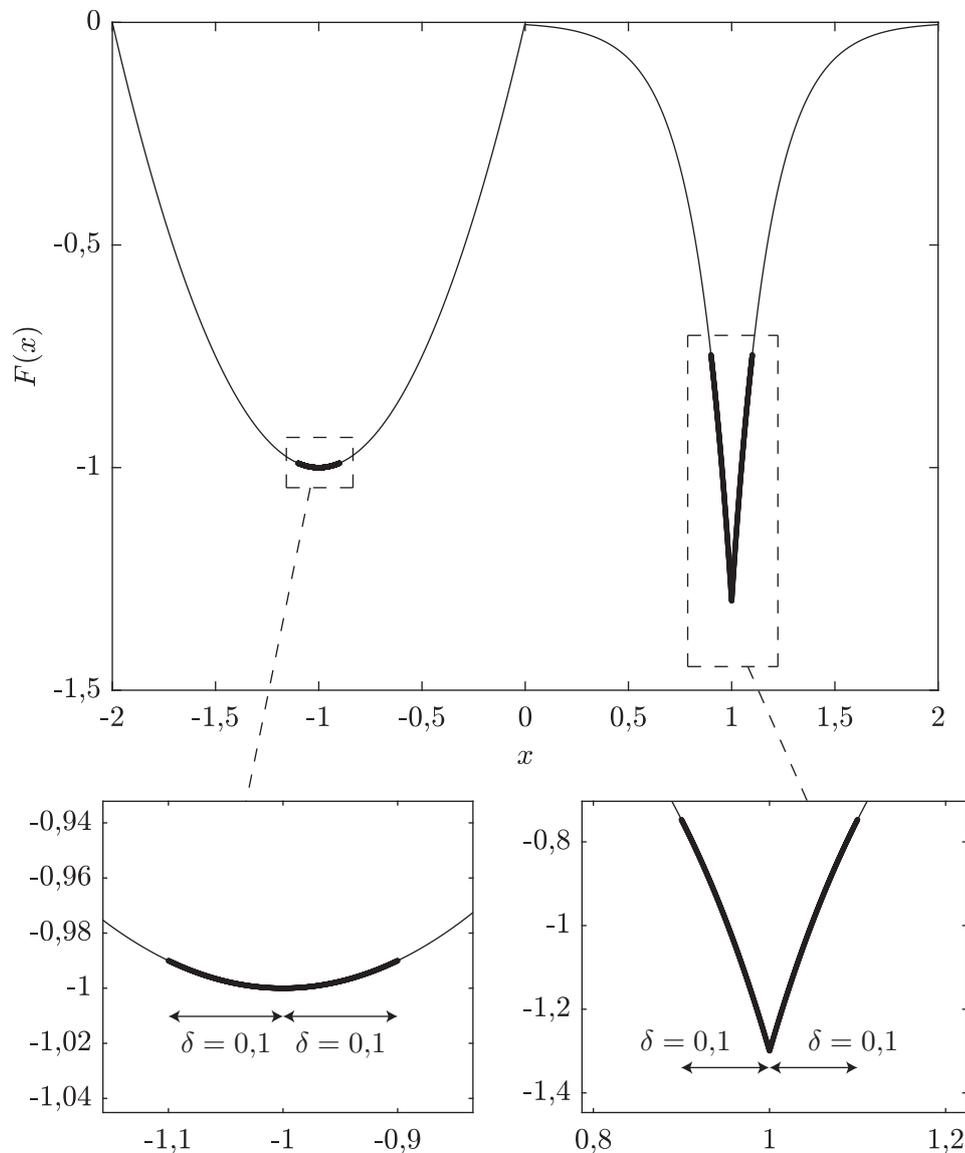
$$F(x) = \begin{cases} (x+1)^2 - 1, & \text{se } -2 \leq x < 0 \\ -c \times 2^{-8|x-1|}, & \text{se } 0 \leq x \leq 2 \end{cases}. \quad (29)$$

Para valores negativos no domínio do problema, a parábola atinge o ponto mínimo quando $x = -1$, tal que $F(-1) = -1$. Por sua vez, o parâmetro c permite deslocar o ponto mínimo da segunda parte da função em relação ao eixo das abscissas, definindo o seu alcance extremo. Para ilustrar a noção de robustez, é conveniente definir $c = 1,3$ e, portanto, para valores positivos de x , a função assume o valor mínimo quando $x = 1$, de tal forma que $F(1) = -c$.

A fim de avaliar os efeitos da análise de robustez, são distribuídas 50 amostras aleatórias ao redor dos dois pontos mínimos do problema (indicados anteriormente), no intervalo $[x - \delta x, x + \delta x]$, sendo x cada uma das soluções analisadas, e considerando uma margem de 10% de incertezas em relação às soluções ótimas, ou seja, fazendo $\delta = 0,1$. A Fig. 14 mostra o comportamento da função que descreve o problema, além das vizinhanças das soluções, nas quais as amostras aleatórias são distribuídas.

Analisando o comportamento da função, é possível perceber que as amostras distribuídas ao redor de $x = 1$ são mais abrangentes do que aquelas localizadas na vizinhança

Figura 14 - Distribuição de amostras aleatórias na vizinhança dos pontos mínimos do problema, com $\delta = 0,1$.



Fonte: O autor, 2018

do ponto mínimo da parábola, que por sua vez estão dispostas de forma menos dispersa, mesmo que o nível de incerteza empregado em ambos os processos de perturbação seja o mesmo. Esta situação se dá pelo fato de que o mínimo global da função apresenta um comportamento muito sensível a qualquer perturbação externa, fazendo com que o valor da função sofra alterações relevantes sob qualquer ligeiro deslocamento dentro do intervalo considerado na distribuição das amostras. Por sua vez, o ponto mínimo da parábola é bem menos sensível às perturbações, haja vista a diferença no valor da função, avaliada nos pontos extremos do intervalo em que as amostras aleatórias são distribuídas, em comparação ao ponto mínimo do problema sem ruído.

Considerando as três técnicas de análise de robustez apresentadas anteriormente, a Tabela 1 mostra o efeito de cada uma destas abordagens, em diferentes níveis de ruído, para cada um dos minimizadores do problema dado pela Eq. (29). Para a técnica de Diferença Normalizada, a função perturbada \mathcal{F} é adotada como o pior valor objetivo dentre todas as amostras avaliadas. Note que, quando o nível de ruído é elevado, o minimizador local $x = -1$ passa a ser o minimizador global, uma vez que a função F apresenta alta sensibilidade em $x = 1$, o minimizador global do problema determinístico.

Tabela 1 - Análise de robustez das soluções $x = -1$ e $x = 1$ do problema dado pela Eq. (29) utilizando diferentes métodos.

Robustez	$x = -1$			$x = 1$		
	Média Efetiva	Dif. Norm.	Penalização	Média Efetiva	Dif. Norm.	Penalização
0%	-1,0000	-1,0000	-1,0000	-1,3000	-1,3000	-1,3000
5%	-0,9992	-1,0000	-0,9992	-1,1353	-0,9852	-1,1733
10%	-0,9967	-1,0000	-0,9967	-0,9979	-0,7467	-1,0676
15%	-0,9925	-1,0000	-0,9925	-0,8826	-0,5659	-0,9789
20%	-0,9867	-1,0000	-0,9867	-0,7855	-0,4289	-0,9042

Fonte: O autor, 2019.

Em geral, pontos extremos mais sensíveis a perturbações tendem a ter um deslocamento mais acentuado após a penalização, uma vez que o seu cálculo considera a diferença média entre diversos os pontos da sua vizinhança. Dependendo do comportamento da função e do nível de ruído determinado, o ponto ótimo do problema determinístico pode deixar de ser a solução do problema de minimização, uma vez que a obtenção da solução robusta pode significar a renúncia do ótimo global do problema determinístico, em favor de uma solução menos sensível às perturbações.

4 OTIMIZAÇÃO BASEADA EM CONFIABILIDADE

Incertezas estão presentes nas observações da maioria dos fenômenos que ocorrem na natureza, o que significa que não podem ser previstos com exatidão. Comumente, observações reiteradas de fenômenos físicos geram múltiplas respostas, sendo que informações que se caracterizam relevantes podem ser obtidas pela diferença na frequência de ocorrência entre elas. Tal falta de padrão de ocorrência é designada como incerteza, aleatoriedade ou estocasticidade. Em geral, todos os parâmetros de um projeto de engenharia que apresentem algum grau de incerteza devem ser considerados como variáveis ou aleatórios.

A tarefa primária do planejamento e projeto de um sistema diz respeito à garantia de desempenho satisfatório. Do ponto de vista das incertezas de um problema, não se pode garantir que o desempenho será adequado. Ao invés disso, a garantia só pode ser atribuída em termos da probabilidade de sucesso relacionada à satisfação de algum critério de desempenho. Na terminologia de engenharia, essa garantia probabilística de desempenho é chamada de confiabilidade.

Existem abordagens tradicionais para o tratamento de falhas em sistemas, que simplificam o problema por meio da consideração dos parâmetros de incerteza através da abordagem dos fatores de segurança. Essa técnica baseia-se em experimentos empíricos e, portanto, não têm garantias quanto à segurança ou rendimento satisfatório, uma vez que não fornece informações no que se refere à influência da variação de parâmetros no sistema. Portanto, o projeto confiável de um sistema de engenharia dinâmico através da utilização da abordagem de fatores de segurança empíricos pode resultar em comportamentos inesperados e dificuldades de variação dos parâmetros que governam o mesmo.

De maneira ampla, as fontes não-cognitivas de incerteza podem ser classificadas em três tipos. Na primeira delas, atribui-se as incertezas às observações físicas. Isso significa que as incertezas são decorrentes das diferenças de resultados coletados durante diversos experimentos, as quais são provocadas pelas flutuações no ambiente, imprecisão de instrumentos, inexatidão nos procedimentos, entre outras ocorrências. Em resumo, dados estatísticos são obtidos através da coleta de uma quantidade considerável de resultados amostrais, os quais estão sujeitos a variações diversas.

Entretanto, a quantidade de experimentos pode estar sujeita a limitações relacionadas à disponibilidade de recursos, como tempo, capital e fatores humanos. Essas restrições levam à segunda fonte de incertezas. Por exemplo, limitação na quantidade de dados para análise, pode resultar em informações imprecisas quanto à variabilidade dos parâmetros observados. Esta incerteza de valores, provocada pela escassez de dados, é conhecida como incerteza estatística.

O terceiro tipo de fonte não-cognitiva de incertezas se refere àquelas relacionadas à modelagem do problema. Modelos computacionais retratam representações aproxima-

das do comportamento de um dado fenômeno ou sistema, capturando suas características essenciais através da idealização por meio de relações matemáticas e procedimentos numéricos. Durante a modelagem, mesmo algumas características ínfimas do sistema ao serem desconsideradas, por imprudência ou de maneira arbitrária, podem acarretar diferenças substanciais entre as previsões computacionais e o comportamento real do fenômeno observado (HALDAR; MAHADEVAN, 2000).

No processo de quantificação de incertezas decorrentes de fontes não-cognitivas, os dados coletados fazem parte do *espaço amostral* ou *domínio amostral*. Uma representação bastante comum da aleatoriedade pode ser dada através da função densidade de probabilidade, a qual é estimada por meio das estatísticas calculadas utilizando os dados amostrais coletados. Com essa ferramenta, é possível estimar o risco de falha de um dado produto ou sistema, para critérios de avaliação específicos. Esses conceitos serão detalhados nas seções subsequentes e fundamentados nas definições apresentadas no Apêndice B.

4.1 Noções de Análise de Confiabilidade

4.1.1 Noção Preliminar de Métodos Probabilísticos

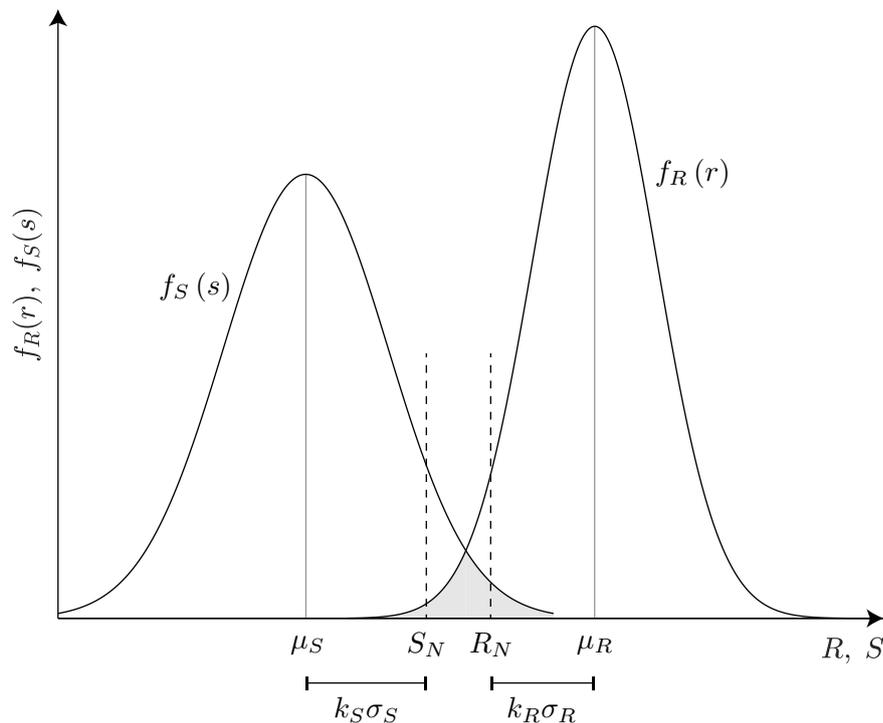
A metodologia de interferência carga-resistência é uma das estratégias mais antigas para a análise de confiabilidade estrutural. Embora existam métodos mais poderosos, isto é, com maior capacidade de generalização e obtenção de resultados mais precisos, essa é uma das ideias mais populares, por conta da sua simplicidade, como será visto mais adiante. Apesar disto, uma grande desvantagem é a necessidade de que carga e resistência sejam estatisticamente independentes, o que pode não ser válido para alguns problemas. Contudo, se essa suposição puder ser justificada, a confiabilidade pode ser calculada de forma relativamente rápida através da definição de fatores de segurança relacionados aos eventos aleatórios avaliados.

O seu nome parece implicar que a confiabilidade é calculada a partir das distribuições de carga e resistência, o que, em certo sentido, é um equívoco, pois ela é aplicável a uma classe mais extensa de problemas. O termo “carga” deve ser considerado em um sentido mais amplo, como qualquer carga aplicada ou quantidade de resposta induzida pela carga que tenha o potencial de causar falha. Por sua vez, o termo “resistência” pode ser compreendido como a capacidade do componente ou sistema em suportar uma certa carga aplicada. No contexto da metodologia de interferência carga-resistência, supõe-se que ocorre uma falha se a carga exceder a resistência, isto é, a probabilidade de falha, ou falta de confiabilidade, é a probabilidade de que a carga seja maior do que a resistência (SUNDARARAJAN; WITT, 1995).

A seguir, é ilustrado um caso simples de avaliação de confiabilidade em um contexto

de análise estrutural, que considera duas variáveis, a saber, a carga imposta sobre uma estrutura, a qual será denotada por S , e a sua capacidade de resistência, representada por R . Em essência, ambas as variáveis têm caráter aleatório e são caracterizadas por suas médias, μ_S e μ_R , desvios-padrão, σ_S^2 e σ_R^2 , e as funções densidade de probabilidade correspondentes, as quais são representada como $f_S(s)$ e $f_R(r)$. A Fig. 15 mostra uma representação gráfica destas informações, exibindo as funções densidade de probabilidade de ambas as variáveis aleatórias no mesmo conjunto de eixos.

Figura 15 - Noção fundamental da avaliação de riscos.



Fonte: Adaptado de Haldar e Mahadevan (2000).

Ainda em relação à Fig. 15, também estão identificados os valores nominais (determinísticos) dos coeficientes de confiabilidade relacionados à abordagem baseada em fatores de segurança, representados por S_N e R_N . Nesta abordagem, a confiabilidade do projeto é assegurada, simplesmente, se a condição $R_N > S_N$ for satisfeita, como indicado anteriormente. Obviamente, o mesmo conceito pode ser aplicado em outras áreas da engenharia, substituindo a concepção de carga e resistência pelas formulações relacionadas à área de interesse.

Em geral, a resistência nominal é um valor conservador, ou seja, adota-se R_N igual a k_R vezes o desvio-padrão da variável aleatória correspondente, abaixo da média. Por sua vez, S_N também é adotado de forma cautelosa, analogamente ao que é estabelecido para a resistência, fazendo $S_N = \mu_S + k_S \sigma_S$. Portanto, o conservadorismo almejado para o projeto, sob a forma de fatores de segurança nominais, depende de diversos fatores associados às

incertezas dos dados relativos às variáveis aleatórias de carga e resistência. Então, os fatores de segurança nominais podem deixar de transmitir a margem real de segurança em um projeto de engenharia, caso as observações dos eventos aleatórios não produzam dados suficientemente consistentes. Obviamente, a decisão sobre os fatores de segurança é influenciada por argumentos econômicos e, por conta da subjetividade ligada ao procedimento, a capacidade de generalização torna-se limitada (HALDAR; MAHADEVAN, 2000).

O propósito de abordagens deste tipo pode ser justificado pela análise da região demarcada entre a interseção das curvas que descrevem o comportamento das funções densidade de probabilidade de R e S , na Fig. 15, a qual denota a ocorrência de falha. De fato, a posição das curvas depende do valor das respectivas médias e desvios-padrão. Através do afastamento dos valores médios, a área referente à interseção entre essas curvas diminui (caso os valores dos desvios-padrão sejam mantidos constantes) e, conseqüentemente, a probabilidade de falha é reduzida. Outro fator que contribui para essa transição relativa é o nível de dispersão dos dados em relação à média, uma vez que a redução do desvio-padrão das variáveis aleatórias também pode favorecer a redução da área de interseção. Apesar da área representada pela interseção das curvas não representar, efetivamente, a probabilidade de falha, ela é qualitativamente proporcional (SUNDARARAJAN; WITT, 1995).

Portanto, o conhecimento de algumas características da análise de confiabilidade através da abordagem de fatores de segurança traz à tona algumas considerações importantes: claramente, esta metodologia tem a desvantagem de, potencialmente, levar a projetos excessivamente conservadores, uma vez que os resultados são derivados de cálculos que consideram problemas com características específicas. Este fato faz com que outra limitação seja evidente, no que diz respeito à generalização da técnica. Além disto, fatores de segurança definidos de forma demasiadamente cautelosa podem acarretar o aumento do custo do projeto. Em vista desses fatos, se faz necessário o desenvolvimento de técnicas de análise de confiabilidade que contribuam com a redução do empirismo, ao mesmo tempo que forneçam recursos para obtenção de resultados consistentes e de maneira mais generalizada possível.

4.1.2 Conceitos Fundamentais

A estratégia de análise de confiabilidade probabilística não enfrenta a limitação sobre a influência na variação dos diferentes parâmetros de um sistema, adversamente ao que ocorre com a abordagem que considera os fatores de segurança, uma vez que são empíricos e, portanto, a experiência obtida por meio de observações é usada como base para mensurar a probabilidade de falha. Neste tipo de análise, a incerteza envolvida em vários parâmetros é modelada usando diferentes funções densidade de probabilidade e o projeto é executado para um nível de confiabilidade especificado (VERMA et al., 2010). O primeiro passo na

avaliação da confiabilidade (ou probabilidade de falha) de um sistema ou produto é a definição acerca do critério de desempenho próprio.

Quando o estado de um sistema excede um limite específico e o seu desempenho é prejudicado, fazendo com que um procedimento previsto não possa ser executado e, conseqüentemente, deixe de gerar o resultado esperado, tal limiar é chamado de *estado limite*. Traçando um paralelo com a análise estrutural, abordada anteriormente, o estado limite indicaria a margem de segurança entre a resistência e a carga das estruturas. Para uma definição matemática, considere que $\mathbf{X} = (X_1, \dots, X_n)$ é um conjunto de n variáveis aleatórias contínuas. A relação funcional, exercida entre essas variáveis, é realizada por meio de uma *função de desempenho*, representada como $Y = g(\mathbf{X})$.

A fronteira entre as regiões segura e de falha, que representa o estado no qual um sistema não pode mais cumprir a função para a qual é projetado, é chamada de *função de estado limite*, representada por $Y = 0$. Neste caso, o critério de falha é definido em relação à inequação $g(\mathbf{X}) \leq 0$. Portanto, o termo fundamental da análise de confiabilidade, que representa a probabilidade de falha de um dado evento aleatório, é definido como

$$p_f = P[g(\mathbf{X}) \leq 0] = \int_{g(\mathbf{x}) \leq 0} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}, \quad (30)$$

onde $f_{\mathbf{x}}(\mathbf{x})$ representa a função densidade de probabilidade conjunta e a integração é realizada sobre a região de falha. A probabilidade de sucesso, denotada por R , equivale ao valor complementar de p_f , isto é,

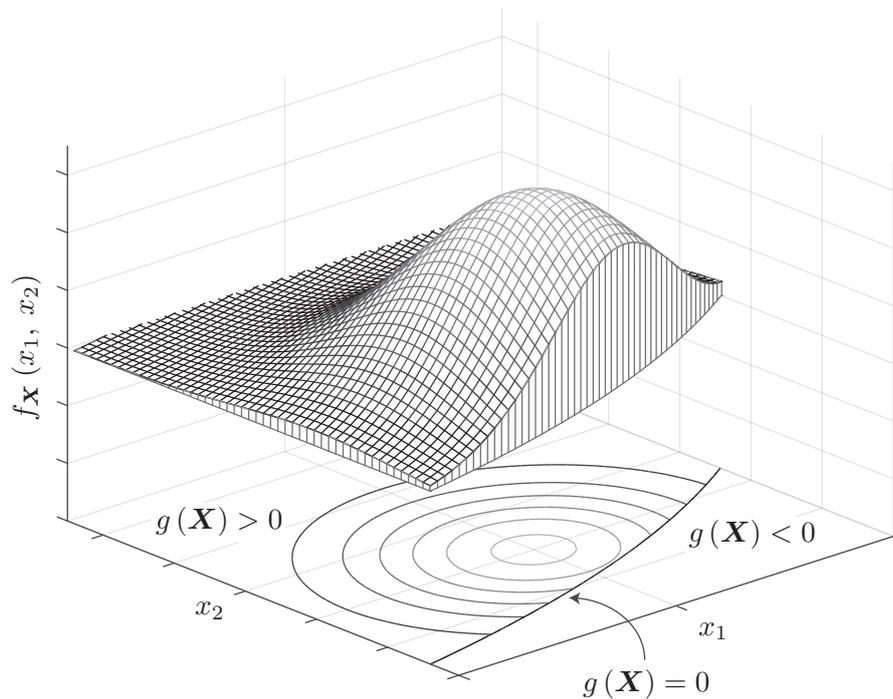
$$R = 1 - p_f = P[g(\mathbf{X}) > 0] = \int_{g(\mathbf{x}) > 0} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}.$$

A Fig. 16 ilustra o volume da função densidade de probabilidade, cuja superfície é limitada pela função de estado limite.

Existem várias dificuldades envolvidas na resolução da integral definida na Eq. (30). Inicialmente, deve-se considerar a possibilidade de que o número de variáveis aleatórias em aplicações práticas seja geralmente alto, envolvendo a necessidade de integração multidimensional. Além disso, $g(\mathbf{X})$ pode ser uma função não-linear e, portanto, o limite de integração também pode ser não-linear. Esses fatores tornam a avaliação da função de desempenho computacionalmente cara, além do fato de que raramente há uma solução de forma fechada para a integração de probabilidade, como ressaltado por Haldar e Mahadevan (2000).

Devido ao esforço computacional excessivo requerido para se obter a solução da Eq. (30), um tratamento possível é usar aproximações analíticas de tais integrais, que são mais simples de se calcular. Nos últimos anos, diversos métodos têm sido propostos, visando a obtenção da probabilidade de falha de sistemas através de uma metodologia mais

Figura 16 - Esquema genérico do cálculo da probabilidade de sucesso através da integração da função densidade de probabilidade em $g(\mathbf{X}) > 0$.



Fonte: O autor, 2018.

eficiente, do ponto de vista computacional. A seguir é feita uma revisão dos principais métodos de análise de confiabilidade da literatura.

4.2 Métodos de Análise de Confiabilidade

Nesta seção, são apresentados os aspectos gerais dos principais métodos capazes de realizar a estimativa da probabilidade de falha de sistemas, considerando possíveis incertezas no modelo matemático. Baseado na classificação proposta por Bichon et al. (2007), os métodos de análise de confiabilidade foram agrupados em três categorias distintas: métodos locais (EL HAMI et al., 2017a), métodos de amostragem (EL HAMI et al., 2017b) e métodos globais (EL HAMI et al., 2017c). Os métodos de confiabilidade de primeira e segunda ordens (conhecidos na literatura como FORM e SORM, em referência aos respectivos acrônimos em inglês), que se encaixam na categoria de métodos locais, serão abordados mais detalhadamente a frente, uma vez que serão essenciais no desenvolvimento desta tese.

É importante salientar que, além dos métodos abordados nesta seção, existem diversos outros que, grosso modo, se tratam de modificações daqueles destacados neste trabalho. Na maior parte dos casos, essas modificações visam melhorar o desempenho no cômputo dos resultados, aumentando a capacidade de obtenção de soluções, através do menor custo

computacional praticável. Evidentemente, apenas as estratégias principais de cada categoria são descritas, de maneira concisa, uma vez que o propósito desta revisão é exibir os tipos distintos de metodologias para análise de confiabilidade, suas principais características e vantagens.

4.2.1 Método Local

4.2.1.1 Método de Superfície de Resposta

Em diversas aplicações práticas, é possível que não se conheça a forma explícita da função de desempenho. Como alternativa a isto, o seu comportamento pode ser ajustado, através da utilização de dados experimentais, por meio de uma função polinomial, cuja simplicidade de manipulação e avaliação são vantagens relevantes. Em geral, métodos de superfície de resposta trata-se de ferramentas estatísticas e matemáticas utilizadas na exploração de relações entre variáveis independentes e variáveis de resposta (MYERS et al., 2016).

O conceito principal na abordagem dos métodos de superfície de resposta aplicados aos problemas de confiabilidade é o ajuste de pontos discretos que aproximem a função de estado limite na vizinhança do ponto mais provável de falha (TAN et al., 2013). A precisão dos resultados depende muito da precisão com que as características da função de estado limite original são aproximadas. Os coeficientes indeterminados do polinômio podem ser estimados pelo Método de Mínimos Quadrados. Em geral, um polinômio de grau dois é suficiente para aproximar soluções, uma vez que polinômios de ordem superior podem conduzir a um problema mal-condicionado (MELCHERS, 1999 apud YANG et al., 2015).

A forma mais comum da função de superfície de resposta, representada por um polinômio quadrático sem termos cruzados, é definida como

$$\hat{g}(\mathbf{x}) = a + \sum_{j=1}^n b_j x_j + \sum_{j=1}^n c_j x_j^2,$$

onde x_j representa a j -ésima variável aleatória, com $j = 1, \dots, n$, tal que n é o número total de variáveis do problema, e os termos a , b_j e c_j correspondem às constantes do polinômio a serem estimadas. Uma vez que a função de superfície de resposta tenha sido aproximada, poderia ser calculada a integral da Eq. (30), o que no entanto não é muito eficiente.

Alternativamente, pode-se utilizar um método de análise de confiabilidade para o cálculo da probabilidade de falha (como por exemplo o FORM, que será visto na seção sobre métodos de confiabilidade de primeira e segunda ordens, o qual se baseia no cálculo do índice de confiabilidade). O uso de interpolação utilizando as informações referentes às

amostras empregadas na obtenção do polinômio que melhor ajusta a superfície de resposta pode ser usada para aperfeiçoar a capacidade de predição do índice de confiabilidade, como proposto por Bucher e Bourgund (1990) e aplicado por Tan et al. (2013).

4.2.2 Métodos de Amostragem

4.2.2.1 Método de Simulação de Monte Carlo

Em geral, métodos de amostragem são empregados em casos onde deseja-se obter o resultado de uma classe de problemas com solução analítica restrita, quando não se pode aplicar métodos diretos. Além disso, por conta dos requisitos menos rigorosos, no que diz respeito à sua aplicabilidade em problemas com propriedades complexas, e a característica de serem facilmente utilizados, o seu uso também torna-se viável na obtenção de soluções para funções do tipo “caixa-preta”, quando não se tem informações relevantes do comportamento do problema (BJERAGER, 1991).

O método elementar de simulação de Monte Carlo, conhecido como método baseado no indicador zero-um, fundamenta-se na implementação de uma função indicadora. Seja $\mathbf{x} \in \mathbb{R}^n$ o vetor de variáveis de decisão e $g : \mathbb{R}^n \rightarrow \mathbb{R}$ a função de desempenho. A função indicadora binária, representada por $J_g(\mathbf{x})$, é definida como

$$J_g(\mathbf{x}) = \begin{cases} 1, & \text{se } g(\mathbf{x}) \leq 0 \\ 0, & \text{se } g(\mathbf{x}) > 0 \end{cases} . \quad (31)$$

Portanto, a integral que estima a probabilidade de falha de um evento, dada pela Eq. (30), pode ser aproximada utilizando a abordagem de Monte Carlo através de

$$p_f = \int_{g(\mathbf{x}) \leq 0} f_x(\mathbf{x}) d\mathbf{x} \approx \int_{-\infty}^{\infty} J_g(\mathbf{x}) f_x(\mathbf{x}) d\mathbf{x} , \quad (32)$$

onde $f_x(\mathbf{x})$ representa a função densidade de probabilidade.

Realizando N simulações sobre o vetor \mathbf{X} , a probabilidade de falha p_f equivale à média dos valores da amostra, isto é,

$$\hat{p}_f = \frac{1}{N} \sum_{i=1}^N J_g(\mathbf{x}^{(i)}) = \frac{N_f}{N} , \quad (33)$$

sendo N_f a quantidade total de amostras testadas nas quais $J_g(\mathbf{x}^{(i)}) = 1$. Levando em conta a variância amostral $\text{Var}[\hat{p}_f] = \hat{p}_f(1 - \hat{p}_f)/N$, a medida da taxa de convergência do método

de Monte Carlo pode ser obtida através do coeficiente de variação dado por

$$CV[\hat{p}_f] = \frac{\sqrt{\text{Var}[\hat{p}_f]}}{\mathbb{E}[\hat{p}_f]} = \sqrt{\frac{1-p_f}{Np_f}}, \quad (34)$$

tal que $\mathbb{E}[\hat{p}_f] = p_f$ representa o estimador não-viesado da probabilidade de falha. De acordo com El Hami et al. (2017b), para $CV[\hat{p}_f] = 0,1$ e probabilidade de falha fixada em $p_f = 10^{-m}$, são necessárias, aproximadamente, $N = 10^{m+2}$ amostras para uma estimativa de probabilidade de falha eficiente.

4.2.2.2 Amostragem de Importância

A ideia fundamental da amostragem de importância gira em torno de concentrar a distribuição dos pontos amostrais na região de maior importância, isto é, na área que apresenta a maior contribuição efetiva para o cálculo da probabilidade de falha, ao invés de propagar as amostras por todo o domínio viável de possíveis valores das variáveis de projeto. Um método capaz de alcançar este propósito foi elaborado por Harbitz (1986). O esquema conceitual do funcionamento do método é apresentado na Fig. 17.

Uma vez que não ocorre falha na região interna da hiperesfera de raio β , sugere-se uma eficiente redução na variância dos dados quando realizada a amostragem apenas na região externa à esta hiperesfera. A prática dessa formulação exige que uma nova função densidade de probabilidade seja escolhida, de forma que as amostras só sejam avaliadas na região de interesse. Nesse contexto, seja $\bar{f}_x(\mathbf{x})$ a função densidade de probabilidade da variável aleatória \mathbf{X} , definida no espaço amostral de interesse, isto é, na região externa à hiperesfera de raio β . Considerando a Eq. (32), a probabilidade de falha pode ser estimada pelo Método de Amostragem de Importância através da modificação estabelecida como

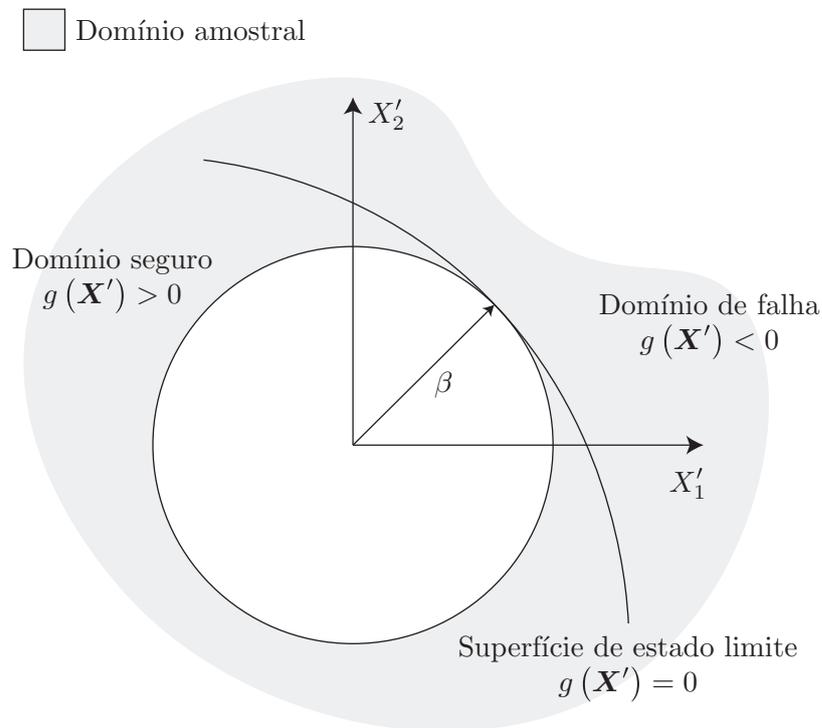
$$p_f = \int_{g(\mathbf{x}) \leq 0} \left[J_g(\mathbf{x}) \frac{f_x(\mathbf{x})}{\bar{f}_x(\mathbf{x})} \right] \bar{f}_x(\mathbf{x}) d\mathbf{x}.$$

Portanto, de maneira análoga à Eq. (33), obtém-se a probabilidade de falha através da amostragem na região de interesse por

$$\hat{p}_f = \frac{1}{N} \sum_{i=1}^N J_g(\mathbf{x}^{(i)}) \frac{f_x(\mathbf{x}^{(i)})}{\bar{f}_x(\mathbf{x}^{(i)})},$$

com N sendo a quantidade de simulações realizadas. Obviamente, a precisão do método é diretamente dependente da escolha de $\bar{f}_x(\mathbf{x})$. Em vista dessa necessidade, diversos autores propuseram maneiras sistemáticas de se realizar tal escolha, dentre as quais pode-se

Figura 17 - Esquema conceitual do Método de Amostragem de Importância de Harbitz.



Fonte: Adaptado de Haldar e Mahadevan (2000).

destacar os trabalhos de Karamchandani et al. (1989) e Melchers (1989).

4.2.2.3 Amostragem por Hipercubo Latino

O método de Amostragem por Hipercubo Latino é sugerido como uma ferramenta alternativa para melhorar a eficiência de diferentes Métodos de Amostragem de Importância. Assim como a maioria dos métodos deste tipo, a precisão desejada para os resultados é absolutamente dependente da quantidade de simulações. A abordagem utiliza um esquema de amostragem estratificada (de maneira que as amostras são divididas em subgrupos) para melhorar a cobertura do espaço amostral (OLSSON et al., 2003).

A amostragem estratificada realizada pelo método de Hipercubo Latino fornece a garantia de que todas as áreas do espaço do projeto estão representadas. O espaço amostral é dividido em m regiões mutuamente exclusivas, denotadas por R_m e amostras aleatórias de cada região são tomadas com base na distribuição atribuída a cada variável de entrada (ZENTNER, 2006 apud ROSHANIAN; EBRAHIMI, 2013). Portanto, a equação para calcu-

lar a confiabilidade é

$$\hat{P}_f = \sum_{j=1}^m \left[P [R_j] \frac{1}{N_j} \sum_{i=1}^{N_j} J_g (\mathbf{x}^{(i)}) \right].$$

Nessa estimativa, $P [R_j]$ representa a probabilidade calculada na região R_j , N_j é o número de amostras distribuídas em uma dada região j e $J_g (\mathbf{x})$ é a função indicadora, como definida na Eq. (31).

4.2.3 Métodos Globais

Os métodos de análise de confiabilidade locais são amplamente utilizados por conta da sua eficiência e moderado custo computacional. Contudo, podem não ser tão eficazes em casos onde a função a ser aproximada não é suficientemente suave na vizinhança do ponto mais provável de falha, ou em casos onde o problema de otimização assume diversas soluções ótimas (problemas multimodais). Além do que, os resultados obtidos podem não ser tão precisos, uma vez que as aproximações da função de estado limite apresentam erros associados ao truncamento da série de Taylor, especialmente quando tais funções possuem alto grau de não-linearidade. Finalmente, em aplicações de engenharia, as funções de desempenho são frequentemente caracterizadas por funções de resposta implícitas, que podem ter custo computacional elevado para avaliação.

Neste cenário, uma alternativa que possibilita contornar essas adversidades é a utilização dos métodos de amostragem, que não apresentam limitações quanto à disposição e comportamento da função de estado limite, tampouco à convergência para o ponto mais provável de falha. A principal limitação desse tipo de método é que, usualmente, necessita-se de uma grande quantidade de avaliações da função, a fim de que seja realizada uma estimativa apropriada, o que pode tornar o procedimento impraticável (no sentido do tempo necessário), caso o problema a ser resolvido apresente custo elevado de avaliação, como por exemplo quando se obtém os dados amostrais através de métodos de elementos finitos. Portanto, é necessário um método de análise de confiabilidade que seja eficiente quando aplicado a funções de desempenho dispendiosas e preciso para uma função de resposta de qualquer forma arbitrária (EL HAMI et al., 2017c).

Considerando essas necessidades, surge uma classe de métodos, denominados globais, que buscam englobar os conceitos essenciais dos métodos locais e de amostragem, com a finalidade de evitar os problemas enfrentados pelas técnicas de ambos os tipos. Duas abordagens típicas deste tipo de método são o *Efficient Global Reliability Analysis* (EGRA) (BICHON et al., 2008) e o *Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation* (AK-MCS) (ECHARD et al., 2011). Essencialmente, em ambas as técnicas,

é utilizada a metodologia de *kriging* para construir um metamodelo que melhor aproxime a função de estado limite original. Em seguida, esse modelo aproximado é utilizado para calcular a probabilidade de falha através de uma técnica de amostragem. A descrição detalhada de ambos pode ser vista nos trabalhos de referência e em El Hami et al. (2017c).

5 MÉTODOS DE CONFIABILIDADE DE PRIMEIRA E SEGUNDA ORDENS

Neste capítulo, são apresentados os métodos de confiabilidade de primeira e segunda ordens, capazes de estimar o índice de confiabilidade de um dado problema. Em seguida, a abordagem de análise de confiabilidade inversa é apresentada, a qual destina-se a obter um resultado que satisfaça os níveis de confiabilidade desejados, diferentemente das demais estratégias que são apresentadas, as quais representam abordagens diretas. Apesar destas distinções, todas compartilham de conceitos em comum, que são detalhados a seguir.

5.1 Método de Confiabilidade de Primeira Ordem

A classe de Método de Confiabilidade de Primeira Ordem (FORM) são baseados na utilização do primeiro e segundo momentos das variáveis aleatórias, ou seja, os valores da tendência central e da dispersão, para o cálculo da confiabilidade. Nessa classe, tem-se os métodos conhecidos como *Mean Value First-order Second-moment* (MVFOSM) e *Advanced First-order Second-moment* (AFOSM), que são os precursores das técnicas de análise de confiabilidade de primeira ordem. É comum ver o método AFOSM ser referido na literatura simplesmente como FORM. Na realidade, qualquer método que aproxime a função de estado limite pela expansão de Taylor até o termo de primeira ordem pode ser referido como FORM. Apesar disso, o constante uso do termo para se referir ao método AFOSM fez com que se criasse um legado. (CHOI et al., 2007).

5.1.1 Mean Value First-order Second-moment

Considere uma situação em que as únicas informações disponíveis sobre as variáveis aleatórias são seus primeiros e segundos momentos, ou seja, é possível calcular as médias, desvios-padrão e coeficientes de correlação. Sob essas circunstâncias, o método MVFOSM foi originalmente formulado, por Cornell (1969), cujo nome também é referido como *First-order Second-moment* (FOSM) na literatura. A sua idealização tem como base a abordagem de interferência carga-resistência, apresentada na Seção 4.1.1. Considere a função de desempenho definida por $Q = g(R, S) = R - S$. Assumindo que as variáveis aleatórias contínuas R e S são estatisticamente independentes e possuem distribuição normal, então Q também apresenta distribuição normal.

Nesse caso, a ocorrência de falhas se dá quando a carga é menor do que a resistência, isto é, $R \leq S$ ou $Q \leq 0$. O objetivo do método é obter uma estimativa para a probabilidade de falha, isto é, a probabilidade de $Q \leq 0$, o que equivale a $p_f = P[Q \leq 0]$. Recorrendo à trans-

formação $Z = (Q - \mu_Q) / \sigma_Q$, onde μ_Q e σ_Q são a média e desvio-padrão de Q , a probabilidade de ocorrência de falha pode ser calculada por

$$p_f = \Phi_Z(0) = \Phi_Z\left(-\frac{\mu_Q}{\sigma_Q}\right), \quad (35)$$

onde $\Phi_Z(z)$ simboliza a função de distribuição acumulada da distribuição normal padrão (HALDAR; MAHADEVAN, 1995).

Como R e S são variáveis aleatórias estatisticamente independentes, é possível afirmar¹ que $\mu_Q = \mu_R - \mu_S$ e $\sigma_Q = \sqrt{\sigma_R^2 + \sigma_S^2}$. Com isso, a Eq. (35) pode ser reescrita, na forma $p_f = \Phi_Y(-\beta)$, com

$$\beta = \frac{\mu_Q}{\sigma_Q} = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}}.$$

A razão definida por β é conhecida como *índice de Cornell*. Multiplicando ambos os lados de β por σ_Q , fica claro que esse índice pode ser interpretado como o número de desvios-padrão, desde a média μ_Q , até a região de falha. Quanto mais desvios-padrão o domínio de falha estiver longe da média, mais segura será a aplicação. Em outras palavras, quanto maior o índice de Cornell, menor é a probabilidade de falha. Por outro lado, valores pequenos desse índice indicam que o domínio da falha está mais próximo da média, o que implica uma maior probabilidade de falha.

A generalização dessa ideia é efetuada por meio da aproximação da função de estado limite por expansão em série de Taylor ao redor da média, como mostrado por Choi et al. (2007). Nesse contexto, considere a representação de uma função de desempenho generalizada, na forma $Q = g(\mathbf{X})$. A expansão em série de Taylor, em torno da média $\mu_{\mathbf{X}}$, tem a forma

$$Q = g(\mu_{\mathbf{X}}) + \sum_{i=1}^n \frac{\partial g(\mu_{\mathbf{X}})}{\partial x_i} (x_i - \mu_{x_i}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 g(\mu_{\mathbf{X}})}{\partial x_i \partial x_j} (x_i - \mu_{x_i})(x_j - \mu_{x_j}) + \dots,$$

onde μ_{x_i} representa a média da variável aleatória X_i , para $i = 1, \dots, n$, e $\mu_{\mathbf{X}} = (\mu_{x_1}, \dots, \mu_{x_n})$. Se a série for truncada após a expansão do termo linear e a expressão correspondente for substituída na Eq. (118), obtém-se uma aproximação para μ_Q , que resulta em $\mu_Q \approx g(\mu_{\mathbf{X}})$.

No caso da aproximação relacionada à variância, deve-se considerar a relação dada por $\sigma_Q^2 = \mathbb{E} \left[(g(\mathbf{X}) - \mu_Q)^2 \right]$, que decorre imediatamente da Eq. (119). Novamente, substituindo $g(\mathbf{X})$ pela aproximação linear dada pela série de Taylor truncada e, após algumas

¹ A demonstração destes resultados pode ser consultada em Meyer (1987, pp. 152, 160)

manipulações algébricas, obtém-se

$$\sigma_Q^2 \approx \sum_{i=1}^n \sum_{j=1}^n \frac{\partial g(\mu_x)}{\partial x_i} \frac{\partial g(\mu_x)}{\partial x_j} \text{Cov}(X_i, X_j),$$

onde $\text{Cov}(X_i, X_j)$ representa a covariância de X_i e X_j , detalhada na Eq. (122). No caso de variáveis aleatórias não-correlacionadas, então a variância resulta em

$$\sigma_Q^2 \approx \sum_{i=1}^n \left(\frac{\partial g(\mu_x)}{\partial x_i} \right)^2 \text{Var}[X_i].$$

O índice de Cornell só apresenta acurácia tolerável quando as variáveis aleatórias que descrevem a função de desempenho têm distribuição normal (THOFT-CRISTENSEN; BAKER, 1982). Além disso, a avaliação da confiabilidade linearizando a função de estado limite ao redor dos valores médios leva a estimativas errôneas para funções de desempenho com alta grau de não-linearidade, ou no caso de coeficientes de variação elevados. O método MVFOSM falha em ser invariante a diferentes formulações matematicamente equivalentes do mesmo problema. Esta é uma adversidade não apenas para formas não-lineares de $g(\mathbf{X})$, mas também para certas formulações lineares. Esse problema é consequência, possivelmente, do fato de que o MVFOSM não utiliza informações acerca da distribuição de probabilidade das variáveis aleatórias, quando disponível, diferentemente do que ocorre com o método que será apresentado a seguir.

5.1.2 Advanced First-order Second Moment

Na Seção 4.1.2, foram evidenciadas as dificuldades de se computar o resultado da integral mostrada na Eq. (30), a qual equivale à probabilidade de falha de um evento aleatório, em termos da função de desempenho $g(\mathbf{X})$. Nesse sentido, o método AFOSM foi proposto, com a intenção pontual de simplificar o cômputo dessa integral e obter o valor da probabilidade de falha através de um procedimento menos complexo, superando as dificuldades relacionadas a esse cálculo.

Na tentativa de simplificar o cômputo da estimativa da probabilidade de falha e também evitar as limitações presentes no método MVFOSM, apresentadas anteriormente, Hasofer e Lind (1974) idealizaram o AFOSM. Os autores propuseram um mapeamento linear das variáveis aleatórias com distribuição normal em um conjunto de variáveis normalizadas e independentes. O método segue três etapas fundamentais: (i) transformação da variável aleatória \mathbf{X} de seu espaço original em um problema equivalente no espaço normal padrão \mathbf{U} , (ii) simplificação do integrando $f_x(\mathbf{x})$ e dos limites de integração da Eq. (30) e (iii) estimativa da probabilidade de falha correspondente à função de estado limite aproximada no

espaço \mathbf{U} (BJERAGER, 1991). Seguindo a tendência da literatura, desse ponto em diante a metodologia a ser apresentada será designada como FORM.

5.1.2.1 Transformação de Variáveis

Na primeira etapa do algoritmo, o procedimento de transformar as variáveis aleatórias do problema original em um novo problema similar, com variáveis no espaço normal padrão (espaço \mathbf{U}), sem qualquer perda de precisão, ocorre através da chamada *transformação de Rosenblatt* (ROSENBLATT, 1952). Seja $F_{\mathbf{x}}(\mathbf{x})$ a função de distribuição acumulada da variável aleatória \mathbf{X} , então a transformação é expressa por

$$\mathbf{u} = \Phi^{-1}(F_{\mathbf{x}}(\mathbf{x})) , \quad (36)$$

onde Φ representa a função de distribuição acumulada da distribuição normal padrão. Transformações deste tipo são sempre possíveis para variáveis aleatórias contínuas. Na verdade, o número de transformações possíveis é infinito, como apontado por Bjerager (1991). Uma transformação particularmente simples pode ser usada quando as variáveis aleatórias são mutuamente independentes. Remetendo à Seção B.5.1, cada variável pode ser transformada, considerando $\mathbf{X} \sim \mathcal{N}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}^2)$, por meio de

$$\mathbf{u} = \Phi^{-1}\left(\Phi\left(\frac{\mathbf{x} - \mu_{\mathbf{x}}}{\sigma_{\mathbf{x}}}\right)\right) . \quad (37)$$

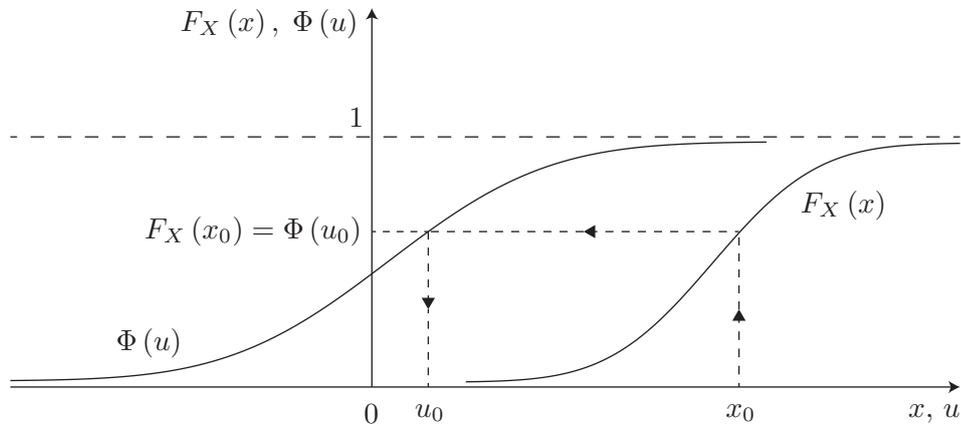
Portanto, uma possível transformação é estabelecida pela relação $\mathbf{x} = \mu_{\mathbf{x}} + \sigma_{\mathbf{x}}\mathbf{u}$. A Fig. 18 ilustra essa transformação, mostrando ambas as funções de distribuição acumulada no mesmo conjunto de eixos.

Com base na transformação da Eq. (37), a média da variável aleatória original \mathbf{X} é mapeada na origem do espaço normal padrão \mathbf{U} . Além disso, a função de estado limite original $g(\mathbf{X}) = 0$ é mapeada para o espaço \mathbf{U} , passando a ser representada por $G(\mathbf{U}) = 0$. Após esta transformação, a probabilidade de falha mostrada na Eq. (30) passa a ser calculada, de forma equivalente, através da expressão

$$p_f = P[G(\mathbf{U}) \leq 0] = \int_{G(\mathbf{U}) \leq 0} \phi(\mathbf{u}) d\mathbf{u} , \quad (38)$$

onde $\phi(\mathbf{u})$ denota a função densidade de probabilidade conjunta da distribuição normal padrão. É importante notar que a transformação conserva a probabilidade e, portanto, este procedimento não insere qualquer tipo de erro adicional ao cálculo da probabilidade de falha.

Figura 18 - Transformação da variável aleatória X para o espaço normal padrão U .



Fonte: Adaptado de Bjerager (1991).

5.1.2.2 Simplificação do Integrandando e dos Limites de Integração

Como todas as variáveis aleatórias são independentes, o produto das funções densidade de probabilidade individuais corresponde à função densidade de probabilidade conjunta (DU, 2005). Com base nessa afirmação,

$$\phi(\mathbf{u}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u_i^2\right) \quad (39)$$

e calcular a probabilidade de falha, por meio da integração da Eq. (38), é equivalente a obter a integral

$$p_f = \int_{G(\mathbf{U}) \leq 0} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \sum_{i=1}^n u_i^2\right) d\mathbf{u} . \quad (40)$$

Após as transformações, a formulação matemática da função de desempenho sofre variações. Naturalmente, a sua representação passa a ser feita por $G(\mathbf{U})$, representando a função de desempenho após a transformação de variáveis para o espaço normal padrão.

Na segunda etapa do FORM, os limites de integração também devem ser simplificados, complementarmente às mudanças realizadas no integrando da expressão responsável pelo cálculo da probabilidade de falha. Como já foi mencionado, a denominação do FORM é justificada pelo fato de que a função de desempenho é simplificada através de série de Taylor de primeira ordem, isto é, $G(\mathbf{u}) \approx L(\mathbf{u})$, onde a aproximação linear é dada por

$$L(\mathbf{u}) = G(\mathbf{u}^*) + \sum_{i=1}^n \frac{\partial G(\mathbf{u}^*)}{\partial u_i} (u_i - u_i^*) . \quad (41)$$

Por meio dessa formulação, manifesta-se uma questão importante: ao redor de que ponto a série de Taylor deve ser expandida, isto é, qual é o valor ótimo para \mathbf{u}^* ?

Diferentemente da transformação realizada no primeiro passo do método, a aproximação da função de desempenho incorpora determinados erros ao cálculo da probabilidade de falha. Truncar a série de Taylor após a expansão do termo linear pode representar uma diferença significativa na estimativa dessa probabilidade. Em contrapartida, a utilização de aproximações de ordem superior pode refletir no aumento do custo computacional, mesmo que isso signifique a obtenção de resultados mais consistentes. Em vista deste cenário, o método FORM tem o objetivo de simplificar os cálculos, propiciando um balanço entre custo computacional e precisão. No caso de uma aproximação linear, a minimização da perda de precisão ocorre quando a expansão da série de Taylor é realizada ao redor do ponto que possui a maior densidade de probabilidade, ou seja, o valor máximo do integrando na Eq. (40). O ponto que apresenta o valor máximo da função densidade de probabilidade, sobre a função de estado limite, é chamado de *ponto mais provável de falha* e representado como

$$\mathbf{u}^* = \begin{cases} \arg \max_{\mathbf{u}} \phi(\mathbf{u}) \\ \text{Sujeito a } G(\mathbf{u}) = 0 \end{cases}, \quad (42)$$

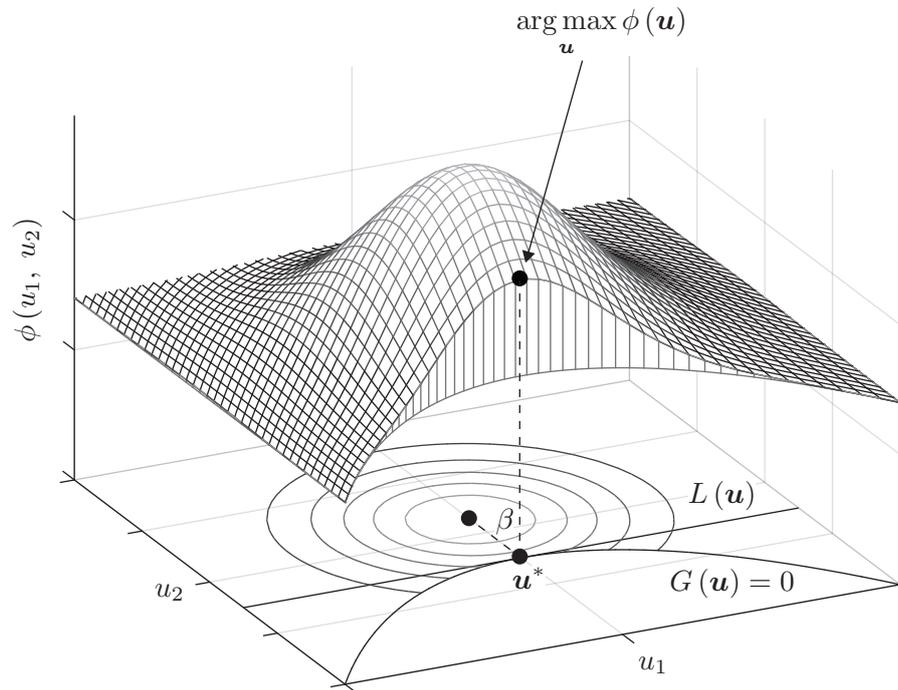
com $\phi(\mathbf{u})$ dado pela Eq. (39).

Assim, o problema de se obter o ponto mais provável de falha da Eq. (42) é equivalente a minimizar o somatório do argumento da função exponencial, cujo problema equivalente pode ser reescrito como

$$\mathbf{u}^* = \begin{cases} \arg \min_{\mathbf{u}} \mathcal{D}(\mathbf{u}) \\ \text{Sujeito a } G(\mathbf{u}) = 0 \end{cases}, \quad (43)$$

onde $\mathcal{D}(\mathbf{u}) = \|\mathbf{u}\|$. Na literatura, esta abordagem é conhecida como RIA (*Reliability Index Approach*) (NIKOLAIDIS; BURDISSO, 1988). Portanto, o uso do FORM (e também do SORM, como será visto na Seção 5.2) essencialmente substitui uma integração multidimensional por um problema de otimização matemática para encontrar o ponto mais provável de falha. Como o valor da função densidade de probabilidade $\phi(\mathbf{u})$ diminui rapidamente à medida que \mathbf{u} se afasta da origem, é natural que a principal contribuição na integral da Eq. (40) seja de valores próximos à origem (BJERAGER, 1991). A menor distância, da origem até o ponto mais provável de falha, representada por $\beta = \|\mathbf{u}^*\|$, é conhecida como *índice de confiabilidade*. A Fig. 19 mostra um esquema desta simplificação.

Figura 19 - Obtenção do ponto mais provável de falha (\mathbf{u}^*) no FORM.



Fonte: O autor, 2019.

5.1.2.3 Obtenção do Ponto Mais Provável de Falha

O problema de otimização apresentado na Eq. (43) foi resolvido, de forma genérica, por Shinozuka (1983). Para a obtenção dos resultados, foi aplicado o método dos multiplicadores de Lagrange, que permite encontrar pontos extremos de funções, sujeitas a restrições. Os resultados são calculados por meio da construção de uma função auxiliar $\Lambda(\mathbf{u}, \lambda)$, que tem a forma

$$\Lambda(\mathbf{u}, \lambda) = \sqrt{\mathbf{u}^T \mathbf{u}} + \lambda G(\mathbf{u}),$$

tal que λ representa o multiplicador de Lagrange. Se $\mathcal{D}(\mathbf{u}^0)$ representa um ponto de máximo para o problema original, então existe um multiplicador de Lagrange, tal que $(\mathbf{u}^0, \lambda^0)$ é um ponto estacionário para a função $\Lambda(\mathbf{u}, \lambda)$, isto é, existe um ponto para o qual as derivadas parciais de $\Lambda(\mathbf{u}, \lambda)$ são iguais a zero.

Portanto, procedendo com $\partial \Lambda(\mathbf{u}, \lambda) / \partial \mathbf{u} = 0$ e $\partial \Lambda(\mathbf{u}, \lambda) / \partial \lambda = 0$, obtém-se um sistema com $n + 1$ equações e $n + 1$ incógnitas, da forma

$$\begin{cases} \frac{\mathbf{u}}{\mathcal{D}(\mathbf{u})} + \lambda \nabla G|_{\mathbf{u}} = 0 \\ G(\mathbf{u}) = 0 \end{cases}.$$

Observe que a segunda equação deste sistema corresponde à restrição do problema de otimização original, mostrado na Eq. (43). Então, para se obter o valor do multiplicador de Lagrange, note que

$$\mathbf{u} = -\lambda \mathcal{D}(\mathbf{u}) \nabla G|_{\mathbf{u}} . \quad (44)$$

Multiplicando ambos os lados da Eq. (44) por \mathbf{u}^T , tem-se

$$\mathbf{u}^T \mathbf{u} = -\lambda \mathbf{u}^T \mathcal{D}(\mathbf{u}) \nabla G|_{\mathbf{u}} ,$$

cujos lado esquerdo pode ser representado como $\mathbf{u}^T \mathbf{u} = (\mathcal{D}(\mathbf{u}))^2$, resultando em

$$\mathcal{D}(\mathbf{u}) = -\lambda \mathbf{u}^T \nabla G|_{\mathbf{u}} . \quad (45)$$

Tomando \mathbf{u}^T na Eq. (44) e substituindo o termo resultante na Eq. (45), obtém-se

$$\mathcal{D}(\mathbf{u}) = -\lambda \left(-\lambda \mathcal{D}(\mathbf{u}) (\nabla G|_{\mathbf{u}})^T \right) \nabla G|_{\mathbf{u}} . \quad (46)$$

Nestas condições, é conveniente dividir a Eq. (46) por $\mathcal{D}(\mathbf{u})$ e, através do agrupamento dos termos semelhantes, o problema reduz-se a

$$1 = \lambda^2 (\nabla G|_{\mathbf{u}})^T \nabla G|_{\mathbf{u}} ,$$

de onde o multiplicador de Lagrange pode ser obtido, o qual resulta em

$$\lambda = \frac{1}{\sqrt{(\nabla G|_{\mathbf{u}})^T \nabla G|_{\mathbf{u}}}} . \quad (47)$$

De posse do resultado do multiplicador de Lagrange, é possível obter uma expressão para o índice de confiabilidade. Quando λ assume o valor dado pela Eq. (47), a restrição do problema definido pela Eq. (43) é satisfeita e, portanto, $\mathcal{D}(\mathbf{u}) = \mathcal{D}(\mathbf{u}^*)$. Com isso, se a solução obtida para o multiplicador de Lagrange, expressa pela Eq. (47), for substituída no termo correspondente da Eq. (45), tem-se que

$$\beta = -\frac{(\mathbf{u}^*)^T \nabla G|_{\mathbf{u}^*}}{\sqrt{(\nabla G|_{\mathbf{u}^*})^T \nabla G|_{\mathbf{u}^*}}} .$$

Analogamente, a solução pode ser expressa em termos de \mathbf{u}^* . Para tanto, basta substituir o

valor de λ na Eq. (44), obtendo

$$\mathbf{u}^* = -\frac{\mathcal{D}(\mathbf{u}^*) \nabla G|_{\mathbf{u}^*}}{\sqrt{(\nabla G|_{\mathbf{u}^*})^T \nabla G|_{\mathbf{u}^*}}}. \quad (48)$$

Como foi mostrado na Seção 5.1.1, a probabilidade de falha pode ser estimada através de $p_f = \Phi(-\beta)$. De acordo com a Eq. (48), o ponto mais provável de falha pode ser representado como $\mathbf{u}^* = -\beta \mathbf{a}$, com as direções de busca dadas por

$$\mathbf{a} = \frac{\nabla G|_{\mathbf{u}}}{\|\nabla G|_{\mathbf{u}}\|}. \quad (49)$$

Neste caso, a direção do gradiente é perpendicular à aproximação linear da função de estado limite no ponto mais provável de falha. Analisando os resultados obtidos, fica claro que a construção dos termos resultantes evidencia a possibilidade de cômputo dos resultados através de uma estratégia de recorrência.

5.1.2.4 Algoritmo do Método de Rackwitz

A metodologia para obtenção do ponto mais provável de falha baseia-se no procedimento iterativo de construções sucessivas de aproximações lineares da função de estado limite. Uma das primeiras técnicas de recorrência com este objetivo foi proposta por Rackwitz (1976). Ela baseia-se na estimativa sucessiva de direções de busca, para cada novo índice de confiabilidade calculado. A fim de oferecer uma visão mais ampla sobre o assunto, os passos principais desta rotina estão listados abaixo (RACKWITZ, 1976 apud HALDAR; MAHADEVAN, 1995):

1. Defina uma estimativa inicial $\mathbf{x}^{(0)}$ para o ponto mais provável de falha e proceda com a transformação da Eq. (37). Através de $\mathbf{u}^{(0)}$ calcule $\beta^{(0)}$ e defina $k = 1$;
2. Calcule o gradiente da função de desempenho no espaço normal padrão, avaliado em $\mathbf{u}^{(k)}$;
3. Calcule a direção de busca $\mathbf{a}^{(k)}$ para a estimativa $\mathbf{u}^{(k)}$, como mostrado na Eq. (49);
4. Obtenha um novo ponto $\mathbf{x}^{(k)}$ em termos de $\beta^{(k)}$, através de $\mathbf{x}^{(k)} = \mu_{\mathbf{x}} - \mathbf{a}^{(k)} \sigma_{\mathbf{x}} \beta^{(k)}$;
5. Substitua a expressão de $\mathbf{x}^{(k)}$, obtida no passo anterior, na função de estado limite original $g(\mathbf{x}) = 0$ e resolva para $\beta^{(k)}$;
6. Utilizando o valor do índice de confiabilidade computado no passo anterior, calcule uma nova estimativa para o ponto mais provável de falha, com $\mathbf{u}^{(k+1)} = -\beta^{(k)} \mathbf{a}^{(k)}$;

7. Faça $k = k + 1$ e repita os passos 2–6 até que o algoritmo convirja.

Pela análise do procedimento apresentado, fica claro que o Método de Rackwitz requer o cálculo da raiz de $g(\mathbf{x}) = 0$ a cada iteração k (observe o quinto passo do algoritmo apresentado acima). A incorporação de dois problemas de otimização, com resultados dependentes entre si, pode significar um aumento expressivo na quantidade de operações totais do método, cuja análise depende de diversos fatores, como a técnica adotada para obtenção de raízes, os parâmetros de controle escolhidos, a necessidade adicional de cálculo de derivadas e o comportamento da função de desempenho envolvida no problema. Em vista desta adversidade, Rackwitz e Fiessler (1978) propuseram um novo método para o cálculo da probabilidade de falha, conhecido como Método de Hasofer-Lind-Rackwitz-Fiessler (HL-RF), por se tratar de uma extensão dos conceitos estabelecidos por Hasofer e Lind (1974).

5.1.2.5 Descrição do Método de Hasofer-Lind-Rackwitz-Fiessler

A ideia fundamental do Método de Hasofer-Lind-Rackwitz-Fiessler baseia-se na obtenção de aproximações sucessivas da função de estado limite, por meio de série de Taylor expandida até o termo de primeira ordem. Na notação vetorial, a relação de recorrência que aproxima a função de estado limite através de uma reta é dada por

$$G(\mathbf{u}^{(k+1)}) = G(\mathbf{u}^{(k)}) + (\nabla G|_{\mathbf{u}^{(k)}})^T (\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}) . \quad (50)$$

Se, para um dado k , o ponto $\mathbf{u}^{(k+1)}$ satisfaz à restrição do problema apresentado na Eq. (43), isso significa que $G(\mathbf{u}^{(k+1)}) = 0$.

Dividindo a Eq. (50) por $\|\nabla G|_{\mathbf{u}^{(k)}}\|$ e introduzindo a direção de busca, definida na Eq. (49), o termo recorrente da aproximação linear assume a forma

$$\frac{G(\mathbf{u}^{(k)})}{\|\nabla G|_{\mathbf{u}^{(k)}}\|} + (\mathbf{a}^{(k)})^T (\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}) = 0 ,$$

tal que, com a expansão dos termos apropriados, a igualdade pode ser reescrita como

$$\frac{G(\mathbf{u}^{(k)})}{\|\nabla G|_{\mathbf{u}^{(k)}}\|} + (\mathbf{a}^{(k)})^T \mathbf{u}^{(k+1)} - (\mathbf{a}^{(k)})^T \mathbf{u}^{(k)} = 0 ,$$

o que equivale a fazer

$$(\mathbf{u}^{(k+1)})^T \mathbf{a}^{(k)} = (\mathbf{u}^{(k)})^T \mathbf{a}^{(k)} - \frac{G(\mathbf{u}^{(k)})}{\|\nabla G|_{\mathbf{u}^{(k)}}\|} . \quad (51)$$

Quando $k \rightarrow \infty$, espera-se que $\mathcal{D}(\mathbf{u}^{(k+1)}) = \beta$ represente o índice de confiabilidade que satisfaz $G(\mathbf{u}^{(k+1)}) = 0$. Assim como no método de Rackwitz, vale a relação de recorrência dada por $\mathbf{u}^{(k+1)} = -\beta^{(k)} \mathbf{a}^{(k)}$ e, conseqüentemente, pode-se tomar a igualdade $\beta^{(k)} = -(\mathbf{u}^{(k+1)})^T \mathbf{a}^{(k)}$. Substituindo o índice de confiabilidade no termo correspondente da Eq. (51), chega-se a

$$\beta^{(k)} = -(\mathbf{u}^{(k)})^T \mathbf{a}^{(k)} + \frac{G(\mathbf{u}^{(k)})}{\|\nabla G|_{\mathbf{u}^{(k)}}\|},$$

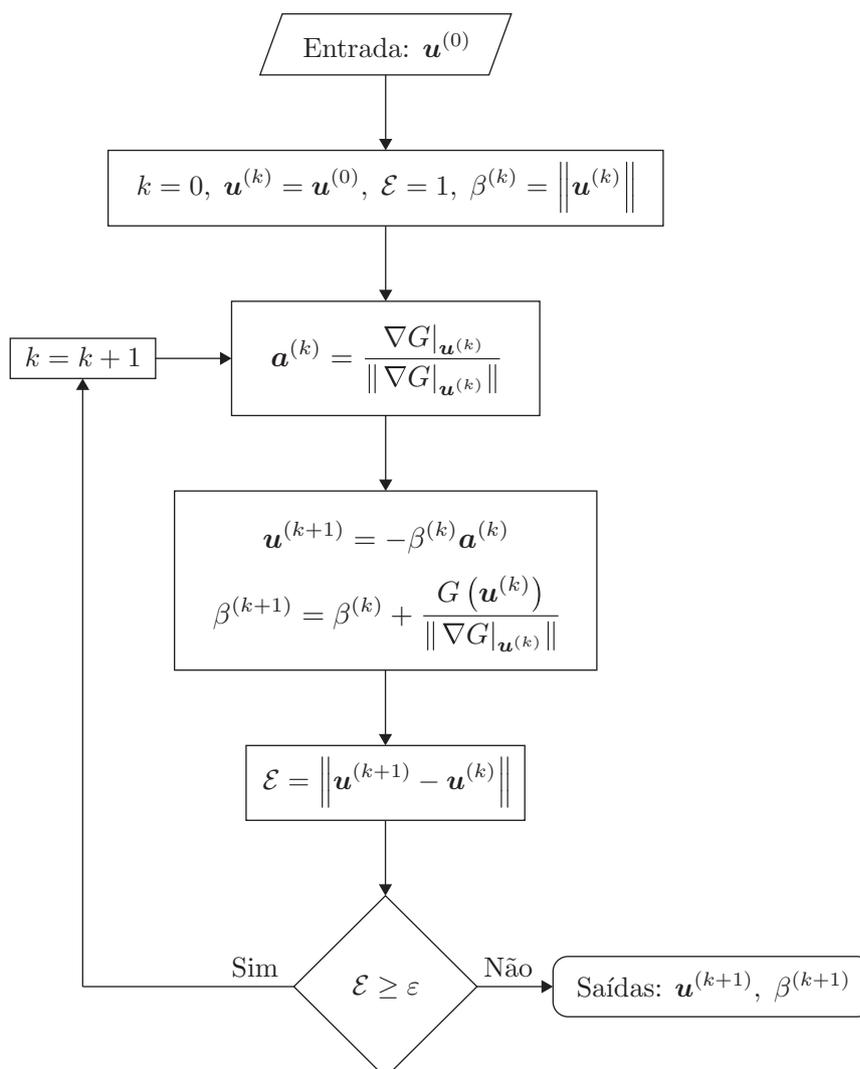
Portanto, através desta igualdade torna-se capaz de calcular o índice de confiabilidade iterativamente. No início do processo, o algoritmo exige a entrada de uma estimativa inicial $\mathbf{u}^{(0)}$. Haja vista que, na primeira etapa do FORM, a média da variável aleatória original \mathbf{X} é mapeada na origem do espaço normal padrão \mathbf{U} , é conveniente escolher $\mathbf{u}^{(0)} = \mathbf{0}$, o que equivale a tomar $\mathbf{x}^{(0)} = \mu_{\mathbf{x}}$. Mais detalhes podem ser vistos em Lemaire et al. (2010).

Além disso, existem três possibilidades distintas, relacionadas aos critérios de parada do algoritmo: a norma euclidiana da diferença entre duas estimativas do ponto mais provável de falha sucessivas, ou de dois vetores gradiente, além da diferença absoluta entre duas estimativas sucessivas para o índice de confiabilidade. O algoritmo termina o processo iterativo quando um desses critérios é menor do que um número ε suficientemente pequeno. A Fig. 20 mostra o fluxograma de uma possível implementação para o Método de Hasofer-Lind-Rackwitz-Fiessler.

Ao fim do processo iterativo, é obtido o índice de confiabilidade—e conseqüentemente o ponto mais provável de falha—que, para a obtenção da probabilidade de falha, deve-se computar $p_f = \Phi(-\beta)$. Analisando o processo iterativo, é possível observar que a operação dominante, no que diz respeito à complexidade do algoritmo, é o cálculo do vetor gradiente. Apesar disso, esta rotina ainda apresenta desempenho superior ao algoritmo proposto por Rackwitz (1976), apresentado anteriormente, uma vez que o processo iterativo é livre do cálculo de raízes.

Dependendo do comportamento da função de desempenho do problema, estimar a probabilidade de falha com o FORM pode resultar em valores não tão precisos. Por razões óbvias, substituir a avaliação de uma função de desempenho altamente não-linear por uma aproximação linear pode ser determinante para uma estimativa excessivamente conservadora ou audaciosa da probabilidade de falha. Nesses casos, considerar a utilização de um método de segunda ordem, como será descrito mais a frente, pode ser uma alternativa apropriada. Outra deficiência do método proposto por Hasofer e Lind (1974) é a necessidade de que todas as variáveis aleatórias devem ter distribuição normal. Desconsiderar esse fato e aplicar o método para variáveis aleatórias com outros tipos de distribuição também pode ocasionar erros significativos. Portanto, antes de apresentar o método de aproximação de segunda ordem, é preciso detalhar algumas alternativas para evitar erros causados por essa particularidade, que serão apresentados a seguir.

Figura 20 - Fluxograma do Método de Hasofer-Lind-Rackwitz-Fiessler.



Fonte: O autor, 2019.

5.1.2.6 Aplicação do FORM em Variáveis Aleatórias com Outros Tipos de Distribuição

Em muitos problemas de engenharia, o modelo adequado pode ser composto por variáveis aleatórias que, por vezes, não apresentam distribuição normal. Em uma situação onde, pelo menos, uma das variáveis aleatórias presente esta característica, torna-se necessário transformá-las em variáveis com distribuição normal equivalente, no caso em que essas informações sejam aplicadas na estimativa da probabilidade de falha de um sistema utilizando o FORM. Em geral, uma transformação exata, do tipo $z = \Phi^{-1}(F_X(x))$, onde Φ^{-1} representa a função inversa da função de distribuição acumulada da distribuição normal padrão e $F_X(x)$ denota a função de distribuição acumulada de uma distribuição arbitrária, pode significar dificuldades computacionais proibitivas.

Em vista desta limitação, Rackwitz e Fiessler (1978) propuseram uma alternativa

para substituir variáveis aleatórias, que sigam uma distribuição qualquer, por uma distribuição normal. Como uma variável aleatória com distribuição normal pode ser descrita, de maneira única, utilizando apenas as suas informações sobre a média e desvio-padrão, a substituição proposta é baseada na condição de que as funções densidade de probabilidade e funções de distribuição acumulada da variável real e daquela com distribuição normal equivalente devem ser iguais, mas apenas em alguns pontos de interesse x_i^* , tal que $i = 1, \dots, k$, onde k representa o número de variáveis aleatórias com distribuição diferente da distribuição normal.

Em um primeiro caso, essa condição corresponde a escolher $\mu_{x_i}^N$ e $\sigma_{x_i}^N$ de tal forma que

$$\Phi\left(\frac{x_i^* - \mu_{x_i}^N}{\sigma_{x_i}^N}\right) = F_{x_i}(x_i^*), \quad (52)$$

onde Φ simboliza a função de distribuição acumulada da distribuição normal padrão, o sobrescrito “N” denota que a média e desvio-padrão correspondentes são relacionados à variável aleatória com distribuição normal, e F_{x_i} representa a função de distribuição acumulada da variável aleatória X . Assim, a Eq. (52) representa a igualdade das funções de distribuição acumulada no ponto de interesse x_i^* . Considerando a propriedade da Eq. (116) e aplicando o operador diferencial em ambos os lados da Eq. (52), tem-se que

$$\frac{1}{\sigma_{x_i}^N} \phi\left(\frac{x_i^* - \mu_{x_i}^N}{\sigma_{x_i}^N}\right) = f_{x_i}(x_i^*), \quad (53)$$

onde ϕ e $f_{x_i}(x_i^*)$ representam a função densidade de probabilidade da distribuição normal padrão equivalente e da distribuição original, respectivamente.

Considerando a transformação mostrada na Eq. (36) e realizando as operações adequadas na Eq. (53), a fim de que seja explicitado o termo $\sigma_{x_i}^N$, o desvio-padrão da distribuição equivalente pode ser calculado como

$$\sigma_{x_i}^N = \frac{\phi\left(\Phi^{-1}\left(F_{x_i}(x_i^*)\right)\right)}{f_{x_i}(x_i^*)}. \quad (54)$$

Por sua vez, se o termo da média for explicitado na Eq. (52), tem-se que

$$\mu_{x_i}^N = x_i^* - \Phi^{-1}\left(F_{x_i}(x_i^*)\right) \sigma_{x_i}^N, \quad (55)$$

a qual depende do valor de $\sigma_{x_i}^N$.

Portanto, tendo determinado a média e o desvio-padrão da distribuição normal equivalente através das Eqs. (55) e (54), respectivamente, pode-se calcular o índice de confiabilidade por meio da metodologia apresentada na Seção 5.1.2, para variáveis aleatórias com

distribuição diferente da distribuição normal. No caso do algoritmo mostrado na Fig. 20, por exemplo, o valor obtido para cada u_i , calculado durante o processo iterativo, é utilizado para recuperar o valor da sua variável original, através de $x_i^* = \mu_{x_i} + \sigma_{x_i} u_i$ e, em seguida, a média e desvio-padrão correspondentes, com distribuição normal, podem ser obtidos pelas Eqs. (55) e (54), respectivamente.

5.2 Método de Confiabilidade de Segunda Ordem

O primeiro estudo envolvendo aproximações de segunda ordem foi realizado por Fiessler et al. (1979), onde alguns resultados da teoria estatística de formas quadráticas de variáveis com distribuição normal foram usados para calcular estimativas melhoradas da probabilidade de falha. Apesar disso, a autoria do Método de Confiabilidade de Segunda Ordem (SORM) pode ser atribuída a Karl Breitung, que estabeleceu os conceitos da aproximação de segunda ordem através da realização de uma série de estudos relacionados (BREITUNG, 1984; BREITUNG; HOHENBICHLER, 1989; BREITUNG, 1989). A seguir, o método SORM é descrito e, em seguida, são apresentadas as etapas principais para implementação do seu algoritmo.

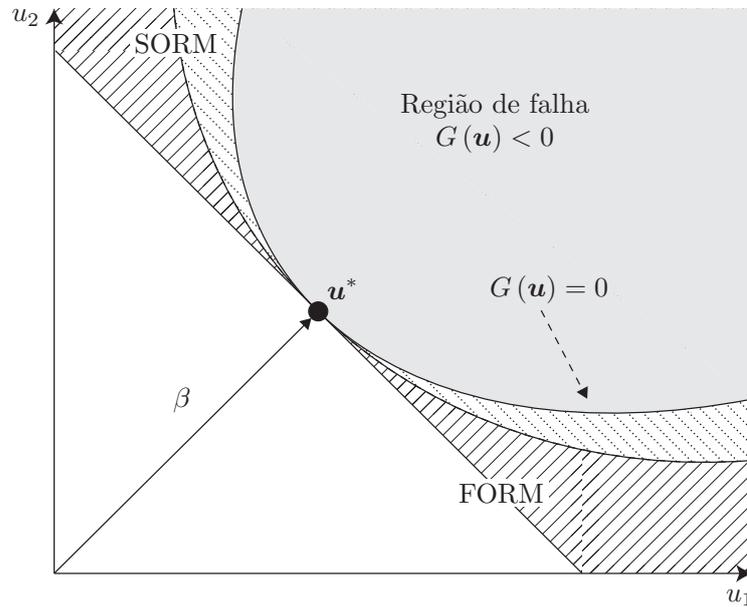
A Fig. 21 ilustra a motivação de se estabelecer uma aproximação quadrática para o cálculo de probabilidades de falha. A área sombreada representa a região de falha, correspondente à integral na Eq. (40). De fato, tanto a função de estado limite original quanto ambas as aproximações apresentam o mesmo ponto de máximo para \mathbf{u}^* . Apesar disso, observe as áreas tracejada e pontilhada, que representam as regiões de falha complementares à área delimitada por $G(\mathbf{u}) = 0$, determinadas pelas aproximações de primeira e segunda ordens, respectivamente. Essas duas áreas podem ser compreendidas, em analogia, como o erro cometido no cálculo da probabilidade de falha, realizado utilizando cada um desses tipos de aproximação. Note que o FORM ignora a curvatura da função de estado limite, o que acarreta um erro de aproximação mais acentuado.

Portanto, a curvatura da função de estado limite, a qual é aproximada ao redor do ponto mais provável de falha, é determinante para a obtenção de soluções mais acuradas. Nesse sentido, considere a aproximação da função de estado limite por meio da série de Taylor de segunda ordem, definida como

$$Q(\mathbf{u}) = G(\mathbf{u}^*) + \sum_{i=1}^n \frac{\partial G(\mathbf{u}^*)}{\partial u_i} (u_i - u_i^*) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 G(\mathbf{u}^*)}{\partial u_i \partial u_j} (u_i - u_i^*)(u_j - u_j^*). \quad (56)$$

Como a curvatura está relacionada às derivadas de segunda ordem, é coerente truncar a série de Taylor após a expansão dos termos quadráticos. Neste caso, \mathbf{u} também sofrerá a transformação de variáveis para o espaço normal padrão, apresentada na Seção 5.1.

Figura 21 - Erros relacionados às aproximações de primeira e segunda ordens.



Fonte: O autor, 2019.

Inicialmente, considere a aproximação de primeira ordem, dada na Eq. (41). Pela definição do problema de obtenção do ponto mais provável de falha, sabe-se que $G(\mathbf{u}^*) = 0$ (ver Eq. (43)). Por sua vez, o índice de confiabilidade pode ser expresso por $\beta = \mathbf{a}^T \mathbf{u}^*$, com \mathbf{a} sendo a direção de busca, no sentido oposto ao que é definido na Eq. (49). Portanto, a linearização da função de desempenho através da aproximação por série de Taylor, em notação vetorial, pode ser representada por

$$L(\mathbf{u}) = \|\nabla G|_{\mathbf{u}^*}\| (\beta - \mathbf{a}^T \mathbf{u}) .$$

Considerando esta representação da aproximação de primeira ordem da função de desempenho ao redor do ponto mais provável de falha, a Eq. (56) pode ser reescrita, também de forma vetorial, como

$$Q(\mathbf{u}) = \|\nabla G|_{\mathbf{u}^*}\| \left(\beta - \mathbf{a}^T \mathbf{u} + \frac{1}{2 \|\nabla G|_{\mathbf{u}^*}\|} (\mathbf{u} - \mathbf{u}^*)^T \nabla^2 G|_{\mathbf{u}^*} (\mathbf{u} - \mathbf{u}^*) \right) , \quad (57)$$

onde $\nabla^2 G|_{\mathbf{u}^*}$ representa a matriz Hessiana.

Para que a Eq. (57) possa ser simplificada, o sistema de coordenadas deve ser rotado, de forma que uma das componentes passe pela origem e pelo ponto mais provável de falha, isto é, uma das componentes deve coincidir com o vetor \mathbf{a} . Habitualmente, a última componente é selecionada para este propósito. Portanto, considere a rotação do sistema de coordenadas, dada por $\mathbf{y} = \mathbf{R}\mathbf{u}$, onde \mathbf{y} são as coordenadas de \mathbf{u} no sistema de coordenadas

resultante da transformação e \mathbf{R} é uma matriz quadrada ortonormal, cuja n -ésima linha é constituída pelo vetor \mathbf{a}^T (DER KIUREGHIAN, 2004).

De acordo com Haldar e Mahadevan (2000), esta matriz de transformação pode ser obtida através do conhecido processo de Gram-Schmidt. Em linhas gerais, primeiro constrói-se uma matriz \mathbf{P} , cujas $n - 1$ primeiras linhas formam a base canônica e a última o vetor \mathbf{a}^T ,

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix}. \quad (58)$$

As linhas de \mathbf{P} são denotadas por $\mathbf{p}_1, \dots, \mathbf{p}_n$. O objetivo é convertê-la em uma matriz \mathbf{R} , com linhas representadas por $\mathbf{r}_1, \dots, \mathbf{r}_n$, ortogonais entre si e em relação ao vetor \mathbf{a} . Com isso, pode-se estabelecer que as n -ésimas linhas de \mathbf{R} e \mathbf{P} devem ser iguais, isto é, $\mathbf{r}_n = \mathbf{p}_n$ e, para $k = n - 1, \dots, 1$, as demais linhas de \mathbf{R} são computadas por

$$\mathbf{r}_k = \mathbf{p}_k - \sum_{j=k+1}^n \frac{\mathbf{r}_j \mathbf{p}_k^T}{\mathbf{r}_j \mathbf{r}_j^T} \mathbf{r}_j.$$

Como \mathbf{R} tem a necessidade de ser ortonormal, deve-se normalizar cada uma das linhas após a construção da matriz.

Do procedimento de rotação, sabe-se que $\mathbf{u} = \mathbf{R}^T \mathbf{y}$, por conta da propriedade que garante que, em relação às matrizes de rotação, $\mathbf{R}^T \mathbf{R} = \mathbf{I}$, sendo que \mathbf{I} representa a matriz identidade. Portanto, substituindo esta relação na Eq. (57) e definindo que $\mathcal{Q}(\mathbf{y}) \equiv Q(\mathbf{R}^T \mathbf{y}) / \|\nabla G|_{\mathbf{R}^T \mathbf{y}^*}\|$, a aproximação quadrática da função de desempenho toma a forma

$$\mathcal{Q}(\mathbf{y}) = \beta - \mathbf{a}^T (\mathbf{R}^T \mathbf{y}) + \frac{1}{2 \|\nabla G|_{\mathbf{R}^T \mathbf{y}^*}\|} (\mathbf{R}^T \mathbf{y} - \mathbf{R}^T \mathbf{y}^*)^T \nabla G|_{\mathbf{y}^*} (\mathbf{R}^T \mathbf{y} - \mathbf{R}^T \mathbf{y}^*), \quad (59)$$

A Eq. (59) ainda pode ser simplificada, tomando-se a matriz

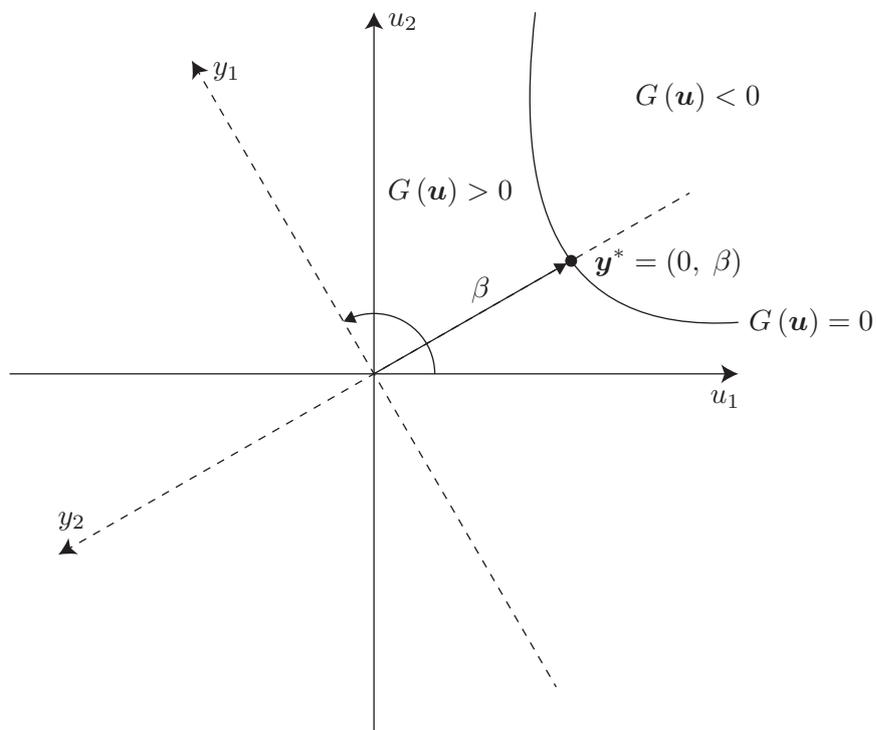
$$\mathbf{A} = \frac{\mathbf{R} \nabla G|_{\mathbf{y}^*} \mathbf{R}^T}{\|\nabla G|_{\mathbf{R}^T \mathbf{y}^*}\|}. \quad (60)$$

e recorrendo ao fato de que $y_n = \mathbf{a}^T (\mathbf{R}^T \mathbf{y})$, uma vez que a n -ésima coordenada de \mathbf{y} corresponde à direção de busca do ponto mais provável de falha. Portanto, a aproximação da função de desempenho da Eq. (59) pode ser reescrita como

$$\mathcal{Q}(\mathbf{y}) = \beta - y_n + \frac{1}{2} (\mathbf{y} - \mathbf{y}^*)^T \mathbf{A} (\mathbf{y} - \mathbf{y}^*). \quad (61)$$

A referida rotação, representada pela transformação $\mathbf{y} = \mathbf{R}\mathbf{u}$, posiciona o ponto mais provável de falha sobre a n -ésima coordenada de \mathbf{y} , fazendo com que $\mathbf{y}^* = (0, \dots, \beta)$, como pode ser visto na Fig. 22. Além disso, ela também garante que a última linha e coluna da matriz \mathbf{A} contenha apenas zeros. Note que, como \mathbf{y}^* é composto apenas por zeros, exceto na sua última coordenada, as operações que envolvem os seus elementos não têm qualquer contribuição para o cálculo da aproximação, uma vez que a matriz \mathbf{A} contém apenas zeros na última linha e coluna. Portanto, o seu uso na Eq. (62) pode ser omitido sem qualquer perda (DER KIUREGHIAN, 2004).

Figura 22 - Procedimento de rotação dos eixos no SORM, para um problema no plano.



Fonte: O autor, 2019.

Considerando estes fatos, pode-se definir uma nova matriz \mathbf{A}_c , composta pelas $n - 1$ primeiras linhas e colunas de \mathbf{A} e um novo vetor $\tilde{\mathbf{y}}$, também formado pelos $n - 1$ primeiros elementos de \mathbf{y} , fazendo com que a Eq. (61) possa ser reescrita como

$$Q(\mathbf{y}) = \beta - y_n + \frac{1}{2} \tilde{\mathbf{y}}^T \mathbf{A}_c \tilde{\mathbf{y}}. \quad (62)$$

Por fim, considere uma nova rotação dos eixos, ao redor de y_n , definida pela transformação $\mathbf{v} = \mathbf{P}\tilde{\mathbf{y}}$, onde \mathbf{P} é uma matriz ortonormal, com dimensão $(n - 1) \times (n - 1)$. Novamente, pode-se afirmar que $\tilde{\mathbf{y}} = \mathbf{P}^T \mathbf{v}$, o que possibilita escrever $\tilde{\mathbf{y}}^T \mathbf{A}_c \tilde{\mathbf{y}} = \mathbf{v}^T \mathbf{P} \mathbf{A}_c \mathbf{P}^T \mathbf{v}$. Com isso, se as colunas de \mathbf{P}^T forem uma base ortonormal de autovalores da matriz \mathbf{A}_c , a expres-

são $\mathbf{P}\mathbf{A}_c\mathbf{P}^T$ diagonaliza \mathbf{A}_c . Segundo Der Kiureghian (2004), a Eq. (62) pode ser reescrita como

$$\mathcal{Q}(\mathbf{P}^T \mathbf{v}, y_n) = \beta - y_n + \frac{1}{2} \sum_{i=1}^{n-1} \kappa_i v_i^2, \quad (63)$$

onde κ_i representa os $n - 1$ autovalores de \mathbf{A}_c . A solução da equação $Q = 0$ representa um hiperparabolóide cujas curvaturas principais são dadas pelos autovalores de \mathbf{A}_c . Estes valores, portanto, são aplicados na solução obtida por Breitung (1984), que determina uma forma simples de se computar a aproximação de segunda ordem, considerando a curvatura da função de desempenho ao redor do ponto mais provável de falha, através de

$$p_f \approx \Phi(-\beta) \prod_{i=1}^{n-1} (1 + \beta \kappa_i)^{-1/2}. \quad (64)$$

Em vista da metodologia apresentada, fica claro que o SORM não é caracterizado por um procedimento iterativo, diferentemente do FORM. Ao invés disto, o SORM pode ser visto como um método corretor, relacionado à estimativa do índice de confiabilidade β , calculado previamente por qualquer outro método de análise de confiabilidade. Obviamente, o custo de se computar a matriz de derivadas de segunda ordem é muito maior do que o cômputo do vetor gradiente no FORM. Por outro lado, é natural que o maior esforço computacional seja revertido em aproximações mais precisas no SORM.

Para que o método SORM possa ser mais bem compreendido, serão sintetizadas abaixo as suas etapas principais, a partir da obtenção de uma estimativa para o ponto mais provável de falha:

1. A partir da matriz \mathbf{P} , dada pela Eq. (58), construir a matriz ortonormal \mathbf{R} , com a sua última linha igual ao vetor de direção de busca \mathbf{a}^T ;
2. Calcular a matriz Hessiana $\nabla^2 G$, cujas derivadas de segunda ordem devem ser avaliadas no ponto mais provável de falha;
3. Definir a matriz \mathbf{A} , como na Eq. (60), e remover sua última linha e coluna, formando a matriz \mathbf{A}_c ;
4. Calcular os autovalores κ_i da matriz \mathbf{A}_c , para $i = 1, \dots, n - 1$;
5. Obter a probabilidade de falha através da Eq. (64), fazendo uso das curvaturas principais κ_i e do índice de confiabilidade β conhecido.

É fato que os métodos FORM e SORM apresentam uma coisa em comum: ambos são abordagens diretas, isto é, através de uma estimativa inicial do ponto mais provável de falha obtém-se um índice de confiabilidade (e, conseqüentemente, a probabilidade de falha). Em

certos casos, pode ser interessante percorrer o caminho inverso, isto é, definir o nível de confiabilidade exigido para um dado projeto e, a partir desta informação, obter o ponto de projeto que satisfaça tais exigências. Mais ainda, pode ser útil definir uma faixa de índices de confiabilidade e obter um conjunto de soluções para os diferentes níveis de satisfação da probabilidade de falha. Formulações deste tipo são conhecidas como análises inversas e a sua construção será apresentada a seguir.

5.3 Análise de Confiabilidade Inversa

Uma das primeiras abordagens presentes na literatura, que considera uma estratégia inversa relacionada ao cálculo da probabilidade de falha, é o trabalho de Der Kiureghian et al. (1994), que propuseram um método para determinação de parâmetros desconhecidos das funções de desempenho, durante o procedimento de obtenção de um dado índice de confiabilidade específico. Apesar disso, a estratégia de análise de confiabilidade inversa só foi estabelecida, de fato, com o estudo de Tu et al. (1999), que desenvolveram a abordagem PMA (*Performance Measure Approach*). A diferença fundamental entre as estratégias RIA e PMA é a forma com a qual o ponto mais provável de falha é calculado.

Considere o caso em que a região de falha, no problema de análise de confiabilidade, é delimitada por $g(\mathbf{X}) \leq \gamma$, com uma função $g(\mathbf{X})$ altamente não-linear. Neste caso, $F_x(\gamma) = P[g(\mathbf{X}) \leq \gamma]$, onde F_x representa a função de distribuição acumulada. Um índice de probabilidade generalizado β_g é uma função não-crescente de γ (observe a Fig. 23), tal que

$$F_x(\gamma) = \Phi(-\beta_g) , \quad (65)$$

onde Φ representa a função de distribuição acumulada da distribuição normal padrão. No caso específico em que a região de falha é determinada por $g(\mathbf{X}) \leq 0$, a Eq. (65) toma a forma

$$F_x(0) \leq \Phi(-\beta_t) , \quad (66)$$

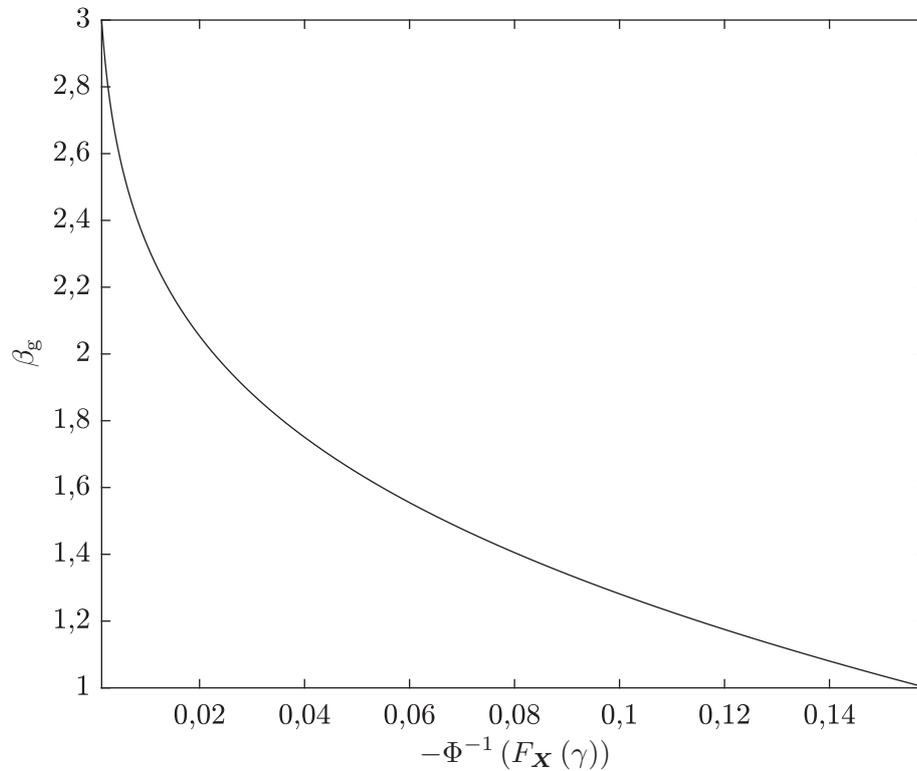
onde β_t é um escalar que representa o índice de confiabilidade alvo.

A expressão da Eq. (66) pode assumir duas formas distintas, levando em conta as devidas transformações inversas, dadas por

$$\beta = -\Phi^{-1}(F_x(0)) \geq \beta_t \quad (67a)$$

Figura 23 - Função não-decrescente do índice de probabilidade generalizado

$$\beta_g(\gamma) = -\Phi^{-1}(F_X(\gamma)).$$



Fonte: O autor, 2019.

e

$$\gamma^* = F_X^{-1}(\Phi(-\beta_t)) \geq 0, \quad (67b)$$

onde β é o índice de confiabilidade e γ^* é chamada de medida de desempenho probabilístico alvo. Note que a Eq. (67a) é a forma compacta de

$$\beta = -\Phi^{-1}\left(\int_{g(\mathbf{x}) \leq 0} f_X(\mathbf{x}) d\mathbf{x}\right),$$

cuja aproximação pode ser calculada utilizando a abordagem de índice de confiabilidade (RIA), pelo problema definido na Eq. (43), após realizadas as devidas transformações da variável aleatória \mathbf{X} para o espaço normal padrão \mathbf{U} .

Por sua vez, a abordagem inversa definida pela Eq. (67b) pode ser compreendida como a determinação quanto a um dado vetor de variáveis de projeto de satisfazer a restrição probabilística para uma determinada probabilidade de falha alvo, uma vez que $F_X^{-1}(\bar{p}_f) \geq 0$, onde $\bar{p}_f = \Phi(-\beta_t) = F_X(0)$ é o limite de probabilidade de falha prescrito. Neste caso, dado

um vetor de variáveis de projeto \mathbf{x} , a restrição probabilística é satisfeita quando $\gamma^* \geq 0$, onde

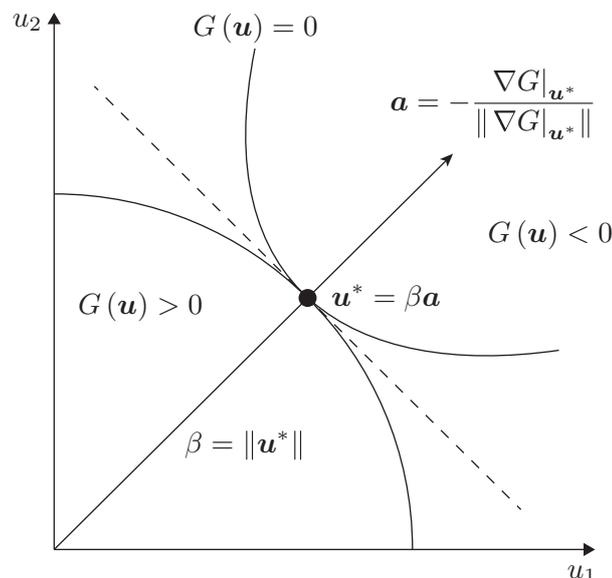
$$\gamma^* = F_x^{-1} \left(\int_{g(\mathbf{x}) \leq \gamma^*} f_x(\mathbf{x}) d\mathbf{x} \right).$$

Portanto, na abordagem de medida de desempenho (PMA), o valor da função de desempenho $G(\mathbf{U})$ deve ser minimizado, enquanto a distância entre a origem do espaço normal padrão e o raio β permanece constante, isto é,

$$\mathbf{u}^* = \begin{cases} \arg \min_{\mathbf{u}} G(\mathbf{u}) \\ \text{Sujeito a } \|\mathbf{u}\| = \beta \end{cases}. \quad (68)$$

Na abordagem inversa, o ponto mais provável de falha \mathbf{u}^* é o ponto tangente do círculo de raio β , centrado na origem do espaço normal padrão \mathbf{U} , em relação à função de estado limite $G(\mathbf{U}) = 0$, como pode ser visto na Fig. 24. Sob um outro ponto de vista, o problema pode ser interpretado como encontrar o valor mínimo de $G(\mathbf{U})$ sobre o perímetro do círculo de raio β , fazendo com que o problema, de fato, resume-se a calcular a direção ótima de busca \mathbf{a} , isto é, o vetor gradiente que aponte para o ponto mais provável de falha que satisfaz o índice de confiabilidade β .

Figura 24 - Esquema de busca do ponto mais provável de falha na abordagem PMA.



Fonte: O autor, 2019.

Melchers e Beck (2018) apontam algumas diferenças relevantes entre as abordagens direta e inversa: em geral, é mais fácil obter soluções de problemas de otimização com funções mais complexas e sujeitas a restrições mais simples do que no cenário oposto. Portanto,

a abordagem PMA tende a ser mais eficiente e convergir mais rápido do que a RIA uma vez que, na primeira, apenas o vetor de direção precisa ser determinado, visto que o índice de confiabilidade é constante, o que é uma vantagem no tratamento da restrição do problema de minimização para obtenção do ponto mais provável de falha. A PMA pode, também, lidar com uma variedade de distribuições de probabilidade sem aumento significativo do custo computacional (YOUN et al., 2003 apud MELCHERS; BECK, 2018).

5.3.1 Visão Geral de Métodos para Análise de Confiabilidade Inversa

Naturalmente, com a consolidação desta formulação, diversos métodos foram propostos, visando a solução do problema definido na Eq. (68). Uma das primeiras propostas, que ganhou destaque por sua eficiência e simplicidade, foi o AMV (*Advanced Mean Value*), sugerida por Wu et al. (1990) e Wu (1994). O método foi originalmente proposto como uma técnica computacionalmente eficiente para gerar a função de distribuição acumulada de uma função resposta, utilizando uma correção simples para compensar os erros introduzidos pela aproximação da série de Taylor utilizada. Mais tarde, o método foi adaptado para ser aplicado na análise de confiabilidade.

Todos os conceitos apresentados na Seção 5.1.2 também são válidos para a análise de confiabilidade inversa. Portanto, inicialmente é preciso realizar a transformação das variáveis aleatórias do problema para o espaço normal padrão. A cada iteração do método AMV, é preciso obter a direção do gradiente normalizado, como na Eq. (49), e uma estimativa do ponto mais provável de falha na iteração $k + 1$ é calculada por meio de

$$\mathbf{u}^{(k+1)} = \beta \frac{\nabla G|_{\mathbf{u}^{(k)}}}{\|\nabla G|_{\mathbf{u}^{(k)}}\|} . \quad (69)$$

Como β_t é constante no decorrer das iterações, assume-se que uma aproximação do vetor de variáveis de projeto que satisfaz o índice de confiabilidade alvo é suficientemente precisa quando $\mathcal{E}^{(k)} = \|\mathbf{u}^{(k)} - \mathbf{u}^{(k-1)}\| < \varepsilon$. Alternativamente, pode-se adotar o mesmo critério de parada, mas considerando o valor de dois vetores gradiente, obtidos consecutivamente.

O método AMV pode apresentar instabilidade e ineficiência para funções de desempenho não-convexas, devido ao uso de informações do gradiente, levando a uma convergência lenta ou podendo até mesmo não convergir, como apontam Youn et al. (2003). Esta dificuldade numérica pode ser resolvida empregando-se um método de busca alternativo, chamado de CMV (*Conjugate Mean Value*), que atualiza a direção de busca através do uso combinado do valor do gradiente de três iterações sucessivas. O método CMV modifica a

direção de busca da Eq. (69), fazendo

$$\mathbf{u}^{(k+1)} = \beta \frac{\mathbf{a}^{(k)} + \mathbf{a}^{(k-1)} + \mathbf{a}^{(k-2)}}{\|\mathbf{a}^{(k)} + \mathbf{a}^{(k-1)} + \mathbf{a}^{(k-2)}\|},$$

onde $\mathbf{a}^{(k)}$ equivale à Eq. (49), com o gradiente sendo avaliado na aproximação $\mathbf{u}^{(k)}$. Pode-se observar que a direção obtida pelo gradiente conjugado é uma soma ponderada de três direções calculadas em sequência. Esta maneira de atualizar a direção de busca melhora a taxa de convergência e a estabilidade em relação àquela do método AMV para funções de desempenho não-convexas.

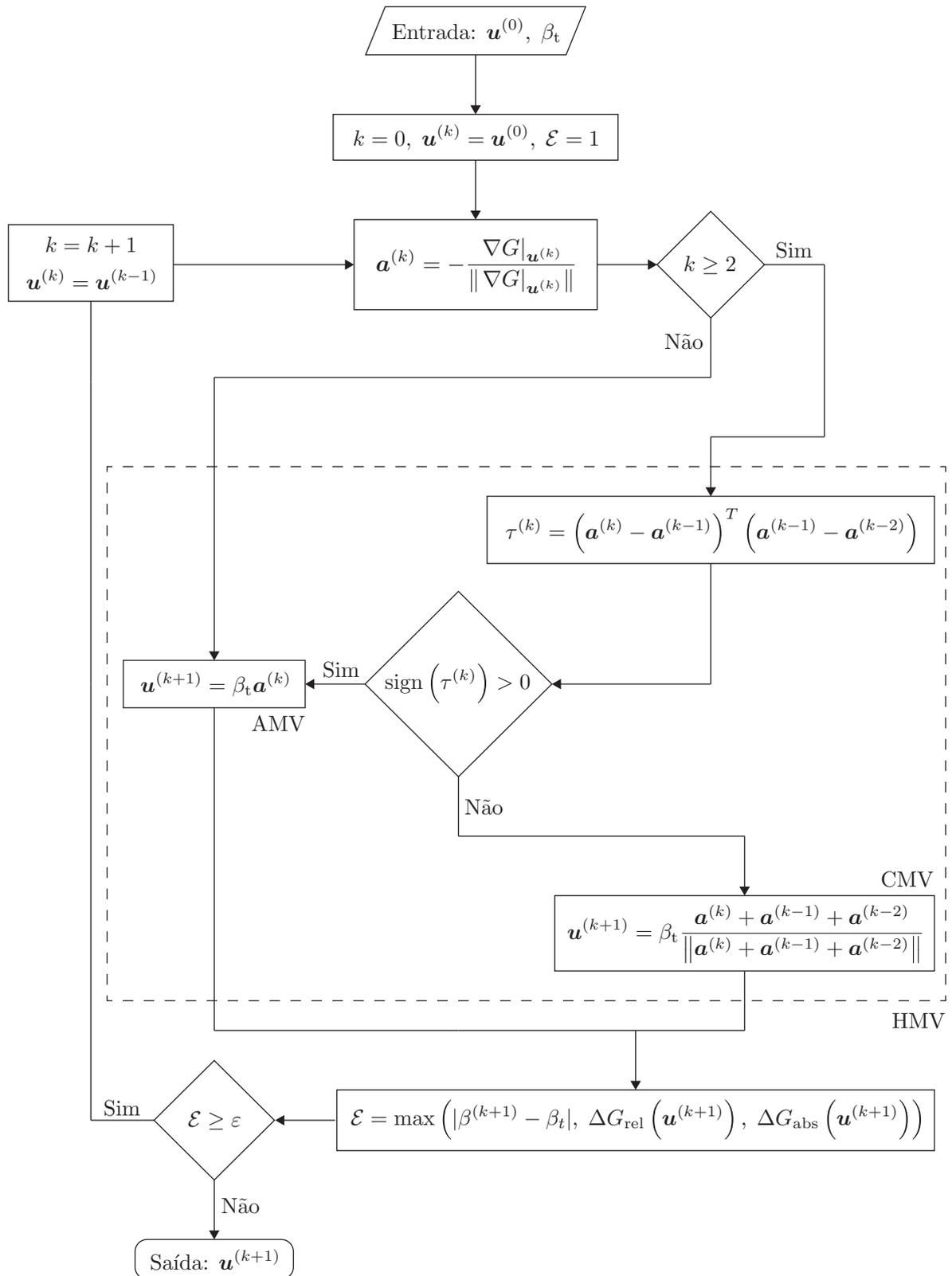
Apesar disso, os mesmos autores mostram que o cálculo da direção de busca pelo gradiente conjugado afeta o desempenho do método CMV nos casos em que a função de desempenho é convexa, fazendo com que a vantagem de se definir uma nova maneira de busca para funções não convexas privasse o funcionamento já estabelecido no método AMV para funções convexas. Visando o balanço entre as melhores características de ambos, Youn et al. (2003) idealizaram o método HMV (Hybrid Mean Value). A proposta é que o método determine, a princípio, o tipo da função de desempenho (côncava ou convexa) na vizinhança de $\mathbf{u}^{(k)}$, com base no valor do gradiente das três iterações mais recentes e, então, selecione um dos dois algoritmos, AMV ou CMV, para o cálculo da nova estimativa do ponto mais provável de falha.

Para a verificação do tipo da função de desempenho no ponto avaliado, considere que $\tau^{(k)} = (\mathbf{a}^{(k)} - \mathbf{a}^{(k-1)})^T (\mathbf{a}^{(k-1)} - \mathbf{a}^{(k-2)})$. Seja, também, sign a função sinal binária. Com isso, foi proposto um critério que define que se $\text{sign}(\tau^{(k)}) > 0$, a função é convexa no ponto avaliado e, portanto, a aproximação $\mathbf{u}^{(k+1)}$ deve ser obtida por meio do método AMV. Caso contrário, a função é côncava no ponto avaliado e um novo ponto é calculado utilizando o método CMV. Obviamente, esta verificação só vale quando $k \geq 2$ e, para as primeiras aproximações, deve-se aplicar o método AMV.

O algoritmo interrompe o processo iterativo quando o critério de parada definido por $\max(|\beta^{(k+1)} - \beta_t|, \Delta G_{\text{rel}}(\mathbf{u}^{(k+1)}), \Delta G_{\text{abs}}(\mathbf{u}^{(k+1)})) \leq \varepsilon$ é satisfeito, levando em conta que $\beta^{(k+1)} = \|\mathbf{u}^{(k+1)}\|$ é o índice de confiabilidade da iteração corrente, β_t representa o índice de confiabilidade alvo (aquele que deseja-se alcançar ao fim das iterações), $\Delta G_{\text{rel}}(\mathbf{u}^{(k+1)})$ e $\Delta G_{\text{abs}}(\mathbf{u}^{(k+1)})$ são as diferenças relativa e absoluta do valor da função de desempenho, avaliadas em duas estimativas sucessivas do ponto mais provável de falha, respectivamente, e ε é um limiar suficientemente pequeno. A Fig. 25 esquematiza o procedimento descrito para o Método HMV, destacando as etapas principais dos Métodos AMV e CMV.

Com o crescimento de aplicações que consideravam a análise de confiabilidade no processo de otimização, a complexidade dos problemas estudados colocaram à prova os métodos existentes. Neste cenário, diversos métodos foram propostos, principalmente visando o aumento da capacidade de convergência e redução do custo computacional. A seguir, são citadas as principais características dos métodos em destaque na literatura. Um dos primei-

Figura 25 - Fluxograma do Método HMV, destacando as principais etapas do AMV e CMV.



ros métodos sugeridos, visando evitar as dificuldades enfrentadas no método HMV, foi o *Enhanced Hybrid Mean Value Method* (YOUN et al., 2005a; YOUN et al., 2005b). O método de análise de probabilidade adaptativa proposto aproxima o ponto mais provável de falha, e então refina adaptativamente este resultado, usando o método de mínimos quadrados.

Mais tarde, foram propostos uma profusão de métodos que levam em conta os conceitos da teoria do caos. A aplicação desses conceitos foi motivada por fenômenos de falha de convergência ocasionados por oscilação periódica e bifurcações em alguns problemas não-lineares. Inicialmente, Yang (2010) propôs o *Chaos Control Method*, baseado no princípio do controle do caos. Fundado nesta ideia, Meng et al. (2015) idealizaram uma técnica empregando o método AMV de forma adaptativa e em conjunto com os conceitos de caos, chamado de *Hybrid Modified Chaos Control Method*. Utilizando ideias relativamente semelhantes, Li et al. (2015) propuseram o *Adaptive Modified Chaos Control Method*.

Seguindo por outra linha, Yi e Zhu (2016) criaram o *Step Length Adjustment Method*, baseado no ajuste do comprimento do passo para controlar a convergência da sequência de aproximações do problema de obtenção do ponto mais provável de falha, visando aumentar a capacidade de convergência do método AMV, mas preservando as suas características, principalmente relacionadas ao baixo custo computacional. Por sua vez, no *Relaxed Mean Value Method*, elaborado por Keshtegar e Lee (2016), um fator de relaxação é adaptativamente implementado ao método AMV. Novamente em relação aos métodos baseados na teoria de caos, Keshtegar et al. (2017) relataram que as propostas existentes até aquele momento eram ineficiente para algumas funções de desempenho não-lineares. Então, os autores sugeriram um método, intitulado *Self-adaptive Modified Chaos Control Method*, com base em um fator de controle dinâmico para melhorar a eficiência do *Hybrid Modified Chaos Control Method*.

O funcionamento do *Hybrid Self-adjusted Mean Value Method* baseia-se em uma modificação do AMV, levando em conta um tamanho de passo auto-adaptativo, isto é, o tamanho do passo para a análise de confiabilidade é ajustado dinamicamente, com base em uma função potência. A adaptação do tamanho de passo funciona em conjunto com uma condição suficiente de descida (KESHTEGAR; HAO, 2017). De forma similar, Keshtegar (2017) elaborou o *Modified Mean Value Method*, em que o AMV é alterado, de tal forma que o ponto mais provável de falha é adaptativamente aproximado utilizando uma direção de busca com base em dois valores anteriores da função de desempenho, além do emprego da ideia de variação do tamanho de passo. Retomando à ideia de aplicação do conceito de caos, Hao et al. (2017) elaboraram o *Enhanced Chaos Control Method*, que verifica e atualiza um fator de controle do algoritmo pelo critério de Wolfe-Powell, a fim de identificar e evitar oscilações periódicas.

Mais tarde, foi criado o *Enriched Self-adjusted Mean Value Method*, onde uma nova metodologia foi aplicada para ajustar a direção de busca, com o intuito de controlar a instabilidade e limitar o tamanho do passo para restrições probabilísticas altamente não-lineares

em problemas de análise de confiabilidade inversa (KESHTEGAR; HAO, 2018b). Com o mesmo propósito, Keshtegar et al. (2018) estabeleceram o *Self-adaptive Conjugate Gradient Method*, visando melhorar a eficiência da busca do ponto mais provável de falha, com base em um fator escalar conjugado adaptativo para problemas côncavos e convexos altamente não-lineares. Com este objetivo, a direção de busca conjugada é calculada de forma adaptativa, usando o valor médio da função de desempenho anterior, com um fator escalar conjugado limitado.

Ainda visando aperfeiçoar o desempenho dos métodos com controle caótico, Keshtegar e Chakraborty (2018) propuseram uma melhoria ao *Chaos Control Method*, com o intuito de favorecer o seu rendimento em problemas de análise de confiabilidade inversa mais complexos, cuja proposta foi identificada como *Dynamical Accelerated Mean Value-based Chaos Control Method*. Nesta técnica, um fator de controle do caos é ajustado dinamicamente e o tamanho do passo adaptável é calculado usando uma condição suficiente de descida. Por outro lado, a utilização dos conceitos de descida rápida (em inglês, *steepest descent*) deram origem aos métodos *Hybrid Descent Mean Value Method* e *Limited Descent-based Mean Value Method*, propostos por Keshtegar e Hao (2018a) e Yaseen e Keshtegar (2018), respectivamente. Ambos se tratam de um aprimoramento do método AMV, modificando a direção de busca do ponto mais provável de falha. A diferença é que, no segundo, o tamanho do passo é limitado, favorecendo a sua convergência.

Mais recentemente, Meng e Keshtegar (2019) elaboraram o *Adaptive-conjugate Method*, onde o vetor de restrições probabilísticas é computado adaptativamente usando o gradiente conjugado, com um fator escalar conjugado dinâmico. Apesar de todas estas propostas, as técnicas de análise de confiabilidade, baseadas na abordagem PMA, permanecem em constante evolução, tendo em vista a diversidade de problemas de otimização, dos quais passou-se a considerar possíveis incertezas. Nesse contexto, o próximo capítulo é destinado à proposta de uma nova técnica de análise de confiabilidade inversa, visando a redução da quantidade de operações necessárias a cada iteração, até a obtenção do ponto mais provável de falha que satisfaça os níveis de confiabilidade estipulados para as restrições probabilísticas. O método é apresentado detalhadamente e, em seguida, são resolvidos alguns problemas de *benchmark* e um problema de engenharia. Os resultados obtidos com a metodologia proposta são comparados com outras técnicas que têm a mesma finalidade.

6 MÉTODO DE TAMANHO DE PASSO DE SEGUNDA ORDEM ADAPTATIVO

6.1 Motivação

Devido à simplicidade e relativa eficiência, o método AMV e suas variantes diretas (CMV e H MV) são comumente usados para obter o ponto de mais provável falha na análise de confiabilidade inversa. No entanto, o AMV sofre dificuldades de convergência quando são necessárias funções de desempenho côncavas. Embora o CMV e o H MV funcionem satisfatoriamente em algumas funções de desempenho convexas e côncavas, quando as funções são fortemente não-lineares, o processo iterativo pode produzir aproximações oscilatórias. Para demonstrar esse comportamento, é aplicado o método H MV (que não será descrito aqui, para mais detalhes consulte Youn et al. (2003)) na função de desempenho

$$g(X_1, X_2) = X_1^4 + 2X_2^4 - 20, \quad (70)$$

com $X_1 \sim \mathcal{N}(10, 5^2)$ e $X_2 \sim \mathcal{N}(12, 5^2)$ mutuamente independentes, quando o índice de confiabilidade alvo é definido como $\beta_t = 2,5$.

A Tabela 2 mostra alguns valores da sequência de aproximações de $\mathbf{u}^{(k)}$ para o ponto mais provável de falha, obtidas na direção $\mathbf{a}^{(k)}$, cujo erro estimado é

$$\mathcal{E}^{(k)} = \|\mathbf{u}^{(k)} - \mathbf{u}^{(k-1)}\| \quad (71)$$

para cada uma das k iterações. Nestes resultados, $G(\mathbf{u})$ representa a função de desempenho, após a transformação das variáveis. O valor ótimo conhecido é $G(\mathbf{u}^*) = 50,3098$, com $\mathbf{u}^* = (-1,5207, -1,9843)$.

A partir de $k = 24$, as aproximações oscilam indefinidamente, com período igual a dois, e ficam presas em um “loop infinito”. Embora o método H MV tente diminuir o comportamento oscilatório, combinando vetores de gradiente de três iterações consecutivas, a convergência não ocorre devido à curvatura excessiva da função de desempenho côncava. Para contornar esse comportamento oscilatório, pode-se usar mais informações sobre a não-linearidade da função desempenho, adotando segundas derivadas para acelerar a convergência. Libotte et al. (2019c) modificaram a direção da busca do método AMV usando a matriz Hessiana da função de desempenho. No entanto, observou-se que a redução do número de iterações é compensada pelo aumento no número de avaliações da função de desempenho (necessárias para calcular as derivadas de segunda ordem). Além disso, embora às vezes a amplitude das oscilações diminua, elas nem sempre são evitadas.

Yi e Zhu (2016) propuseram um método, com base no *steepest descent*, denominado Step Length Adjustment Method (SLA), a fim de evitar esses problemas e preservar a simplicidade do método AMV. A cada iteração, o método calcula um novo ponto em duas etapas,

movendo-se na direção oposta do gradiente, com

$$\mathbf{u}_\lambda^{(k+1)} = \mathbf{u}^{(k)} - \lambda^{(k)} \nabla G|_{\mathbf{u}^{(k)}} \quad (72a)$$

$$\mathbf{u}^{(k+1)} = \beta_t \frac{\mathbf{u}_\lambda^{(k+1)}}{\|\mathbf{u}_\lambda^{(k+1)}\|}, \quad (72b)$$

onde a segunda etapa garante que o índice de confiabilidade seja β_t , e o tamanho do passo $\lambda^{(k)}$ é discutido a seguir.

Tabela 2 - Resultados do processo iterativo do método HMV para a função de desempenho fornecida pela Eq. (70).

k	$a_1^{(k)}$	$a_2^{(k)}$	$u_1^{(k)}$	$u_2^{(k)}$	$G(\mathbf{u}^{(k)})$	$\mathcal{E}^{(k)}$
0	—	—	0,0000	0,0000	51452	—
1	-0,2780	-0,9606	-0,6949	-2,4015	1793,3772	49658,6228
2	-1,0000	0,0000	-2,5000	0,0000	41491,0630	39697,6858
3	-0,5447	-0,8386	-1,3618	-2,0966	94,3073	41396,7557
4	-0,8525	-0,5228	-2,1312	-1,3070	1764,2909	1669,9836
5	-0,4027	-0,9153	-1,0067	-2,2883	588,5586	1175,7323
6	-0,8523	-0,5230	-2,1309	-1,3075	1761,1695	1172,6109
7	-0,4461	-0,8950	-1,1152	-2,2375	363,8911	1397,2785
8	-0,8916	-0,4528	-2,2290	-1,1320	3212,8618	2848,9707
⋮	⋮	⋮	⋮	⋮	⋮	⋮
24	-0,8901	-0,4558	-2,2252	-1,1394	3137,7987	2761,0616
25	-0,4431	-0,8965	-1,1079	-2,2411	376,7371	2761,0616
26	-0,8901	-0,4558	-2,2252	-1,1394	3137,7987	2761,0616
27	-0,4431	-0,8965	-1,1079	-2,2411	376,7371	2761,0616
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Fonte: O autor, 2019.

Na concepção original do *steepest descent*, o tamanho de passo ótimo é calculado a cada iteração, usando um procedimento de busca linear. No entanto, o SLA adota uma estratégia diferente, em favor do desempenho: $\lambda^{(k)}$ é mantido constante ou diminui de acordo com o seguinte critério. Define-se os parâmetros $\lambda_{\max} > \lambda_{\min} > 0$, limitando os valores de $\lambda^{(k)}$, e $c_{\text{SLA}} > 1$, um fator de ajuste. Inicialmente, $\lambda^{(0)} = \lambda_{\max}$ e de acordo com cada nova estimativa, se $\mathcal{E}^{(k)} \geq \mathcal{E}^{(k-1)}$, onde $\mathcal{E}^{(k)}$ é dado pela Eq. (71), o tamanho do passo é reduzido através de $\lambda^{(k+1)} = \lambda^{(k)} / c_{\text{SLA}}$, sempre que $\lambda^{(k+1)} > \lambda_{\min}$. Caso contrário, o tamanho do passo não muda, $\lambda^{(k+1)} = \lambda^{(k)}$. Pode acontecer que $\lambda^{(k)}$ seja constante durante todo o processo iterativo.

Nos problemas de *benchmark* considerados por Yi e Zhu (2016) usando o SLA, a

técnica de ajuste do tamanho do passo sempre foi mais eficiente em comparação com outras abordagens, principalmente no número necessário de avaliações da função de desempenho (critério adotado pelos autores para análise de custo computacional). No entanto, supor que a sequência de erros esteja diminuindo estritamente, o que manteria o tamanho do passo constante, pode ser excessivamente otimista. Portanto, em geral, o tamanho do passo seria reduzido e pode se tornar muito pequeno, diminuindo a capacidade de convergência. Nesse sentido, existem dois aspectos a serem considerados. Primeiro, o critério de ajuste do tamanho do passo não leva significativamente em conta o erro de uma determinada iteração. Segundo, ele também não usa o valor da função de desempenho. Ao aumentar o tamanho do passo, pode-se diminuir o erro. Particularmente no caso de funções altamente não-lineares, avaliar o tamanho do passo a cada iteração, utilizando uma boa estratégia, pode acelerar a convergência do método.

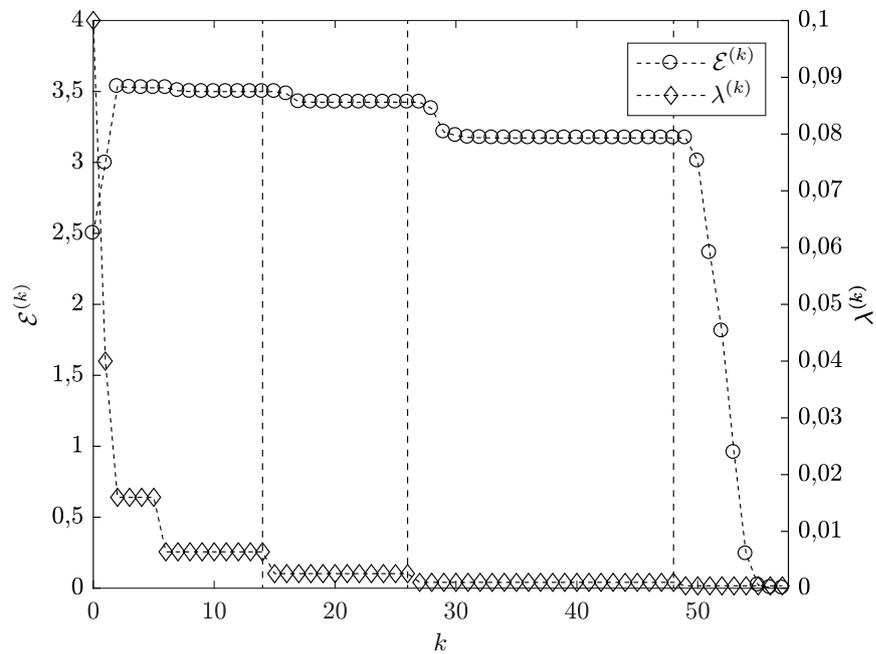
Considere novamente o problema definido pela Eq. (68), para g dado pela Eq. (70). Seguindo as recomendações de Yi e Zhu (2016), o SLA é executado com $\lambda_{\max} = 0,1$, $\lambda_{\min} = 0$, e o coeficiente de ajuste é tomado como $c_{\text{SLA}} = 2,5$. Além disso, o critério de parada é definido como $\mathcal{E}^{(k)} \leq 10^{-3}$, e a estimativa inicial $\mathbf{u}^{(0)} = (0, 0)$. A Fig. 26a mostra a relação entre o tamanho de passo e o erro em cada iteração k , até a convergência. É possível subdividir a sequência em quatro grupos, delimitados por linhas verticais na Fig. 26a. Para os três primeiros grupos, portanto, para a maioria das iterações, existe uma relação direta entre o tamanho do passo e o erro. Nesse caso, fica claro que manter o tamanho do passo constante por vários passos iterativos não é uma prática benéfica, pois o erro permanece praticamente inalterado, evidenciando a dificuldade de explorar o domínio e obter o ponto mínimo. Supõe-se que este fenômeno ocorre porque o critério de ajuste do tamanho do passo não permite que o seu valor aumente, independentemente dos valores da função de desempenho na direção da busca.

Mesmo que a sequência $\mathbf{u}^{(k)}$ apresente um comportamento agudo em zigue-zague (veja a Fig. 26b), usual nos métodos do tipo *steepest descent*, há apenas uma pequena variação de erro entre os valores pertencentes a cada subgrupo analisado (veja a Fig. 26a). Não se pode concluir que o comportamento em zigue-zague do método SLA se deve à manutenção de um tamanho de passo constante, sempre que a condição $\mathcal{E}^{(k)} < \mathcal{E}^{(k-1)}$ for atendida.

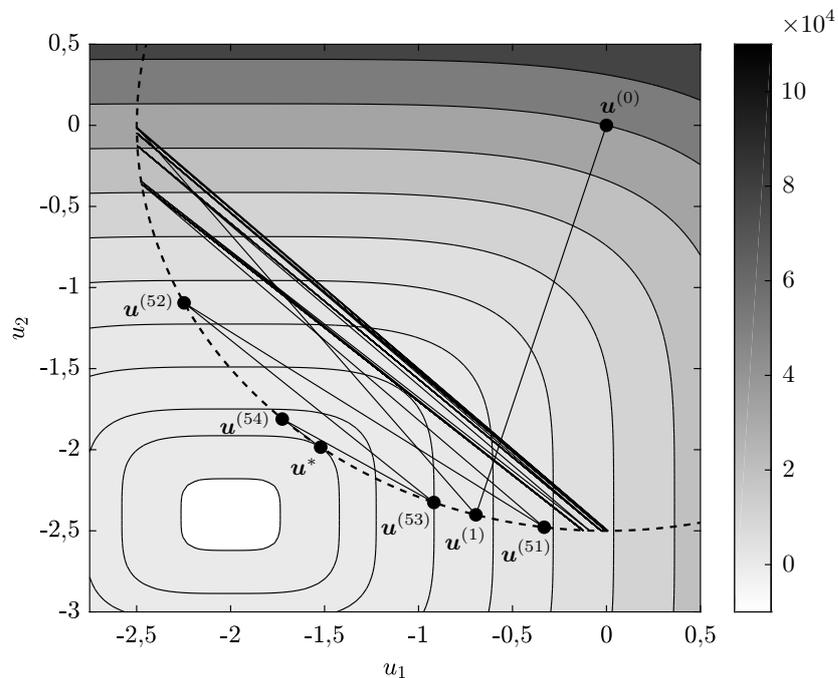
A estratégia adotada para o ajuste do tamanho do passo no método SLA é fortemente motivada pela redução do custo computacional. Em geral, calcular um tamanho de passo ideal em cada iteração representa um custo computacional adicional. Apesar disso, o cálculo de um tamanho de passo aprimorado em cada iteração (permitindo que seja maior ou menor que o da iteração anterior) pode levar a uma convergência mais acelerada para o ponto mais provável de falha.

O objetivo é propor um método para análise de confiabilidade inversa, capaz de aproximar o ponto mais provável de falha usando um tamanho de passo adaptável de segunda ordem, seguindo a estratégia apresentada por Andrei (2004) em outro contexto, conforme

Figura 26 - Comportamento do método SLA no problema fornecido pela Eq. (70). Nesta simulação, $\beta_t = 2,5$.



(a) Relação entre erro $\mathcal{E}^{(k)}$ e tamanho de passo $\lambda^{(k)}$ para cada iteração k . As leituras de $\mathcal{E}^{(k)}$ estão à esquerda e as de $\lambda^{(k)}$ estão à direita.



(b) Comportamento em zigue-zague da sequência $\mathbf{u}^{(k)}$ (pouco mais de um quarto da hipersfera (um círculo) é representado na figura, com a linha tracejada).

Fonte: O autor, 2019.

discutido na próxima seção. Os resultados obtidos para problemas de *benchmark* e de engenharia são comparados com outras técnicas recentes.

6.2 Cálculo do Tamanho de Passo Ótimo

Considere o cálculo do minimizador $\mathbf{u}^* \in \mathbb{R}^n$ de $G(\mathbf{u})$, onde $G : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função de desempenho contínua e duplamente diferenciável, sem restrições, construindo uma sequência minimizadora $\mathbf{u}^{(k)}$, por uma modificação do método *steepest descent*, em que as direções de busca são dadas pelo gradiente na aproximação atual, $\mathbf{d}^{(k)} = (\nabla G|_{\mathbf{u}^{(k)}})^\top$, e a relação de recorrência

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \tau^{(k)} \mathbf{d}^{(k)} \quad (73)$$

define a sequência que, espera-se, aproxima \mathbf{u}^* quando $k \rightarrow \infty$ (veja a Fig. 27a). Nesta relação, $\tau^{(k)}$ representa o tamanho do passo na k -ésima iteração e é definido por um problema de minimização restrito (em vez do problema irrestrito que o *steepest descent* normalmente adota)

$$\tau^{(k)} = \arg \min_{0 < t \leq \bar{t}^{(k)}} G(\mathbf{u}^{(k)} - t \mathbf{d}^{(k)}), \quad (74)$$

com $\bar{t}^{(k)}$ definido como segue.

Inicialmente, $\bar{t}^{(0)} = 1$. Dada uma estimativa inicial $\mathbf{u}^{(0)}$, a Eq. (73) permite calcular $\mathbf{u}^{(1)}$. Utiliza-se aproximações de segunda ordem de $\mathcal{G}_k(t) = G(\mathbf{u}^{(k)} - t \mathbf{d}^{(k)})$, a restrição de G para a busca linear, a fim de definir $\bar{t}^{(k)}$. Considerando a fórmula de Taylor em torno de $t = 0$, tem-se que

$$\mathcal{G}_k(t) = G(\mathbf{u}^{(k)}) - t (\mathbf{d}^{(k)})^\top \mathbf{d}^{(k)} + \frac{1}{2} t^2 (\mathbf{d}^{(k)})^\top (\nabla^2 G|_{\mathbf{z}_t}) \mathbf{d}^{(k)},$$

com $\mathbf{z}_t \in [\mathbf{u}^{(k)}, \mathbf{u}^{(k)} - t \mathbf{d}^{(k)}]$. Estas aproximações são utilizadas de duas maneiras, construindo um polinômio de segundo grau aproximando

$$\Gamma(t) = G(\mathbf{u}^{(k)}) - t (\mathbf{d}^{(k)})^\top \mathbf{d}^{(k)} + \frac{1}{2} t^2 (\mathbf{d}^{(k)})^\top (\nabla^2 G|_{\mathbf{u}^{(k)}}) \mathbf{d}^{(k)}, \quad (75)$$

e uma aproximação de $\mathcal{G}_{k-1}(\tau^{(k-1)}) = G(\mathbf{u}^{(k)})$, dada por

$$G(\mathbf{u}^{(k)}) \approx G(\mathbf{u}^{(k-1)}) - \tau^{(k-1)} (\mathbf{d}^{(k-1)})^\top \mathbf{d}^{(k-1)} + \frac{1}{2} (\tau^{(k-1)})^2 (\mathbf{d}^{(k-1)})^\top (\nabla^2 G|_{\mathbf{u}^{(k)}}) \mathbf{d}^{(k-1)}, \quad (76)$$

onde, em cada caso, substitui-se \mathbf{z}_t por uma extremidade diferente.

Uma possibilidade de estabelecer o tamanho do passo seria encontrar o minimizador

de $\Gamma(t)$, definido pela Eq. (75), que é dado por

$$\tilde{t}^{(k)} = \frac{\left(\mathbf{d}^{(k)}\right)^T \mathbf{d}^{(k)}}{\left(\mathbf{d}^{(k)}\right)^T \left(\nabla^2 G|_{\mathbf{u}^{(k)}}\right) \mathbf{d}^{(k)}}. \quad (77)$$

Se G é convexa, o intervalo $[0, \tilde{t}^{(k)}]$ seria uma escolha razoável para o algoritmo de busca linear encontrar um mínimo local de G ao redor de $\mathbf{u}^{(k)}$, restrito pela reta $\mathbf{u}^{(k)} - t\mathbf{d}^{(k)}$, como é explicado brevemente na Fig. 27b. No entanto, o cálculo de $\tilde{t}^{(k)}$ envolve derivadas de segunda ordem e, em vez disso, uma aproximação mais rápida é usada.

Note que $\tilde{t}^{(k)}$ é o inverso do quociente de Rayleigh, avaliado na direção da busca $\mathbf{d}^{(k)}$, que é, portanto, um valor entre o inverso do menor e maior autovalores da matriz Hessiana simétrica $\nabla^2 G|_{\mathbf{u}^{(k)}}$. Em particular, é homogêneo de grau zero em $\mathbf{d}^{(k)}$. Seguindo Andrei (2004), para evitar o cálculo da Hessiana de G , aproxima-se $\tilde{t}^{(k)}$ avaliando o inverso do quociente de Rayleigh de $\nabla^2 G|_{\mathbf{u}^{(k)}}$ na direção de busca anterior (mantendo a homogeneidade) e, além disso, usando a Eq. (76) para aproximar $\nabla^2 G|_{\mathbf{u}^{(k)}}$ como

$$\tilde{t}^{(k)} \approx \frac{\left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}}{\left(\mathbf{d}^{(k-1)}\right)^T \left(\nabla^2 G|_{\mathbf{u}^{(k)}}\right) \mathbf{d}^{(k-1)}} \approx \frac{\left(\tau^{(k-1)}\right)^2 \left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}}{2\left(G\left(\mathbf{u}^{(k)}\right) - G\left(\mathbf{u}^{(k-1)}\right) + \tau^{(k-1)}\left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}\right)}.$$

Finalmente, define-se $\bar{t}^{(k)}$ como um meta-modelo para $\tilde{t}^{(k)}$ usando o lado direito da equação anterior,

$$\bar{t}^{(k)} = \frac{\left(\tau^{(k-1)}\right)^2 \left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}}{2\left(G\left(\mathbf{u}^{(k)}\right) - G\left(\mathbf{u}^{(k-1)}\right) + \tau^{(k-1)}\left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}\right)}, \quad (78)$$

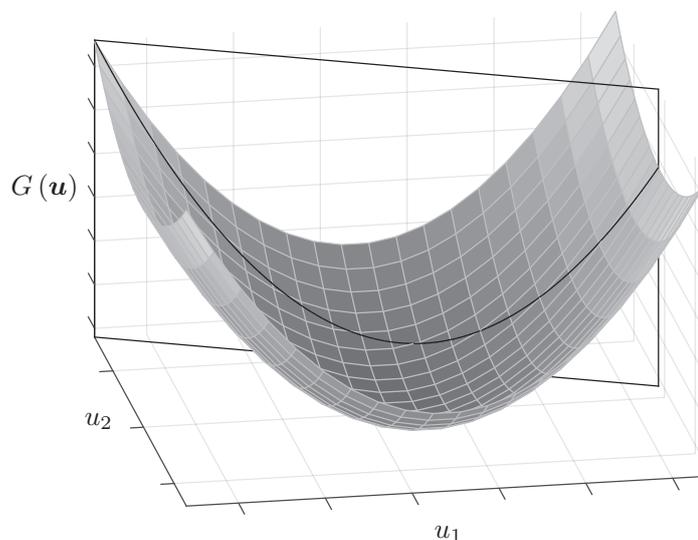
para ser usado na definição de $\tau^{(k)}$ na Eq. (74).

Quando G é convexa, o gráfico de G está acima da linha tangente (veja a Fig. 27b). Em particular,

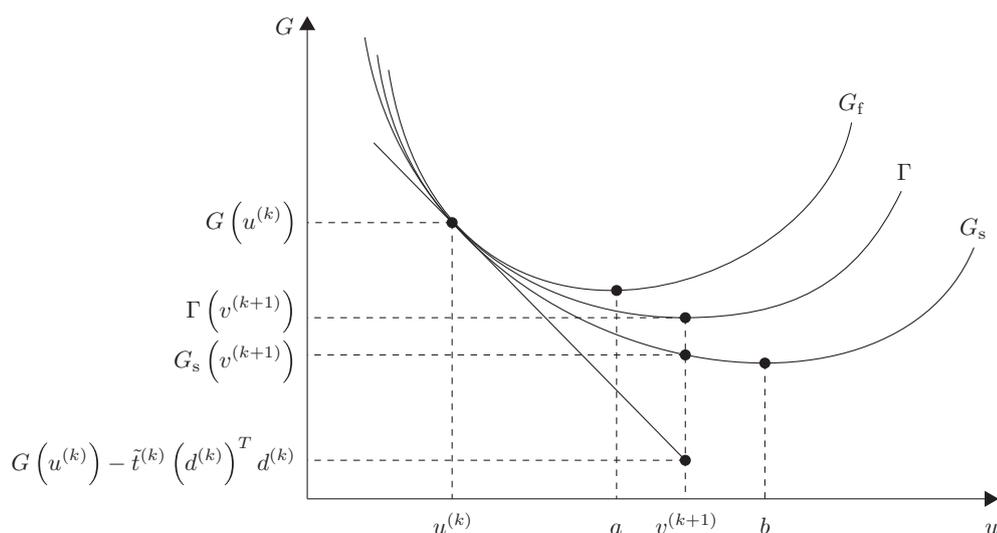
$$G\left(\mathbf{u}^{(k)}\right) \geq G\left(\mathbf{u}^{(k-1)}\right) - \tau^{(k-1)}\left(\mathbf{d}^{(k-1)}\right)^T \mathbf{d}^{(k-1)}.$$

Neste caso, o denominador da Eq. (78) é positivo e, portanto, $\bar{t}^{(k)}$ é positivo. Por outro lado, se G é não-convexa, o valor de $\bar{t}^{(k)}$ pode ser negativo, enquanto $\tilde{t}^{(k)}$ na Eq. (77) é o maximizador de $\Gamma(t)$, não seu minimizador (veja a Fig. 28). No entanto, como $\mathcal{G}'_k(0) = -\left(\mathbf{d}^{(k)}\right)^T \mathbf{d}^{(k)} < 0$ (a menos que seja atingido um ponto crítico, caso em que seria igual a zero), pelo menos para um $\tau^{(k)}$ positivo e suficientemente pequeno na Eq. (73), $\mathcal{G}_k(t)$ decresce.

Figura 27 - Obtenção do minimizador de uma função convexa.



(a) Buscando o minimizador ao longo de uma direção de busca $\mathbf{d}^{(k)}$.

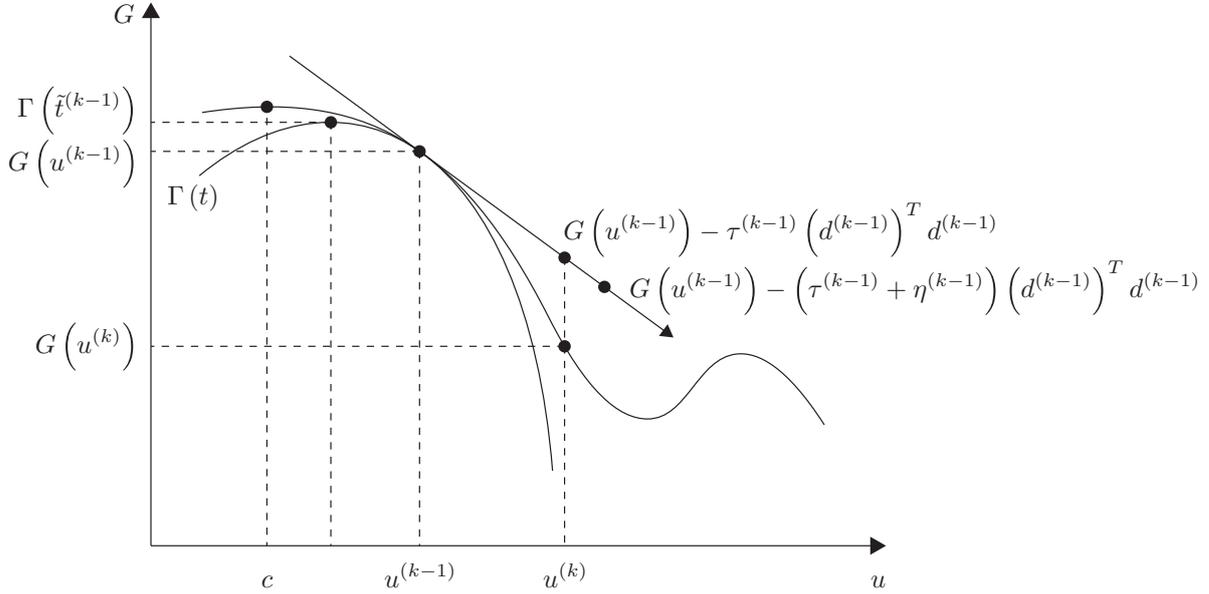


(b) Quando G é convexa, e procura-se o ponto mínimo ao longo de uma direção de busca, tipicamente próximo a $u^{(k)}$ a função G altera o seu comportamento rapidamente, G_f , comparado com sua aproximação polinomial de segunda ordem, Γ , ou lentamente, G_s . No primeiro caso, o minimizador a está no intervalo de busca $[u^{(k)}, v^{(k+1)}]$, com $v^{(k+1)} = u^{(k)} - \tilde{\mathbf{t}}^{(k)} \mathbf{d}^{(k)}$. No segundo caso, o minimizador local b possivelmente não está no intervalo $[u^{(k)}, v^{(k+1)}]$, e não pode ser alcançado em um passo. No entanto, por conta da convexidade, $G(u^{(k)}) > G(v^{(k+1)})$, então um ponto com menor valor do que $G(u^{(k)})$, possivelmente $v^{(k+1)}$, será obtido, e o algoritmo avança em direção ao mínimo.

Fonte: O autor, 2019.

Um possível limite superior ($\tilde{t}^{(k)}$) para o intervalo de busca, no caso de uma função não-convexa, pode ser obtido considerando uma modificação da função G ao longo da di-

Figura 28 - Extensão do limite superior ($\tilde{t}^{(k)}$) para o intervalo de busca, quando a aproximação polinomial de segunda ordem é não-convexa. Neste caso, $\mathbf{v}^{(k)} = \mathbf{u}^{(k-1)} - \tilde{t}^{(k-1)} \mathbf{d}^{(k-1)}$ e o minimizador de Γ é representado por c , fora do intervalo de busca.



Fonte: O autor, 2019.

reção da busca, mantendo algumas de suas características, mas transformando-a em uma função convexa. Seja $\Lambda(t)$ a modificação proposta. Considera-se um polinômio de segundo grau, $\Lambda(t) = p_1 + p_2 t + p_3 t^2$, tal que $\Lambda(0) = G(\mathbf{u}^{(k-1)})$ e $\Lambda'(0) = -(\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}$. Portanto,

$$\Lambda(t) = G(\mathbf{u}^{(k-1)}) - (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)} t + p_3 t^2.$$

A fim de definir p_3 , considera-se a condição de que $\Lambda(\tau^{(k-1)} + \eta^{(k-1)}) = G(\mathbf{u}^{(k)})$, portanto

$$p_3 = \frac{G(\mathbf{u}^{(k)}) - G(\mathbf{u}^{(k-1)}) + (\tau^{(k-1)} + \eta^{(k-1)}) (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}}{(\tau^{(k-1)} + \eta^{(k-1)})^2},$$

onde $\eta^{(k-1)}$ é uma extensão do intervalo de busca, necessária para Λ ser convexa (o que é garantido por $p_3 \geq 0$). Escolhendo $\eta^{(k-1)}$, tal que $p_3 = 0$, mas adicionando ainda uma pequena quantidade representada por δ_η , tem-se que

$$\eta^{(k-1)} = \frac{G(\mathbf{u}^{(k-1)}) - G(\mathbf{u}^{(k)}) - \tau^{(k-1)} (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}}{(\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}} + \delta_\eta. \quad (79)$$

Assim como quando G é convexa, o tamanho do passo pode ser estabelecido calcu-

lando o minimizador de $\Lambda(t)$, que é dado por

$$\bar{t}^{(k)} = \frac{(\tau^{(k-1)} + \eta^{(k-1)})^2 (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}}{2 \left(G(\mathbf{u}^{(k)}) - G(\mathbf{u}^{(k-1)}) + (\tau^{(k-1)} + \eta^{(k-1)}) (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)} \right)}. \quad (80)$$

com $\eta^{(k-1)}$ dado pela Eq. (79). Esta definição é válida quando $\bar{t}^{(k)}$ na Eq. (78) é negativo. Em particular,

$$G(\mathbf{u}^{(k)}) \geq G(\mathbf{u}^{(k-1)}) - (\tau^{(k-1)} + \eta^{(k-1)}) (\mathbf{d}^{(k-1)})^T \mathbf{d}^{(k-1)}$$

e a extensão do intervalo de busca força a convexidade da aproximação dada por $\Lambda(t)$.

6.3 Algoritmo do Método de Tamanho de Passo de Segunda Ordem Adaptativo

Nesta seção, o Método de Tamanho de Passo de Segunda Ordem Adaptativo (ASOSL, do inglês *Adaptive Second Order Step Length*) é proposto para análise de confiabilidade inversa. Ele é baseado na abordagem PMA, definida pela Eq. (68). Fundamentalmente, o algoritmo é uma modificação do *steepest descent*. A proposta é que o tamanho do passo possa ser aumentado ou diminuído iterativamente, de acordo com o problema que está sendo resolvido. De fato, o método pode ser escrito como na Eq. (72), com o tamanho do passo, indicado aqui como $\tau^{(k)}$, sendo calculado a cada iteração, por um procedimento de busca linear. Usa-se a estratégia de cálculo do tamanho de passo, proposta por Andrei (2004) no contexto da otimização geral sem restrições, para acelerar a convergência da busca linear.

Para descrever o método ASOSL, seja $G : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de desempenho contínua e duplamente diferenciável após a transformação das variáveis aleatórias envolvidas no problema para o espaço normal padrão. Considere, a princípio, uma estimativa inicial do ponto mais provável de falha $\mathbf{u}^{(0)} \in \mathbb{R}^n$, calcule $G(\mathbf{u}^{(0)})$ e o vetor gradiente $\mathbf{d}^{(0)} = \nabla G|_{\mathbf{u}^{(0)}}$. Então, o cálculo da próxima estimativa $\mathbf{u}^{(1)}$ para o ponto mais provável de falha depende do tamanho de passo $\tau^{(0)}$, dado pela Eq. (74) com $\bar{t}^{(0)} = 1$, que pode ser obtido por um procedimento de busca linear. Neste caso, é utilizada a técnica de *backtracking*, descrita no Apêndice A.

Usando esses valores, torna-se possível calcular $\mathbf{u}_\tau^{(1)} = \mathbf{u}^{(0)} - \tau^{(0)} \mathbf{d}^{(0)}$. No entanto, dado um índice de confiabilidade alvo, denotado como β_t , deve ser assegurado que $\|\mathbf{u}^{(k)}\| = \beta_t$ para cada iteração $k > 0$. Portanto, a nova aproximação para o ponto mais provável de falha é calculada normalizando-se $\mathbf{u}_\tau^{(1)}$, tal que $\mathbf{u}^{(1)} = \beta_t \left(\mathbf{u}_\tau^{(1)} / \|\mathbf{u}_\tau^{(1)}\| \right)$. Isso faz com que $\mathbf{u}^{(1)}$ pertença à hipersfera do raio β_t . Imediatamente, é possível calcular $G(\mathbf{u}^{(1)})$ e $\mathbf{d}^{(1)} = \nabla G|_{\mathbf{u}^{(1)}}$. Em seguida, calcula-se o limite superior para o intervalo da busca linear do tamanho de passo, através dos conceitos mostrados na Seção 6.2, que usa dois pontos

consecutivos. Escolha δ_η e avance com o processo iterativo descrito a seguir. Para $k \geq 1$, compute $\bar{t}^{(k)}$ usando a Eq. (78). Se o valor calculado for negativo, realize a sua extensão, que é feita calculando uma nova estimativa para $\bar{t}^{(k)}$, usando a Eq. (80), onde $\eta^{(k-1)}$ é obtido pela Eq. (79). Desta forma, é garantido que $\bar{t}^{(k)} > 0$ e este valor pode ser usado para calcular o tamanho de passo ótimo, $\tau^{(k)}$, obtido resolvendo-se o problema de busca linear fornecido pela Eq. (74), utilizando a técnica de *backtracking*.

O próximo passo do processo iterativo é o cálculo de

$$\mathbf{u}_\tau^{(k+1)} = \mathbf{u}^{(k)} - \tau^{(k)} \mathbf{d}^{(k)},$$

e então a nova estimativa do ponto mais provável de falha é obtida pela normalização $\mathbf{u}_\tau^{(k+1)}$, resultando em

$$\mathbf{u}^{(k+1)} = \beta_\tau \frac{\mathbf{u}_\tau^{(k+1)}}{\|\mathbf{u}_\tau^{(k+1)}\|}$$

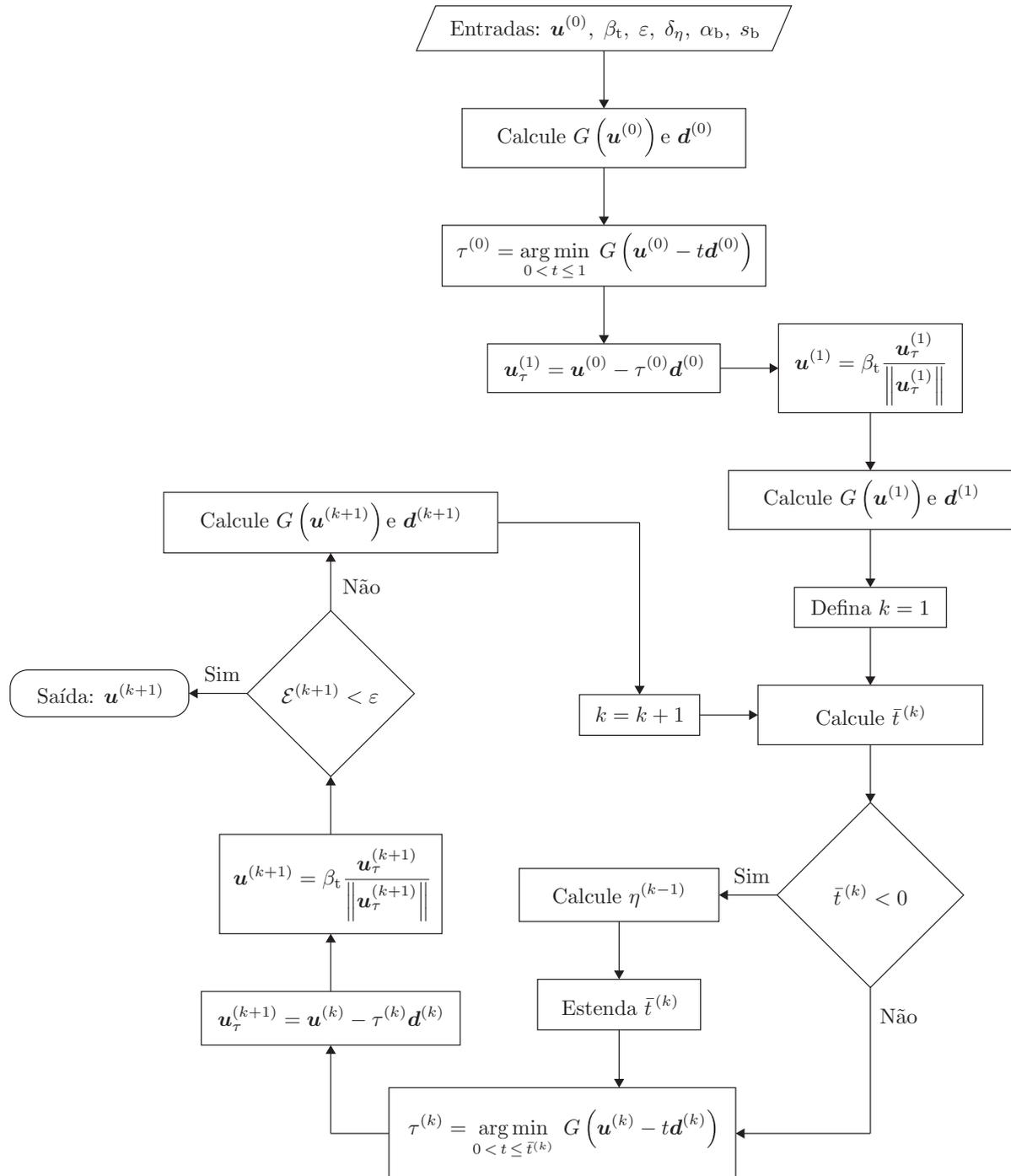
para $k \geq 1$. O resultado obtido é utilizado para avaliar $G(\mathbf{u}^{(k+1)})$ e calcular $\mathbf{d}^{(k+1)}$.

Para a maioria dos problemas, o cálculo de derivadas usando uma estratégia de aproximação de diferenças finitas de dois pontos é uma alternativa viável para o vetor gradiente. Ele não afeta significativamente a determinação dos resultados corretos e a precisão das derivadas, além de exigir apenas uma única avaliação da função de desempenho, além do cálculo de $G(\mathbf{u}^{(k+1)})$. Além disso, é alcançado com baixo custo computacional. Por sua vez, o critério de parada para o processo iterativo é o seguinte. Dado um limiar pré-estabelecido $\varepsilon > 0$, suficientemente pequeno, o procedimento é interrompido quando $\mathcal{E}^{(k+1)} < \varepsilon$, onde $\mathcal{E}^{(k+1)}$ é definido pela Eq. (71). A Fig. 29 mostra o fluxograma do Método de Tamanho de Passo de Segunda Ordem Adaptativo.

6.4 Resultados

O objetivo desta seção é ilustrar a operação do método proposto e comparar seu desempenho—principalmente em custo computacional—com outros métodos para calcular a probabilidade de falha, determinando o ponto mais provável de falha. Este estudo mostra um aumento de desempenho na solução da classe de problemas considerados, quando comparado com alguns métodos recentes, o que pode indicar um ganho computacional na solução de problemas semelhantes.

Figura 29 - Fluxograma do Método de Tamanho de Passo de Segunda Ordem Adaptativo.



Fonte: O autor, 2019.

6.4.1 Uma Primeira Comparação Entre SLA e ASOSL

Primeiro será avaliado o desempenho do método ASOSL, com relação ao comportamento do erro $\mathcal{E}^{(k)}$, o tamanho do passo $\tau^{(k)}$ e a sequência até o ponto mais provável de

falha, para o problema definido na Eq. (68), com a função de desempenho especificada pela Eq. (70). O desempenho do SLA neste problema, conforme discutido na Seção 6.1, foi a motivação para desenvolver o ASOSL.

Em todas as simulações apresentadas nesta seção, o mesmo critério de parada é usado e coincide com o adotado por Yi e Zhu (2016); o algoritmo é interrompido quando $\varepsilon^{(k)} \leq 10^{-3}$. Além disso, a estimativa inicial é sempre escolhida como a origem do espaço normal padrão. O parâmetro usado na extensão do intervalo da busca linear é $\delta_\eta = 10^{-4}$ e a técnica de *backtracking* foi executada com $\alpha_b = 10^{-4}$ e $s_b = 0,5$. Após apenas 8 iterações, o ASOSL obteve $\mathbf{u}^* = (-1,5234, -1,9822)$, o ponto ótimo.

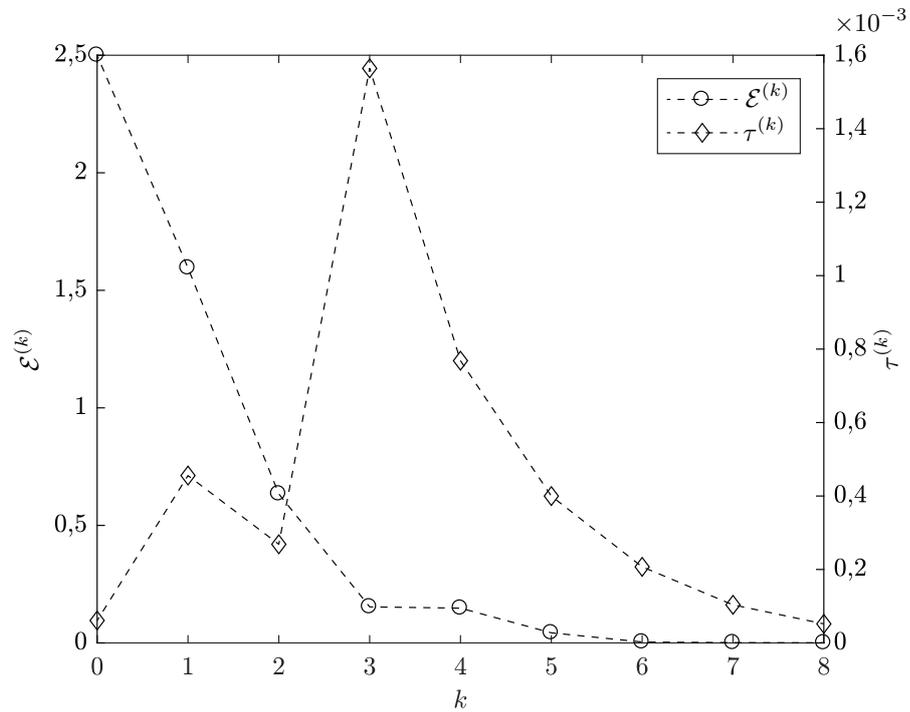
Comparando as Figs. 26 e 30, supõe-se que o ASOSL se beneficie da estratégia de tamanho do passo adaptável, para que o método atinja a solução com menos iterações que o SLA. No início, o ASOSL adota um tamanho de passo reduzido, em comparação com os valores usados para $\lambda^{(k)}$ no SLA. Este comportamento revela um recurso importante, que diz respeito à não dependência de uma boa escolha para o tamanho de passo inicial no ASOSL, em detrimento do que ocorre com o SLA. Parece claro também a capacidade do método em aumentar o valor de $\tau^{(k)}$, quando necessário, enquanto o erro continua diminuindo.

Embora a estimativa inicial do tamanho de passo no SLA tenha sido maior que a calculada inicialmente pelo ASOSL, o tamanho de passo do SLA, $\lambda^{(k)}$, manteve-se constante por várias iterações. Isso revela que o SLA não reduz o tamanho de passo adequadamente. Além disso, Yi e Zhu (2016) declarar que o menor valor que o tamanho de passo deve atingir é $\lambda = 0,01$. Nos testes realizados, esta condição não foi levada em consideração, como pode ser visto na Fig. 26a, para que fosse possível analisar o comportamento dos valores assumidos para $\lambda^{(k)}$. Se o tamanho de passo fosse restringido da maneira proposta pelos autores, O SLA não convergiria para o problema em análise.

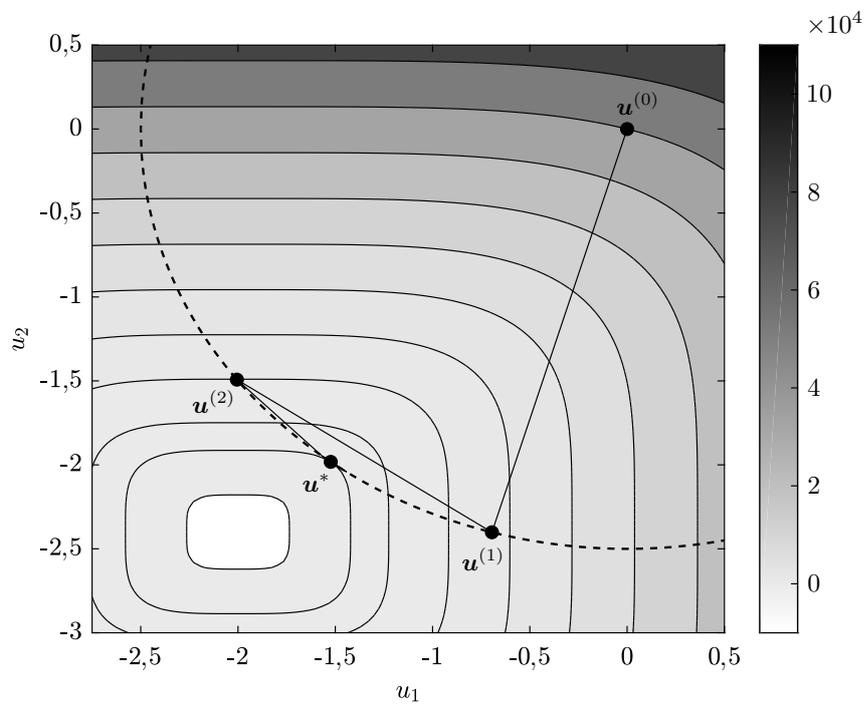
6.4.2 Análise Comparativa de Desempenho

A análise de desempenho realizada por Yi e Zhu (2016) será tomada como base, especialmente no que se refere à comparação dos custos computacionais. Como o método proposto é uma modificação da maneira como o tamanho de passo é calculado no SLA, é razoável considerar esta análise. Além disso, os métodos AMV e HMV, bem como os métodos Chaos Control (CC) e Modified Chaos Control (MCC) estão incluídos na comparação. Os dois últimos métodos foram projetados para evitar problemas de convergência enfrentados pelo AMV e HMV (discutido na Seção 6.1), portanto, é importante considerá-los na análise, porque eles adotam estratégias muito diferentes daquelas propostas no SLA e no ASOSL.

Figura 30 - Comportamento do método ASOSL no problema dado pela Eq. (70).



(a) Relação entre erro $\varepsilon^{(k)}$ e tamanho de passo $\tau^{(k)}$ para cada iteração k . As leituras de $\varepsilon^{(k)}$ estão à esquerda e as de $\tau^{(k)}$ estão à direita.



(b) Sequência de aproximações de $\mathbf{u}^{(k)}$.

Fonte: O autor, 2019.

6.4.2.1 Problema 1—Função Convexa

O primeiro problema de *benchmark* é definido pela função de desempenho convexa

$$g_1(\mathbf{X}) = 10 - \exp(X_1 - 7) - X_2, \quad (81)$$

com $X_1, X_2 \sim \mathcal{N}(6, 0,8^2)$ sendo duas variáveis aleatórias independentes e índice de confiabilidade alvo $\beta_t = 3$. A Tabela 3 apresenta \mathbf{u}^* obtido por cada método, bem como o valor da função de desempenho, o número de iterações até a convergência (n_{iter}) e o número de avaliações da função de desempenho (n_g). A técnica de *backtracking* no ASOSL foi executada com $\alpha_b = 10^{-4}$ e $s_b = 0,5$ em todos os casos.

Tabela 3 - Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_1(\mathbf{U})$ derivado da Eq. (81), para diferentes técnicas. O SLA é executado com $\lambda_{\text{max}} = 10$ e $c_{\text{SLA}} = 2,5$. O ASOSL é executado com $\delta_\eta = 10^{-4}$.

	\mathbf{u}^*	$G(\mathbf{u}^*)$	n_{iter}	n_g
AMV	(2,898, 0,775)	-0,3579	8	25
HMV	(2,898, 0,777)	-0,3579	24	73
CC	(2,894, 0,786)	-0,3553	87	262
MCC	(2,895, 0,785)	-0,3579	74	223
SLA	(2,898, 0,776)	-0,3579	11	34
ASOSL	(2,898, 0,775)	-0,3579	8	35

Fonte: O autor, 2019.

Nenhum método teve dificuldade em obter a solução correta. Observe que o AMV se beneficia do caráter convexo da função de desempenho, apresentando o menor valor para n_g , enquanto ASOSL e SLA tiveram custos computacionais praticamente iguais. Isso demonstra que o procedimento de busca linear proposto para o cálculo do tamanho de passo ótimo não torna o método mais lento. Além disso, pode-se observar uma redução no número de iterações realizadas, o que também interfere no custo computacional, uma vez que as demais operações realizadas durante o processo iterativo também consomem recursos computacionais. Nesse caso, portanto, o procedimento de busca linear afeta o número de avaliações da função de desempenho em cada iteração, enquanto o ponto mais provável de falha é obtido com menos iterações, em comparação com o SLA. Essa diferença se torna mais acentuada à medida que a complexidade do problema aumenta, como é visto nos pró-

ximos exemplos.

6.4.2.2 Problema 2—Função Côncava

Agora considere a função de desempenho côncava

$$g_2(\mathbf{X}) = 0,3X_1^2X_2 - X_2 + 0,8X_1 + 1, \quad (82)$$

onde $X_1 \sim \mathcal{N}(1,2, 0,42^2)$ e $X_2 \sim \mathcal{N}(1, 0,42^2)$ são variáveis aleatórias independentes e $\beta_t = 6$. Os resultados são relatados na Tabela 4. Como esperado, o método AMV não converge e, portanto, confirma-se que, de fato, seu bom desempenho é limitado a funções convexas. Observe, no entanto, que o conhecimento prévio deste recurso pode não ser uma tarefa trivial em certos problemas.

Tabela 4 - Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_2(\mathbf{U})$ derivado da Eq. (82), para diferentes técnicas. O SLA é executado com $\lambda_{\max} = 10$ e $c_{\text{SLA}} = 2,5$. O ASOSL é executado com $\delta_\eta = 1$.

	\mathbf{u}^*	$G(\mathbf{u}^*)$	n_{iter}	n_G
AMV	—	—	—	—
HMV	(-3,104, 5,135)	-2,2293	52	157
CC	(-3,105, 5,126)	-2,2259	67	202
MCC	(-3,106, 5,133)	-2,2293	28	85
SLA	(-3,105, 5,134)	-2,2293	27	82
ASOSL	(-3,109, 5,132)	-2,2293	7	36

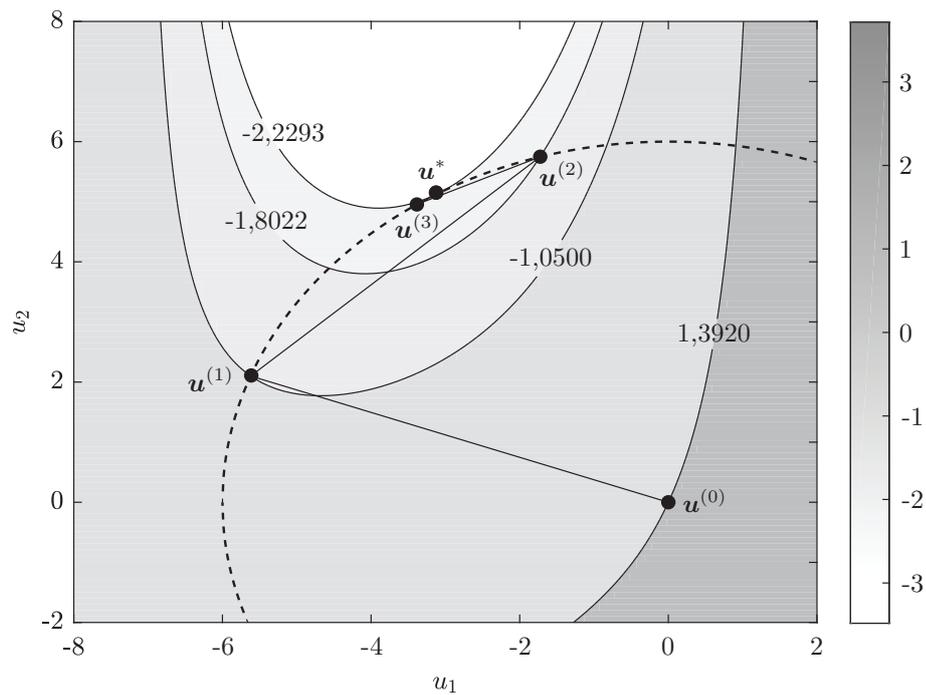
Fonte: O autor, 2019.

Embora o método HMV possa manipular funções côncavas, ele converge lentamente. O mesmo ocorre com CC e MCC. Mesmo o MCC, que tende a melhorar o desempenho do CC, ainda apresenta um desempenho ruim. Em relação ao SLA, Yi e Zhu (2016) apontam que o método converge, para este problema em particular, mais rapidamente se $\lambda_{\max} = 5$ é escolhido e mantido constante durante as iterações (isto é, escolhendo $c_{\text{SLA}} = 1$), diferentemente do que eles propuseram no algoritmo. Embora esta escolha represente uma vantagem, do ponto de vista do esforço computacional, a decisão sobre quando priorizar a estratégia de ajuste do tamanho do passo no SLA ou qual valor definir para $\lambda^{(k)}$ pode ser uma tarefa não trivial e afeta o desempenho do método.

O número de iterações do ASOSL é muito menor do que de todos os outros métodos.

Isso se deve principalmente à estratégia de calcular um tamanho de passo ótimo. Ao evitar o comportamento oscilatório, o processo iterativo é acelerado. A Fig. 31 mostra a sequência $\mathbf{u}^{(k)}$ calculada pelo ASOSL que, com apenas três iterações, já está muito próximo de \mathbf{u}^* . As demais iterações refinam a solução e não podem ser distinguidas devido à escala da imagem na Fig. 31.

Figura 31 - Histórico de iterações do método ASOSL na solução do problema $G_2(\mathbf{U})$, com $\delta_\eta = 1$.



Fonte: O autor, 2019.

6.4.2.3 Problema 3–Função Altamente Não-linear

Agora considere o problema dado por

$$g_3(\mathbf{X}) = 4 - (X_1 + 0,25)^2 + (X_1 + 0,25)^3 + (X_1 + 0,25)^4 - X_2, \quad (83)$$

onde X_1 e X_2 são variáveis aleatórias independentes que seguem a distribuição normal padrão. O índice de confiabilidade alvo é definido em $\beta_t = 3$. A Tabela 5 apresenta os resultados obtidos. É indicado por um traço quando o método não converge. Observe que os métodos que não usam o controle de tamanho do passo não convergem para a solução correta ou exibem comportamento oscilatório.

Tabela 5 - Resultados obtidos no problema definido pela Eq. (68), com a função de desempenho $G_3(\mathbf{U})$ derivado da Eq. (83), para diferentes técnicas. O SLA é executado com $\lambda_{\max} = 10$ e $c_{\text{SLA}} = 2,5$. O ASOSL é executado com $\delta_\eta = 1$.

	\mathbf{u}^*	$G(\mathbf{u}^*)$	n_{iter}	n_G
AMV	—	—	—	—
HMV	—	—	—	—
CC	—	—	—	—
MCC	—	—	—	—
SLA	(-1,357, 2,676)	0,2441	99	298
ASOSL	(-1,350, 2,679)	0,2440	7	33

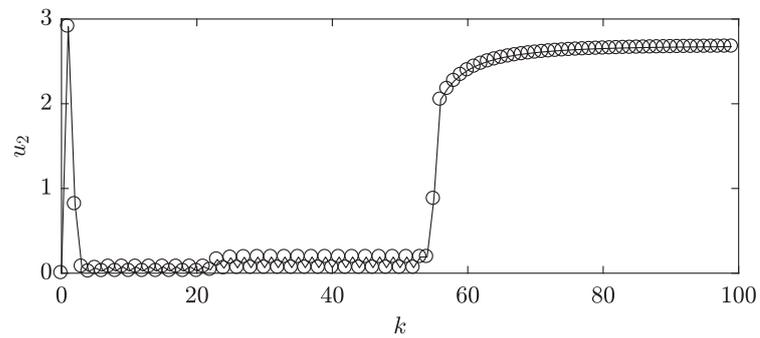
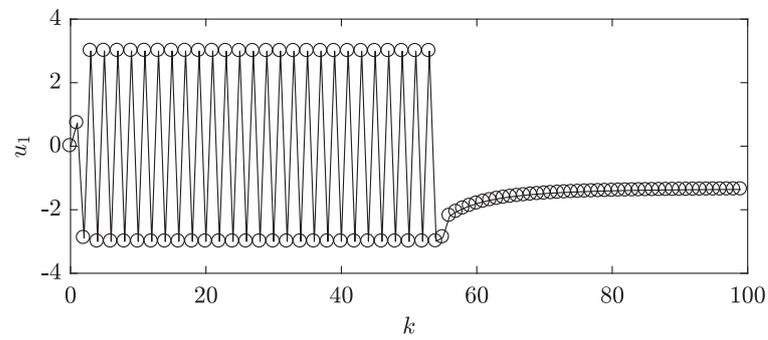
Fonte: O autor, 2019.

Embora o SLA e o ASOSL convirjam para resultados muito próximos, o SLA leva muito mais iterações para convergir. O método mostra um comportamento oscilatório, que é mantido durante várias iterações, antes que um valor diferente para $\lambda^{(k)}$ seja calculado. Esse comportamento é mostrado na Fig. 32a. Durante quase metade das iterações, a aproximação ao ponto mais provável de falha permanece oscilando perto de dois valores diferentes, até que o método atinja uma região na qual a sequência evolua constantemente para a solução do problema. No entanto, mesmo após o término das oscilações, o SLA exige várias iterações até que a solução seja refinada. Mesmo quando $\mathbf{u}^{(k)}$ está na vizinhança da solução, o critério para calcular o tamanho do passo é determinante para que o ponto resultante da iteração atual apresente pouca diferença em relação ao obtido na anterior, fazendo com que o algoritmo demore mais para convergir. Este fato também corrobora a afirmação de que a falta de uma estratégia para aumentar o tamanho do passo no SLA pode ser determinante para a lenta convergência do método. Por outro lado, o ASOSL não oscila nem refina a solução lentamente, convergindo muito rapidamente para a solução, como pode ser visto na Fig. 32b.

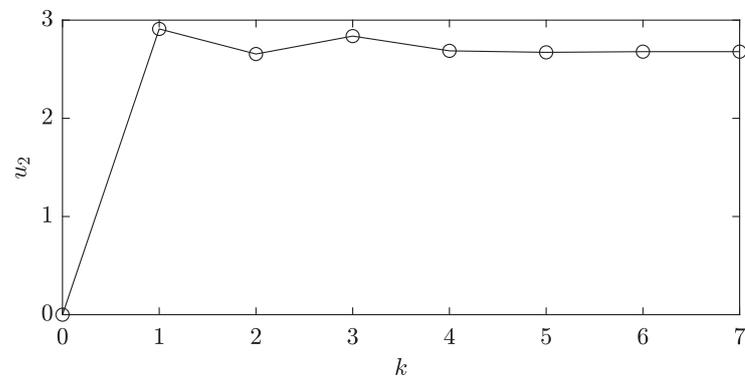
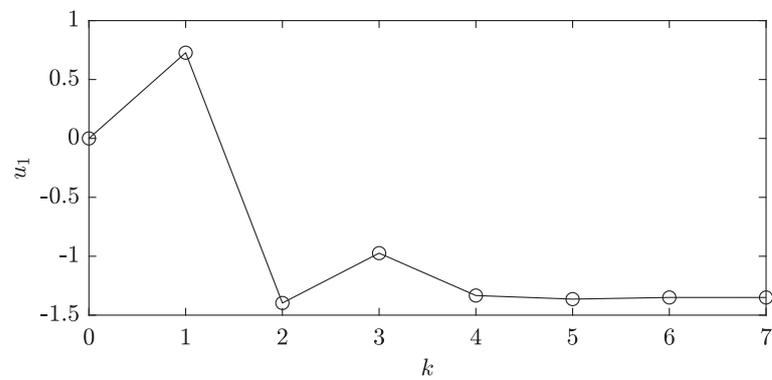
6.4.3 Análise de Sensibilidade do ASOSL

É importante avaliar o comportamento do ASOSL em relação à variação do parâmetro δ_η usado na extensão do intervalo da busca linear. Além disso, a análise de sensibilidade do ASOSL com relação aos parâmetros de controle α_b e s_b , para o procedimento de *backtracking*, é fornecido no contexto de um problema de engenharia. Por enquanto, considere novamente o problema definido pela Eq. (68), com a função performance dada pela Eq. (83).

Figura 32 - Histórico de iterações de SLA e ASOSL na resolução do problema $G_3(\mathbf{U})$.



(a) O SLA é executado com $\lambda_{\max} = 10$ e $c_{\text{SLA}} = 2,5$.



(b) O ASOSL é executado com $\delta_{\eta} = 1$.

Fonte: O autor, 2019.

A escolha de δ_η refere-se à correção de $\bar{t}^{(k)}$, calculado na Eq. (80). Portanto, se o algoritmo realizar muitas correções durante as iterações, o valor de δ_η pode ter uma grande influência no desempenho, mesmo considerando que as escolhas subsequentes do tamanho ótimo do passo dependem dos valores de $\mathbf{u}^{(k)}$ calculados anteriormente. Nesse sentido, é importante avaliar o número de correções necessárias na Eq. (78), quando $\bar{t}^{(k)} < 0$, e qual é o efeito no custo computacional. Os resultados são mostrados na Tabela 6.

Tabela 6 - Convergência do ASOSL para diferentes valores de δ_η , com $\alpha_b = 10^{-4}$ e $s_b = 0,5$.

δ_η	\mathbf{u}^*	$G(\mathbf{u}^*)$	n_{iter}	n_G	$\bar{\tau}$	$\bar{t}^{(k)} < 0$
0,0001	(-1,351, 2,679)	0,2440	7	45	0,8590	1
0,0005	(-1,352, 2,678)	0,2440	8	47	0,6385	1
0,0010	(-1,352, 2,678)	0,2440	8	46	0,6388	1
0,0050	(-1,348, 2,680)	0,2440	7	40	0,6216	1
0,0100	(-1,348, 2,680)	0,2440	7	39	0,6242	1
0,0500	(-1,350, 2,679)	0,2440	7	36	0,9184	1
0,1000	(-1,345, 2,682)	0,2441	9	43	0,5127	1
0,5000	(-1,350, 2,679)	0,2440	6	30	0,7530	1
1,0000	(-1,350, 2,679)	0,2440	7	33	0,8944	1

Fonte: O autor, 2019.

Além disso, o valor médio do tamanho do passo, representado por $\bar{\tau}$, e o número de correções, indicadas como $\bar{t}^{(k)} < 0$, são reportados. Primeiro, observe que em todos os casos, o ASOSL converge para a solução correta. Este fato revela que o método não apresenta sensibilidade à escolha deste parâmetro, o que pode ser uma vantagem na resolução de problemas mais complexos.

Observa-se que o número de avaliações da função de desempenho varia, mesmo quando o método realiza o mesmo número de iterações. Entretanto, a convergência do algoritmo não é influenciada e, mesmo com um possível aumento de n_G , O ASOSL ainda tem um desempenho muito melhor que o SLA neste problema, pois, como pode ser visto na Tabela 5, $n_{\text{iter}} = 99$ e $n_G = 298$ para o SLA. No caso do tamanho médio do passo, não existe uma relação simples com δ_η . Mesmo assim, $\bar{\tau}$ sempre permaneceu em uma faixa relativamente estreita em todos os casos analisados neste problema, o que confirma a baixa sensibilidade do método na escolha deste parâmetro.

6.4.4 Problema de Engenharia

Nesta seção, o desempenho do ASOSL é comparado a outras técnicas concorrentes em um problema de engenharia (um sistema dinâmico com dois graus de liberdade) exibindo maior complexidade do que as funções de desempenho anteriores. Além do método SLA, são consideradas técnicas mais recentes nesta análise, a saber, o método SMV e sua variante híbrida, conhecida como HSMV, desenvolvidas por Keshtegar e Hao (2017), bem como os métodos LDMV e DCC, propostos por Yaseen e Keshtegar (2018) e Keshtegar e Chakraborty (2018), respectivamente. Por razões de concisão, os métodos não são detalhados neste trabalho, mas apenas o esboço de seus principais atributos.

Os métodos SMV, HSMV, LDMV e DCC, bem como vários outros na literatura, calculam o ponto mais provável de falha através da relação de recorrência dada por

$$\tilde{\mathbf{u}}^{(k+1)} = \mathbf{u}^{(k)} + \lambda^{(k)} \left(\mathbf{u}_p^{(k)} - \mathbf{u}^{(k)} \right), \quad (84)$$

onde $\lambda^{(k)}$ é o tamanho do passo iterativo e $\mathbf{u}_p^{(k)}$ está relacionado a um valor estimado de $\mathbf{u}^{(k+1)}$, geralmente calculado pelo AMV. Então, $\tilde{\mathbf{u}}^{(k+1)}$ é normalizado para a hipersfera de raio β_t , resultando na aproximação $\mathbf{u}^{(k+1)}$. Os métodos diferem em como o tamanho do passo $\lambda^{(k)}$ é calculado.

A seguir, é realizada uma análise acerca do desempenho dos métodos em relação à convergência das soluções obtidas, os custos computacionais e da sensibilidade nos parâmetros de controle. Usa-se o mesmo critério de parada de antes: os processos iterativos são interrompidos quando $\mathcal{E}^{(k+1)} \leq 10^{-3}$. Além disso, a estimativa inicial é $\mathbf{u}^{(0)} = \mathbf{0}$ para todos os casos.

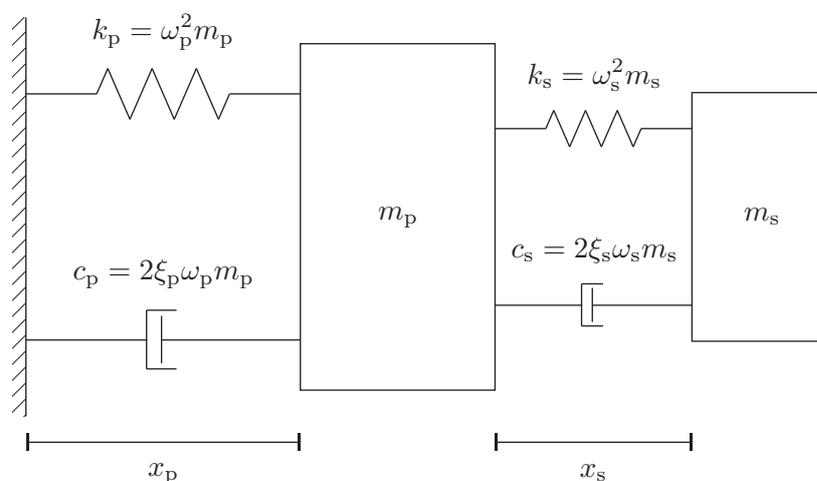
6.4.4.1 Sistema Dinâmico com Dois Graus de Liberdade

Considere um sistema dinâmico com dois graus de liberdade, representado esquematicamente na Fig. 33, discutido por Igusa e Der Kiureghian (1985). O sistema possui subsistemas primário e secundário acoplados e é definido pelas massas (m_p e m_s), rigidez da mola (k_p e k_s), taxas de amortecimento (ξ_p e ξ_s) e deslocamento do oscilador em relação à base (x_p e x_s), onde os subscritos se referem ao sistema primário e secundário.

A função de desempenho é avaliada com base na capacidade da mola do sistema secundário (F_s), considerando seu deslocamento relativo quadrado médio, cuja expressão, obtida por Igusa e Der Kiureghian (1985), é

$$g(\mathbf{X}) = F_s - 3k_s \sqrt{\frac{\pi S_0}{4\xi_s \omega_s} \left(\frac{\xi_a \xi_s}{\xi_p \xi_s (4\xi_a^2 + \theta^2) + m_r \xi_a^2} \frac{(\xi_p \omega_p^3 + \xi_s \omega_s^3) \omega_p}{4\xi_a \omega_a^4} \right)},$$

Figura 33 - Esquema de sistema de vibração dinâmico com dois graus de liberdade.



Fonte: Adaptado de Igusa e Der Kiureghian (1985).

com $\mathbf{X} = (m_p, m_s, k_p, k_s, \xi_p, \xi_s, F_s, S_0)$. Todas as variáveis aleatórias deste problema têm distribuições log-normal, com médias e desvios-padrão listados na Tabela 7. A massa relativa $m_r = m_s/m_p$ e o parâmetro $\xi_a = (\xi_p + \xi_s)/2$ representa o amortecimento médio. S_0 representa a intensidade do ruído branco e $\theta = (\omega_p - \omega_s)/\omega_a$ é um parâmetro de ajuste, com $\omega_a = (\omega_p + \omega_s)/2$ denotando a frequência média, com $\omega_p = \sqrt{k_p/m_p}$ e $\omega_s = \sqrt{k_s/m_s}$.

Tabela 7 - Propriedades estatísticas das variáveis aleatórias do problema de vibração com dois graus de liberdade.

Variável aleatória	m_p	m_s	k_p	k_s	ξ_p	ξ_s	F_s	S_0
Média	1	0,01	1	0,01	0,05	0,02	15	100
Desvio-padrão	0,1	0,001	0,2	0,002	0,02	0,01	1,5	10

Fonte: Reproduzido de Kiureghian e de Stefano (1991, p. 2918).

Este é um bom estudo de caso, pois sua função de desempenho é altamente não-linear e suas variáveis têm diferentes ordens de magnitude. O índice de confiabilidade alvo é $\beta_t = 1,75$, o mesmo usado por Yaseen e Keshtegar (2018). Esse problema também foi analisado por Keshtegar e Hao (2017) e Keshtegar e Chakraborty (2018) no contexto da análise de confiabilidade inversa. Embora todos esses três estudos tenham apresentado 0,001 como o desvio-padrão da rigidez da mola no sistema secundário, acredita-se que o valor correto, coerente com os resultados relatados, é 0,002, que é o valor adotado nos testes realizados (veja a Tabela 7) e é o mesmo que em Kiureghian e de Stefano (1991).

6.4.4.2 Comparação de Custo Computacional

Inicialmente, considera-se uma execução típica de cada um dos métodos, a fim de verificar possíveis problemas, como a ocorrência de oscilações ou a obtenção de soluções não ideais. Para isso, todos os métodos são executados nas mesmas condições, considerando as seguintes opções para os parâmetros de controle específicos de cada um: no ASOSL, $\delta_\eta = 1$ e os demais parâmetros são mantidos como nos testes anteriores ($\alpha_b = 10^{-4}$ e $s_b = 0,5$), bem como para o SLA, que foi executado com $\lambda_{\max} = 10$ e $c_{\text{SLA}} = 2,5$. No caso dos métodos SMV e HSMV, que têm parâmetros em comum, o coeficiente de ajuste do tamanho do passo foi escolhido como $C = 0,95$ e o fator de potência como $P = 1,25$. Por outro lado, os fatores limitantes e adaptativos do LDMV foram definidos como $\rho_{\text{LDMV}} = 0,8$ e $\delta_{\text{LDMV}} = 0,96$, respectivamente. Finalmente, o único parâmetro do DCC, o fator de escala adaptativo, foi escolhido como $c_{\text{DCC}} = 0,5$. Todas as escolhas seguem as sugestões dos autores para cada uma destas técnicas.

A Fig. 34 mostra o comportamento das sequências aproximando-se do ponto mais provável de falha calculado por cada método. Para o índice de confiabilidade estabelecido, $G(\mathbf{u}^*) \approx 0,875$. A maioria das técnicas convergiu com poucas iterações, exceto o SLA, que exigiu 55 passos até que o resultado ótimo fosse obtido (apesar disso, apenas as 20 primeiras iterações são mostradas). Como esperado, nenhuma técnica analisada apresentou comportamento oscilatório.

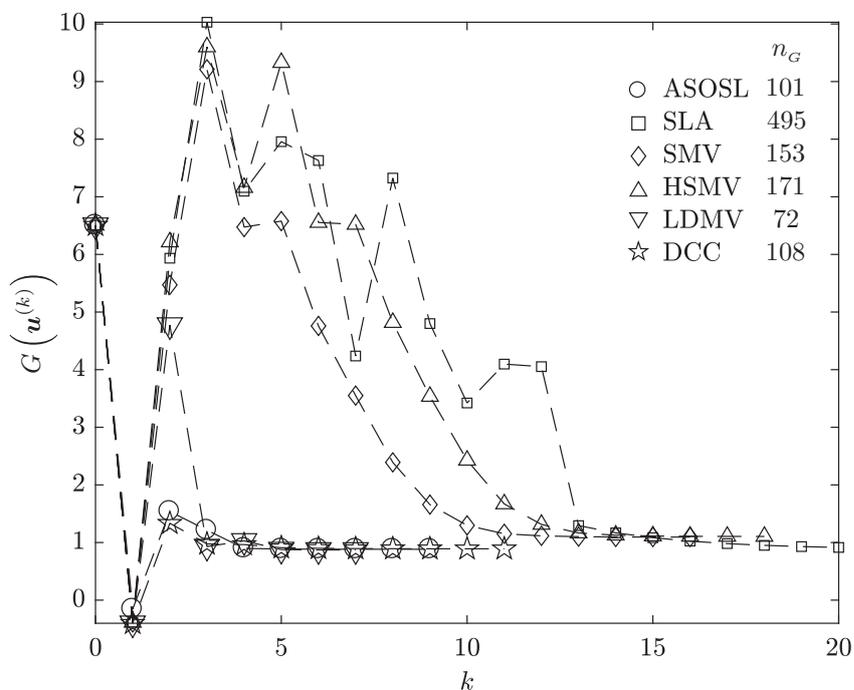
O método SMV não alcançou a solução correta do problema, uma vez que o valor obtido foi $G(\mathbf{u}_{\text{SMV}}^*) = 1,0943$. Mesmo usando sua variante híbrida, a sequência de aproximações mostrou-se promissora nas primeiras iterações, mas o resultado final também não foi alcançado, pois $G(\mathbf{u}_{\text{HSMV}}^*) = 1,1056$. Pelo menos no problema analisado, o uso da variante híbrida do SMV não se justifica, pois o aumento no número de iterações não levou a melhores resultados.

Por outro lado, os métodos LDMV e DCC mostraram-se eficientes para solucionar esse problema, juntamente com o ASOSL, obtendo os resultados esperados com poucas iterações e avaliações da função de desempenho, especialmente o LDMV que foi o melhor em ambos os aspectos. Apesar disso, os resultados correspondem ao comportamento dos métodos apenas para os parâmetros utilizados nesta execução e, portanto, é razoável que a sensibilidade dos algoritmos, relacionada à variação de seus parâmetros de controle, seja analisada, o que é feito a seguir.

6.4.4.3 Análise de Sensibilidade dos Parâmetros de Controle

Primeiro, considere o método LDMV. O parâmetro ρ_{LDMV} representa o valor mínimo que o tamanho do passo pode assumir em uma determinada iteração k e δ_{LDMV} é o fator

Figura 34 - Valor da função de desempenho em cada iteração k para cada método. Na caixa inserida, o número de avaliações da função de desempenho (n_G) até que a convergência é indicada, juntamente com a sigla de cada método.



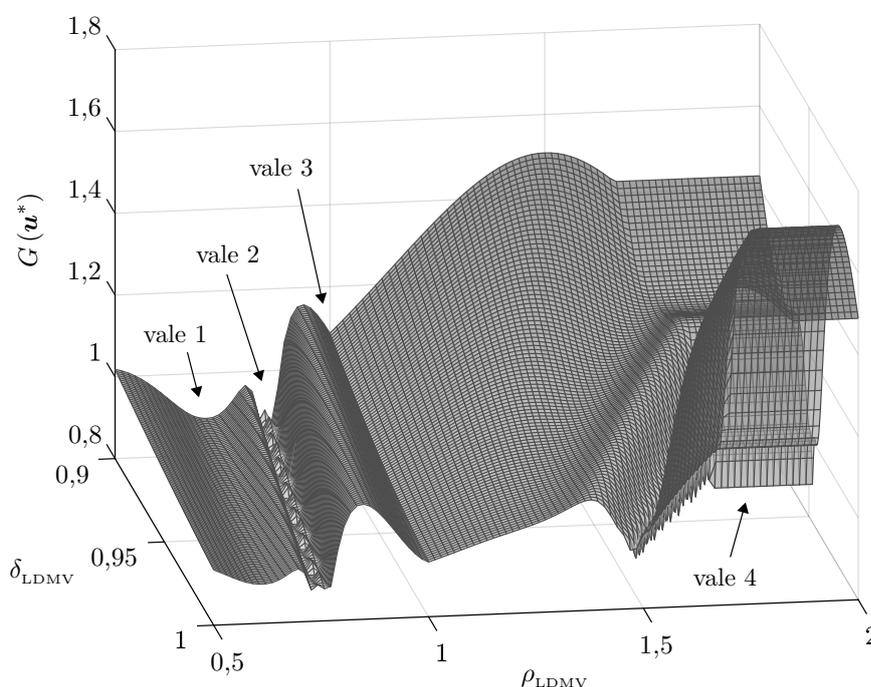
Fonte: O autor, 2019.

de amplificação do tamanho do passo. De acordo com Yaseen e Keshtegar (2018), deve-se escolher $\rho_{\text{LDMV}} > 0$ e $0,9 \leq \delta_{\text{LDMV}} < 1$. O método foi executado para diferentes combinações de tais parâmetros e a função de desempenho foi avaliada em cada ponto resultante, como pode ser visto na Fig. 35.

Observe que a superfície tem quatro vales. Neste problema, o LDMV obtém o otimizador do problema de confiabilidade somente se for executado com ρ_{LDMV} e δ_{LDMV} escolhido como uma combinação de valores localizados no vale 2, indicado na Fig. 35. Obviamente, a ocorrência deste fato é garantida apenas se houver algum conhecimento prévio do comportamento da função de desempenho. Apesar disso, a região cujas escolhas são eficazes para obter o ponto ótimo é muito limitada. Portanto, fica claro que o LDMV é muito sensível à escolha dos parâmetros, afetando o desempenho do algoritmo.

Por sua vez, o método DCC é um dos mais consistentes, dentre as técnicas que utilizam a Eq. (84). Seu parâmetro de controle, o fator adaptativo c_{DCC} , é usado para a determinação do tamanho do passo iterativo e deve satisfazer $0 < c_{\text{DCC}} < 1$, como apontado por Keshtegar e Chakraborty (2018). A Fig. 36 apresenta a convergência do algoritmo, bem como o número de avaliações da função de desempenho até a convergência, para diferentes valores de c_{DCC} .

Figura 35 - Resultados obtidos pelo LDMV para diferentes combinações dos parâmetros de controle ρ_{LDMV} e δ_{LDMV} .



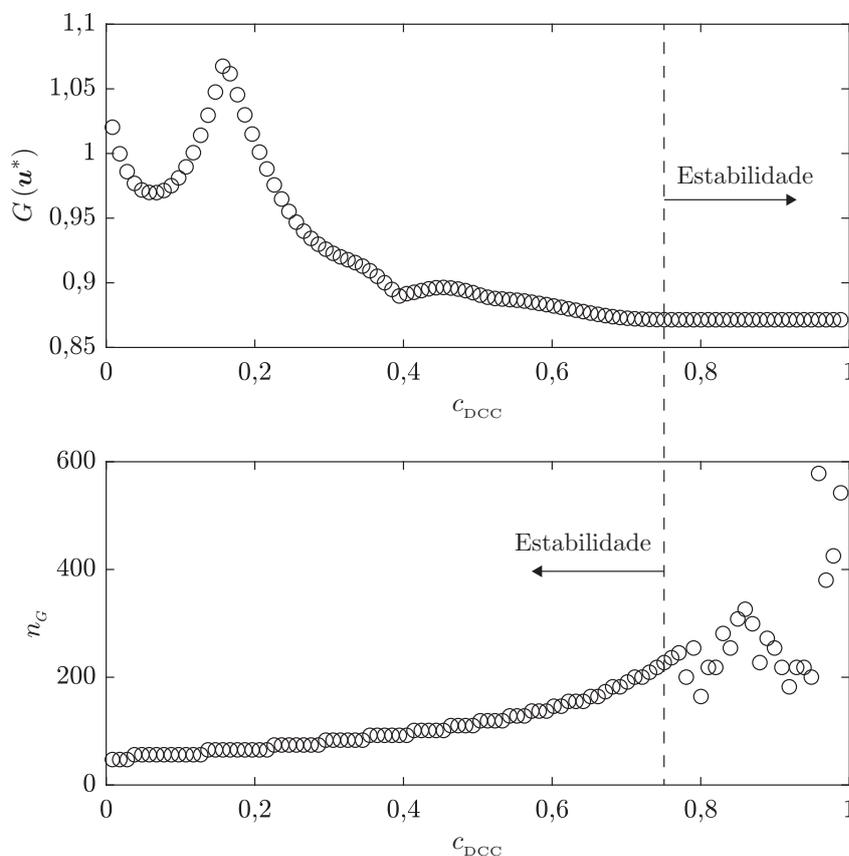
Fonte: O autor, 2019.

Na maior parte do intervalo em que c_{DCC} pode ser especificado, especialmente para valores que fornecem os melhores resultados em relação à função de desempenho, n_G aumenta severamente e apresenta instabilidade, conforme mostrado na Fig. 36. Neste problema, observa-se que a escolha de valores próximos ao limite superior em que c_{DCC} é definido, representa menos chance de convergência prematura, mas reflete na estabilidade do número de avaliações da função de desempenho, mesmo quando o fator adaptativo sofre apenas pequenas variações.

Para realizar uma análise equivalente à apresentada para os outros métodos, primeiro considere a variação de δ_η no ASOSL. Os parâmetros relacionados à busca linear, α_b e s_b , são mantidos constante, com os mesmos valores usados nos testes anteriores (ver a Seção 6.4.2). Fazendo δ_η variar no intervalo $10^{-8} \leq \delta_\eta \leq 1$, produzindo uma sequência de 100 valores igualmente espaçados, o método proposto foi executado considerando cada um destes valores. Em média, o valor da função de desempenho foi 0,8831, com um desvio-padrão aproximadamente igual a $2,2316 \times 10^{-15}$, o que significa que o método convergiu praticamente para a mesma solução em todas as 100 execuções. Em relação ao número de avaliações da função desempenho, o método realizou, em média, 102,08 avaliações, com desvio-padrão igual a 10,8.

Mesmo que as escolhas de $\alpha_b = 10^{-4}$ e $s_b = 0,5$ provaram-se ser eficientes nos testes anteriores, é importante que o desempenho do método seja avaliado quando estes parâme-

Figura 36 - Relação entre resultados obtidos e quantidade de avaliações da função de desempenho do método DCC para diferentes valores do parâmetro c_{DCC} .



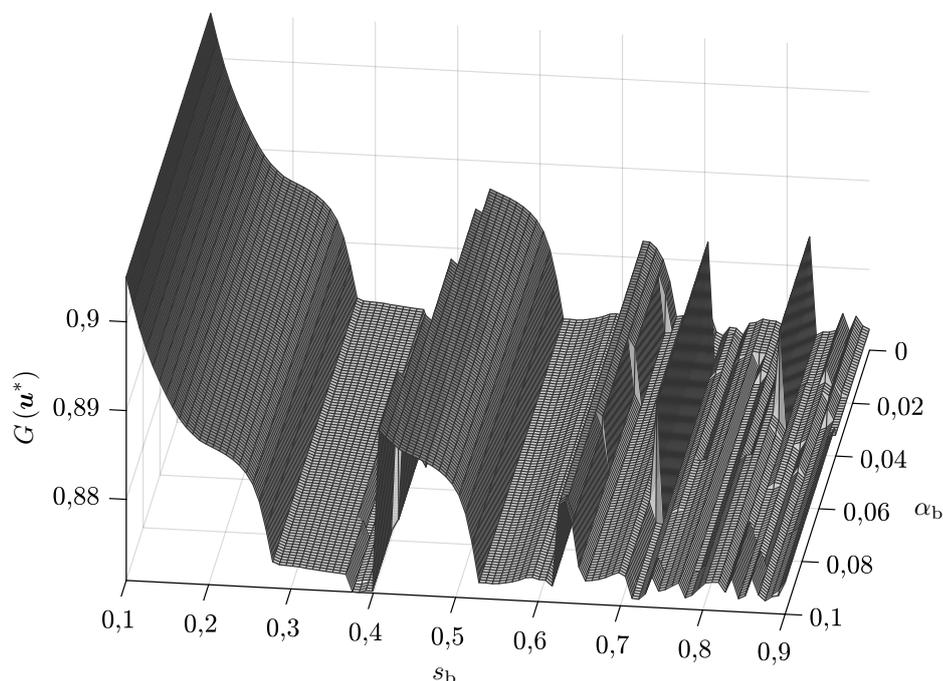
Fonte: O autor, 2019.

tros são modificados. A Fig. 37 apresenta os resultados obtidos para diferentes combinações de α_b e s_b no ASOSL, com $10^{-8} \leq \alpha_b \leq 10^{-1}$, $0,1 \leq s_b \leq 0,9$, mantendo $\delta_\eta = 1$.

Observe que o método não mostra grandes variações em relação aos valores calculados para a função de desempenho, que estão entre 0,8709 e 0,9051, particularmente em comparação com os alcançados pelo LDMV, que variam entre 0,8705 e 1,6355, como mostrado na Fig. 35. Além disso, a média e o desvio-padrão do conjunto de valores de $G(\mathbf{u}^*)$ usando o LDMV são 1,13 e 0,1858, respectivamente, enquanto para o ASOSL são 0,88 e 0,0079. No caso do DCC, o conjunto de valores obtidos para $G(\mathbf{u}^*)$ variam entre 0,8705 e 1,0666, com média e desvio-padrão iguais a 0,91 e 0,0545, respectivamente, cujos valores se referem à Fig. 36.

Quanto à taxa de convergência dos métodos, o ASOSL obteve $G(\mathbf{u}^*) \leq 0,875$ em 59,66% do número total de execuções com diferentes combinações de parâmetros, enquanto o LDMV alcançou esse resultado em apenas 1,43% dos casos e, em relação ao DCC, esse valor é igual a 34%. Considerando esses resultados em conjunto, fica claro que os valores ótimos computados da função de desempenho usando ASOSL apresentam dependência mais

Figura 37 - Resultados obtidos pelo ASOSL para diferentes combinações dos parâmetros de controle α_b e s_b , com $\delta_\eta = 1$.



Fonte: O autor, 2019.

Tabela 8 - Resultados estatísticos referentes ao custo computacional e à sensibilidade dos parâmetros de controle relacionados aos métodos ASOSL, LDMV e DCC, no problema do sistema dinâmico com dois graus de liberdade.

Método	Parâmetros	$G(\mathbf{u}^*)$				n_g	
		Intervalo	Média	Desvio-padrão	Conv. (%)	Média	Desvio-padrão
	$10^{-8} \leq \alpha_b \leq 10^{-1}$						
ASOSL	$0,1 \leq s_b \leq 0,9$ $\delta_\eta = 1$	(0,8709, 0,9051)	0,88	0,0079	59,66	151,56	88,71
LDMV	$0,5 \leq \rho_{LDMV} \leq 2$ $0,9 \leq \delta_{LDMV} < 1$	(0,8705, 1,6355)	1,13	0,1858	1,43	77,18	9,62
DCC	$0 < c_{DCC} < 1$	(0,8705, 1,0666)	0,91	0,0545	34	145,62	98,95

Fonte: O autor, 2019.

fraca dos parâmetros escolhidos, no problema de engenharia analisado.

No caso do número de avaliações da função de desempenho, o ASOSL executa 151,56 avaliações em média de $G(\mathbf{u})$ até a convergência (para os resultados apresentados na Fig. 37), com desvio-padrão igual a 88,71, enquanto o DCC realizou 145,62 avaliações da função de desempenho, em média, com desvio-padrão igual a 98,95 (para os resultados mostrados na Fig. 36). Por sua vez, o LDMV obteve, levando em consideração as execuções mostradas na

Fig. 35, os valores 77,18 e 9,62 para média e desvio-padrão relacionados ao número de avaliações da função de desempenho, respectivamente. A Tabela 8 resume todos os resultados obtidos usando os métodos ASOSL, LDMV e DCC.

7 OTIMIZAÇÃO MULTI-OBJETIVO BASEADA EM ROBUSTEZ E CONFIABILIDADE

Neste capítulo, é proposta uma formulação para problemas de otimização multi-objetivo considerando robustez e confiabilidade, e descrita a estrutura para acoplamento da avaliação da sensibilidade do vetor de variáveis de decisão e das restrições probabilísticas no procedimento de otimização. A formulação proposta é analisada através da solução de um problema de *benchmark* e, em seguida, são resolvidos alguns problemas de engenharia, a fim de validá-la e verificar a influência de cada tipo de incerteza no resultado ótimo do problema de otimização multi-objetivo.

7.1 Definição do Problema de Otimização com Incertezas

Considere o problema de otimização multi-objetivo, mostrado na Eq. (1), como base da definição de um novo problema do mesmo gênero. A proposta é que a formulação do novo problema, inicialmente determinística, isto é, sem que sejam consideradas as variações estocásticas do modelo, passe a considerar tais características através da avaliação de incertezas por meio de dois aspectos: robustez e confiabilidade.

No contexto analisado, um modelo matemático é descrito por variáveis independentes determinísticas e estocásticas, que compõem o conjunto das variáveis de projeto. O primeiro tipo, as variáveis determinísticas, são denotadas por \mathbf{d} e delimitadas por $\mathbf{d}_{\text{inf}} \leq \mathbf{d} \leq \mathbf{d}_{\text{sup}}$. Por sua vez, as variáveis estocásticas são representadas por \mathbf{X} e definidas no intervalo $\mathbf{X}_{\text{inf}} \leq \mathbf{X} \leq \mathbf{X}_{\text{sup}}$. Essencialmente, a formulação proposta diferencia-se da determinística quanto à avaliação das incertezas em torno das variáveis de projeto.

Como apresentado no Cap. 3, a otimização robusta diz respeito à obtenção de soluções ótimas, mas considerando a influência da sensibilidade das variáveis de decisão no valor da função-objetivo. Seja δ o vetor de incertezas associadas a cada componente de \mathbf{d} . A sensibilidade das variáveis de decisão devem ser avaliadas tanto na função-objetivo, quanto nas restrições de igualdade e desigualdade.

Por sua vez, considerar as incertezas do modelo no sentido da sua confiabilidade corresponde a transformar as restrições de desigualdade do problema em restrições probabilísticas e, conseqüentemente, obter o conjunto de soluções viáveis do problema de otimização multi-objetivo nestas condições significa que tais soluções satisfazem a um nível de confiabilidade prescrito, caracterizado através de uma probabilidade de falha, como abordado na Seção 4.1.2. Na formulação proposta, a probabilidade de falha é determinada por $\Phi(-\beta_i)$, onde Φ é a função de distribuição acumulada da distribuição normal padrão e β_i representa o índice de confiabilidade alvo.

Seja $F(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) = (f_1(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}), \dots, f_m(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}))^T$. A definição do problema de otimização multi-objetivo baseado em robustez e confiabilidade é dada por

$$\begin{aligned}
 & \arg \min_{\mathbf{d}, \mathbf{X}} F(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) \\
 & \max \quad \beta \\
 \text{Sujeito a} \quad & P[g_i(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) \leq 0] \leq \Phi(-\beta), \\
 & h_j(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) = 0 \\
 & \mathbf{d}_{\text{inf}} \leq \mathbf{d} \leq \mathbf{d}_{\text{sup}} \\
 & \beta_{\text{inf}} \leq \beta \leq \beta_{\text{sup}}
 \end{aligned} \tag{85}$$

onde $i = 1, \dots, p$ e $j = 1, \dots, q$, tal que p e q representam a quantidade de restrições de desigualdade e igualdade, respectivamente. A seguir, é proposta uma metodologia genérica que demonstra como avaliar a influência do vetor de incertezas no valor dos objetivos, assim como a manipulação das restrições probabilísticas durante o procedimento de otimização.

7.2 Avaliação de Incertezas na Otimização Multi-objetivo

Esta seção destina-se a detalhar um esquema acerca da estruturação de um procedimento de otimização capaz de obter o conjunto ótimo de Pareto da Eq. (85). O procedimento descrito considera o uso de métodos de otimização multi-objetivo baseados em metaheurísticas. Primeiramente, considere uma população inicial dada por $\mathcal{P}^{(0)} = (\mathbf{P}_1^{(0)}, \dots, \mathbf{P}_v^{(0)})^T$, onde $\mathbf{P}_\ell^{(0)} = (d_1^{(0)}, \dots, d_n^{(0)}, \beta^{(0)})$, para $\ell = 1, \dots, v$. Neste caso, a população é representada por uma matriz $v \times (n + 1)$, sendo que a última coluna de \mathcal{P} é formada por candidatos a índice de confiabilidade, o qual deve ser maximizado.

A cada iteração k do método de otimização, deve-se gerar um conjunto $\mathbf{R}^{(k)}$ com H vetores de números aleatórios, cada um deles com $n + 1$ componentes. Estes números aleatórios devem ser uniformemente distribuídos no intervalo $(0, 1)$. O conjunto $\mathbf{R}^{(k)}$ é necessário para a avaliação da robustez dos candidatos a solução. É importante que o mesmo conjunto de números aleatórios seja usado para avaliar todos os v candidatos a solução presentes na população, em uma dada iteração.

A cada avaliação da função-objetivo, primeiro é preciso levar em conta a análise de robustez. Para exemplificar, considere a técnica de Média Efetiva, apresentada na Seção 3.3.1.2. Cada uma das linhas da matriz $\mathbf{R}^{(k)}$ deve ser utilizada para gerar amostras aleatórias ao redor de um dado candidato $\mathbf{P}_\ell^{(k)}$. As amostras, denotadas por $\mathcal{R}^{(k)}$, cuja matriz também tem dimensão $H \times (n + 1)$, assim como $\mathbf{R}^{(k)}$, devem estar dispostas no intervalo dado por $\mathcal{J} = [\mathbf{P}_\ell^{(k)} - \boldsymbol{\delta}\mathbf{P}_\ell^{(k)}, \mathbf{P}_\ell^{(k)} + \boldsymbol{\delta}\mathbf{P}_\ell^{(k)}]$, onde $\boldsymbol{\delta}$ é o vetor de incertezas associadas às variáveis $d_1^{(k)}, \dots, d_n^{(k)}$, dado por $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n, 0)$. A última coordenada de $\boldsymbol{\delta}$ assegura que a

análise de robustez não envolverá o índice de confiabilidade. Cada uma das linhas da matriz de amostras aleatórias são dadas por $\mathcal{R}_j^{(k)} = \mathcal{J}_{\text{inf}} + \mathbf{R}_j^{(k)} (\mathcal{J}_{\text{sup}} - \mathcal{J}_{\text{inf}})$, para $j = 1, \dots, H$, onde os subscritos “inf” e “sup” representam os limites inferior e superior do intervalo, respectivamente. Portanto, uma nova função, responsável por quantificar a robustez, é definida como

$$F_\gamma^{\text{eff}}(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) = \frac{1}{H} \sum_{j=1}^H F_\gamma(\mathcal{R}_j, \mathbf{X}, \boldsymbol{\delta}), \quad (86)$$

para $\gamma = 1, \dots, m$.

Para cada uma das amostras avaliadas na Eq. (86) pela análise de robustez, é preciso verificar se as restrições probabilísticas são satisfeitas. Essa verificação é realizada utilizando um método de análise de confiabilidade inversa. Desta forma, pode-se determinar se uma amostra, originada da perturbação de um candidato a solução, satisfaz a restrição probabilística para um determinado índice de confiabilidade alvo, dado por $\mathcal{R}_{j,n+1}^{(k)}$. Para cada restrição probabilística do problema de otimização, deve-se executar a técnica de análise de confiabilidade inversa independentemente. Neste caso, considere uma estimativa inicial $\mathbf{u}^{(0)}$ para as variáveis aleatórias do problema, após a sua transformação para o espaço normal padrão \mathbf{U} , levando em conta a sua média e desvio-padrão. O resultado da execução do método de análise de confiabilidade inversa é representado por G_i^* , isto é, o valor resultante da função de desempenho, para $i = 1, \dots, p$.

Quando um dado vetor de variáveis de projeto atende às condições de uma restrição probabilística, o resultado obtido pelo método de análise de confiabilidade inversa satisfaz $G_i^* > 0$. Uma nova função-objetivo \mathcal{F} , que considera as restrições probabilísticas do problema, para a penalização da função-objetivo robusta \mathbf{F}^{eff} , é definida como

$$\mathcal{F}_\gamma(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) = F_\gamma^{\text{eff}}(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) + \Psi \sum_{i=1}^p \max(-G_i^*, 0), \quad (87)$$

para $\gamma = 1, \dots, m$, onde Ψ é o termo de penalização, geralmente adotado como um valor suficientemente grande para os casos em que \mathbf{F} é minimizada, e F_γ^{eff} é definida pela Eq. (86). Cada vetor de amostras $\mathcal{R}_j^{(k)}$ deve ser avaliado pela Eq. (87). Por meio desta técnica de penalização, quando todas as restrições probabilísticas são satisfeitas, tem-se que $\mathcal{F}(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) = \mathbf{F}^{\text{eff}}(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta})$.

A Eq. (87) representa uma generalização da função capaz de avaliar robustez e confiabilidade do problema de otimização multi-objetivo definido pela Eq. (85). O procedimento descrito é executado durante a aplicação dos operadores genéticos do método de otimização multi-objetivo. Então, candidatos a solução (ou mesmo amostras aleatórias geradas na vizinhança destes candidatos) que são penalizados, tendem a não prosseguir para as gerações seguintes. Note que, na formulação proposta na Eq. (87), a função \mathcal{F} é composta pela

soma das contribuições da análise de robustez e da otimização baseada em confiabilidade. Isso deixa claro que, ao final do procedimento de otimização, o conjunto de pontos obtidos no espaço objetivo é composto por indivíduos pouco sensíveis a perturbações externas (robustos) e que satisfazem as restrições probabilísticas do sistema, isto é, que estão dentro dos níveis de confiabilidade almejados. A Fig. 38 apresenta um fluxograma com as etapas principais da metodologia proposta.

7.2.1 Custo Computacional

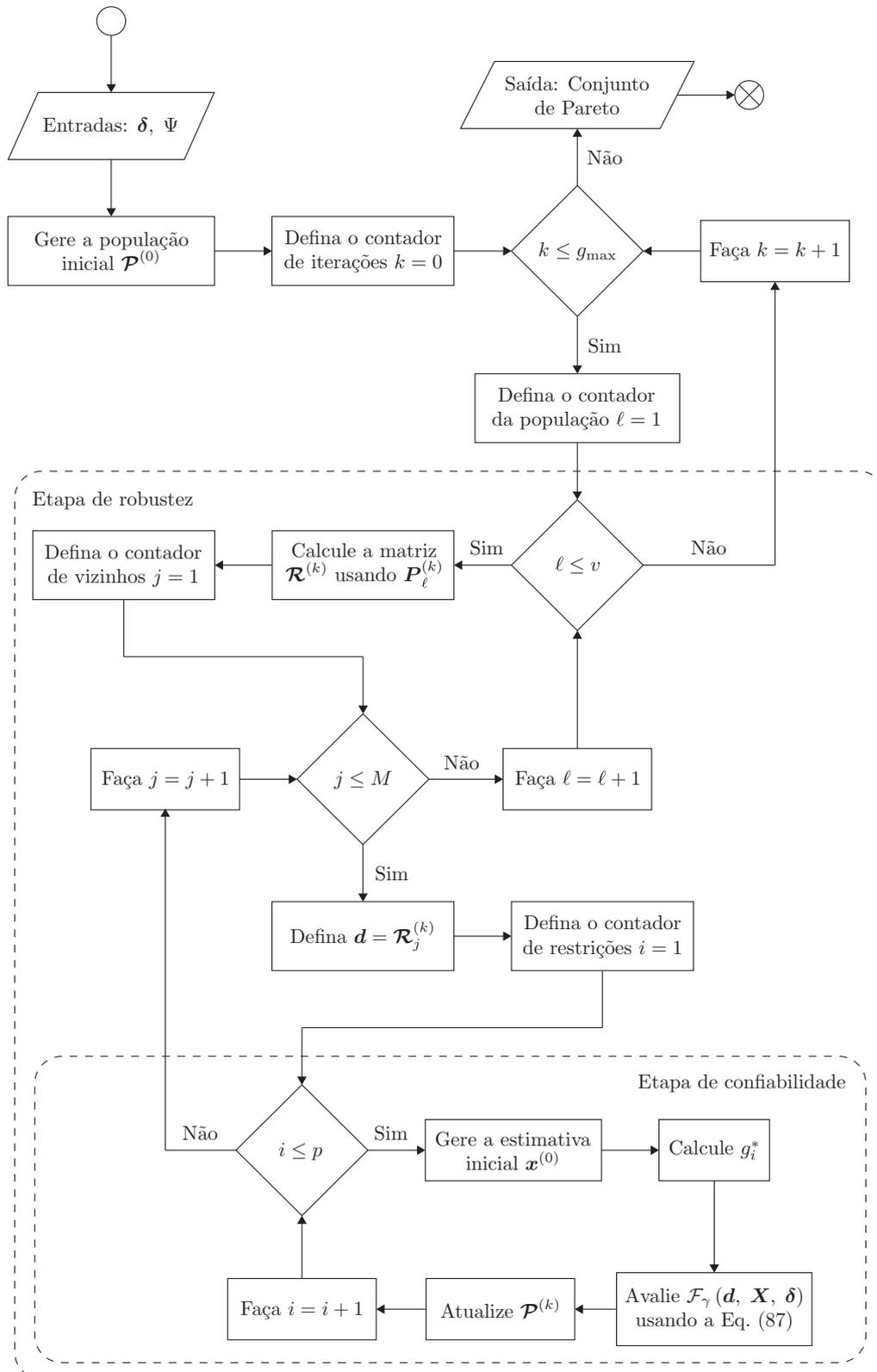
Um ponto muito importante a se destacar é o custo computacional para obtenção de soluções robustas e confiáveis. Admita um custo mensurado pelo número de avaliações da função-objetivo e das funções de desempenho. Pela formulação sugerida, gera-se H amostras aleatórias na vizinhança de NP indivíduos da população por G_{\max} gerações da metaheurística. Todas as restrições probabilísticas são resolvidas para cada uma destas amostras e as potenciais soluções que satisfazem tais restrições são empregadas na avaliação da função-objetivo. Assim, inicialmente considere C_m o custo de se computar o conjunto de soluções ótimas de um problema de otimização multi-objetivo determinístico (sem considerar incertezas). Este custo pode ser representado por $C_m \approx G_{\max} \times NP \times C_v$, onde C_v é um custo variável, que depende da estratégia evolutiva adotada na metaheurística. Por sua vez, admita que n_g é o custo necessário para se computar uma restrição probabilística. Aproximadamente, um conjunto ótimo de Pareto para um problema com incertezas é calculado, por meio da formulação proposta, com custo dado por $C_m \times H \times p \times n_g$, onde p é o número de restrições probabilísticas.

Portanto, note que a obtenção de soluções levando em conta as incertezas do modelo, em geral, está associada ao emprego de mais recursos computacionais do que no processo de obtenção de soluções determinísticas, como esperado. Comparado ao custo de se obter as soluções determinísticas (C_m), o cômputo de resultados robustos e confiáveis é, aproximadamente, $H \times p \times n_g$ maior. Naturalmente, essa grandeza é variável, uma vez que n_g representa um custo que não é constante, isto é, o número de avaliações de uma dada função de desempenho depende de diversos fatores, como o método de análise de confiabilidade inversa utilizado e o seu critério de parada. Tal variação depende, ainda, do fator C_v relacionado à metaheurística.

7.3 **Avaliação da Formulação Proposta**

A seguir, a formulação proposta para obtenção de soluções robustas e confiáveis em problemas de otimização multi-objetivo é avaliada. Inicialmente é resolvido um problema de *benchmark* clássico da literatura, a fim de fornecer uma visão geral do comportamento

Figura 38 - Fluxograma da metodologia geral para incorporação de robustez e confiabilidade em problemas de otimização multiobjetivo.



Fonte: O autor, 2019

dos resultados obtidos. Em seguida, são resolvidos alguns problemas de engenharia química. Em todos os casos, o método de análise de confiabilidade inversa, empregado na solução das restrições probabilísticas, é o Método de Tamanho de Passo de Segunda Ordem Adaptativo, proposto no Cap. 6.

7.3.1 Problema Matemático de *Benchmark*

Inicialmente, considere o problema de otimização mono-objetivo restrito, dado por

$$\arg \min_{\mathbf{d}} f(\mathbf{d})$$

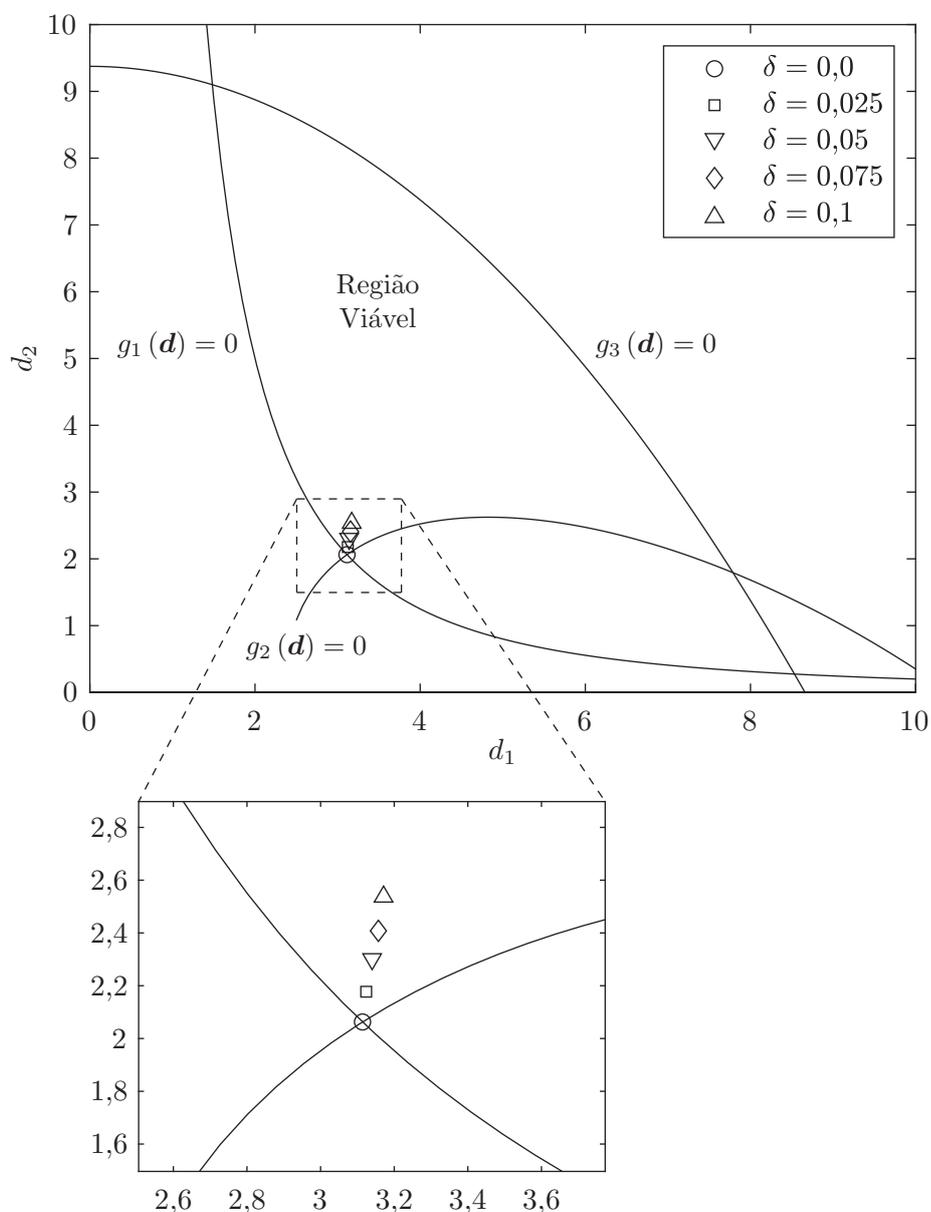
$$\begin{aligned} \text{Sujeito a } g_1(\mathbf{d}) &= \frac{d_1^2 d_2}{20} - 1 > 0 \\ g_2(\mathbf{d}) &= \frac{(d_1 + d_2 - 5)^2}{30} + \frac{(d_1 - d_2 - 12)^2}{120} - 1 > 0, \\ g_3(\mathbf{d}) &= \frac{80}{d_1^2 + 8d_2 + 5} - 1 > 0 \\ 0 &\leq d_1, d_2 \leq 10 \end{aligned} \quad (88)$$

onde $f(\mathbf{d}) = d_1 + d_2$. A sua solução pode ser obtida utilizando um método numérico de otimização capaz de considerar restrições de desigualdade, ou através de uma metaheurística associada a uma técnica que considere tais restrições durante o procedimento de otimização. Para o segundo caso, o método da Evolução Diferencial (STORN; PRICE, 1997) pode ser empregado, onde a função-objetivo é penalizada caso alguma das restrições seja violada. Nestas condições, a execução da DE por 100 gerações, com $NP = 50$ indivíduos, fator de amplificação $F = 0,5$ e probabilidade de cruzamento $CR = 0,8$ obtém o minimizador do problema dado pela Eq. (88), $\mathbf{d}^* = (3,113885, 2,062648)$, onde $f(\mathbf{d}^*) = 5,176532$.

Agora, admita o mesmo problema, mas no contexto da análise de robustez. Mais especificamente, considere a técnica de Média Efetiva, apresentada na Seção 3.3.1.2. A solução do problema sob análise, para diferentes níveis de robustez, isto é, para diferentes valores de δ , pode exprimir o comportamento do conjunto de soluções em relação à região viável do problema e, mais ainda, em comparação à sua solução determinística. Nestas circunstâncias, a DE é executada novamente, a fim de obter resultados independentes de problemas para cada nível de robustez, utilizando os mesmos valores de parâmetros do caso anterior. No caso da técnica de Média Efetiva, são usadas $H = 50$ amostras aleatórias, distribuídas ao redor de cada candidato a solução, geradas pela estratégia de Hipercubo Latino. Os resultados para níveis de sensibilidade das variáveis de projeto variando entre 0% e 10% são mostrados na Fig. 39.

Analisando os resultados obtidos, primeiro note que a solução determinística, representada pelo círculo na Fig. 39 e obtida quando $\delta_k = 0$, para $k = 1, 2$, está localizada em um extremo da região viável, delimitada pelas restrições g_1 e g_2 . O aumento do valor de δ

Figura 39 - Soluções do problema da Eq. (88) para diferentes níveis de robustez das variáveis de projeto.



Fonte: O autor, 2019

significa que a metaheurística tende a renunciar de pontos ótimos cuja sensibilidade é elevada, isto é, pontos nos quais qualquer ligeira perturbação no vetor de variáveis de projeto tende a produzir significativa diferença no valor da função-objetivo. Em geral, soluções determinísticas tendem a ser sensíveis. Portanto, o aumento do parâmetro de ruído δ consiste em abdicar do minimizador com melhor valor da função-objetivo, em favor da obtenção de uma solução mais robusta. Este comportamento confirma-se na Fig. 39, onde as soluções computadas mostram um aumento gradual da distância em relação ao otimizador determinístico, o mais sensível, haja vista que uma ligeira perturbação é capaz de movê-lo para fora

da região viável.

Por sua vez, considere as incertezas do problema do ponto de vista da análise de confiabilidade. Um problema de otimização baseada em confiabilidade pode ser formulado de diversas formas, a depender do modelo matemático e dos interesses inerentes ao projeto. No caso do problema estudado, uma possível formulação, proposta por Tu (1999) e utilizada em diversos trabalhos na literatura, é dada por

$$\begin{aligned} & \arg \min_{\mathbf{d}} f(\mathbf{d}) \\ \text{Sujeito a } & P[g_i(\mathbf{d}, \mathbf{X}) \leq 0] \leq \Phi(-\beta_t) \cdot \\ & 0 \leq d_1, d_2 \leq 10 \end{aligned} \quad (89)$$

onde $i = 1, 2, 3$. Neste problema, \mathbf{X} é um conjunto de variáveis aleatórias independentes com distribuição normal, onde $\mu_x = (d_1, d_2)$ e $\sigma_x = (0,3, 0,3)$, e o índice de confiabilidade alvo é $\beta_t = 3$. As funções que aparecem nas restrições do problema são

$$g_1(\mathbf{d}, \mathbf{X}) = \frac{X_1^2 X_2}{20} - 1 \quad (90a)$$

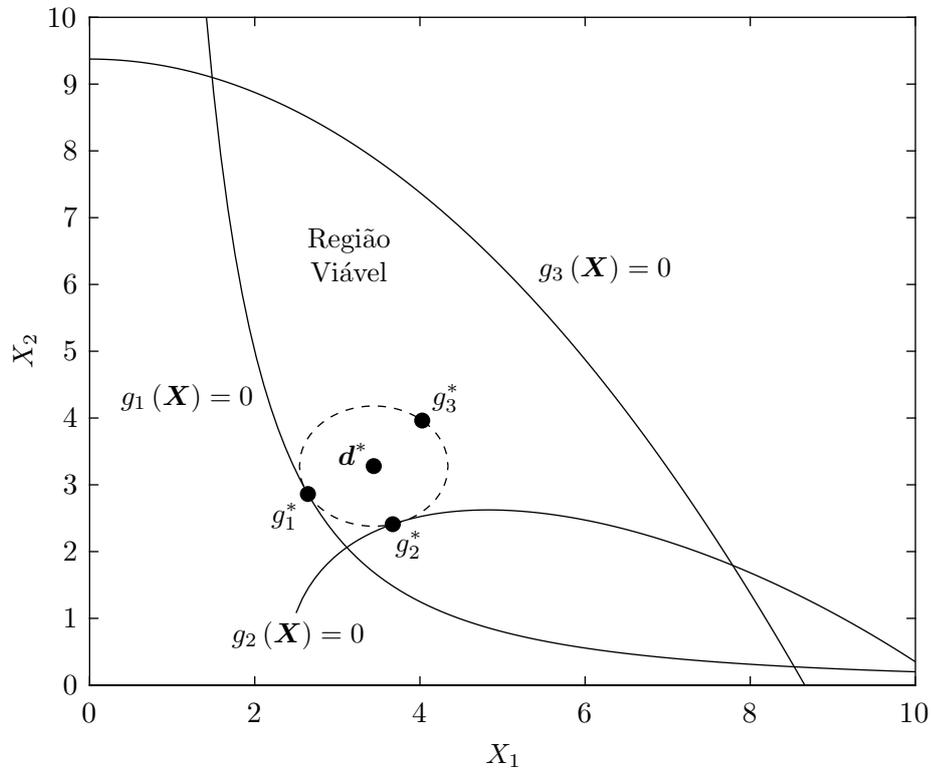
$$g_2(\mathbf{d}, \mathbf{X}) = \frac{(X_1 + X_2 - 5)^2}{30} + \frac{(X_1 - X_2 - 12)^2}{120} - 1 \quad (90b)$$

$$g_3(\mathbf{d}, \mathbf{X}) = \frac{80}{X_1^2 + 8X_2 + 5} - 1. \quad (90c)$$

Para a solução do problema definido pelas Eqs. (89)–(90), novamente é empregado o método DE, tendo sido utilizados os mesmos parâmetros de controle dos casos anteriores (os problemas determinísticos e com análise de robustez). As restrições probabilísticas são resolvidas com o método ASOSL, onde $\delta_\eta = 1$ e os parâmetros de busca local utilizando *backtracking* são $\alpha_b = 10^{-4}$ e $s_b = 0,5$. A solução do problema quando $\beta_t = 3$ equivale à confiabilidade de 99,86%, aproximadamente. A Fig. 40 mostra o resultado obtido na otimização baseada em confiabilidade. O valor ótimo alcançado para a média das variáveis aleatórias é $\mathbf{d}^* = (3,440563, 3,279963)$. O círculo tracejado, mostrado na Fig. 40, representa a região confiável, delimitada pelo índice de confiabilidade no espaço \mathbf{X} .

Os pontos g_i^* , para $i = 1, 2, 3$, representam os valores das restrições probabilísticas, avaliadas no ponto mais provável de falha, obtidos durante as respectivas execuções do método ASOSL. Note que apenas as restrições g_1 e g_2 estão ativas, com $g_1(2,6435, 2,8619) \approx 0$ e $g_2(3,6709, 2,4099) \approx 0$. No caso da terceira restrição, $g_3(4,0275, 3,9623) \approx 0,5118$, e portanto $g_3^* > 0$, o que significa que todas as restrições probabilísticas são satisfeitas. Nesta análise, fica claro que as soluções confiáveis sempre estão dispostas no centro do círculo (para problemas no plano) de raio β_t , os quais tendem a se afastar cada vez mais do mi-

Figura 40 - Otimização baseada em confiabilidade do problema dado pela Eq. (89), para $\beta_t = 3$.



Fonte: O autor, 2019

nimizador determinístico, à medida que eleva-se o valor de β_t , fornecendo soluções mais confiáveis, mas com o prejuízo do aumento do valor ótimo, assim como na otimização robusta.

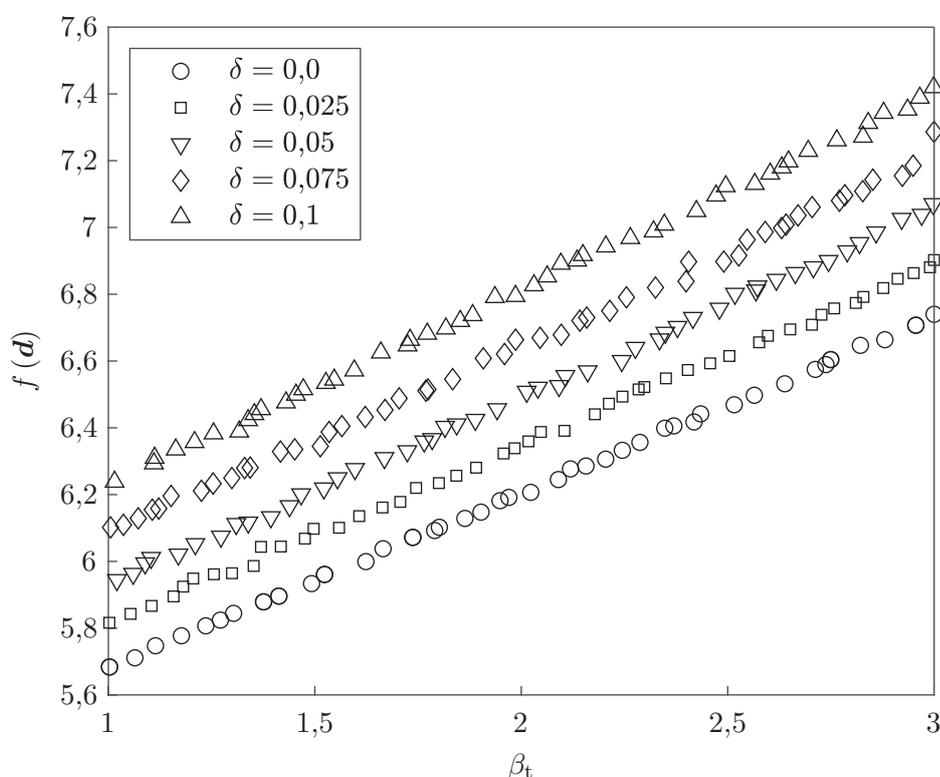
Finalmente, considere a formulação do problema multi-objetivo, como proposto na Eq. (85), levando em conta as incertezas pela análise de robustez e otimização baseada em confiabilidade. Neste cenário, a proposta é que o problema seja formulado como

$$\begin{aligned}
 & \arg \min_{\mathbf{d}} f(\mathbf{d}, \boldsymbol{\delta}) \\
 & \max \beta_t \\
 \text{Sujeito a} & P[g_i(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) \leq 0] \leq \Phi(-\beta_t), \\
 & 0 \leq d_1, d_2 \leq 10 \\
 & 1 \leq \beta_t \leq 3
 \end{aligned} \tag{91}$$

onde $i = 1, 2, 3$ e as expressões das restrições são dadas pelas Eqs. (90a)–(90c). Assim como no caso anterior, \mathbf{X} é um conjunto de variáveis aleatórias independentes com distribuição normal, onde $\mu_x = (d_1, d_2)$ e $\sigma_x = (0,3, 0,3)$. O intervalo no qual o índice de confiabilidade é definido equivale à confiabilidade entre $\Phi(1) \approx 0,8413$ e $\Phi(3) \approx 0,9987$.

Para obtenção do conjunto de soluções deste problema é empregado o algoritmo MODE, apresentado na Seção 2.4.2, com fator de amplificação $F = 0,5$, probabilidade de cruzamento $CR = 0,8$, parâmetro de redução $r = 0,9$ e número de pseudo-frentes $R = 10$. O algoritmo é executado por 500 gerações para cada nível de robustez δ , com $NP = 50$ indivíduos. No caso das restrições probabilísticas, é empregado o método ASOSL com os mesmos parâmetros do caso anterior, quando foi avaliado o problema de otimização mono-objetivo baseado em confiabilidade. As aproximações computadas para as frentes de Pareto, com diferentes níveis de robustez, são mostradas na Fig. 41.

Figura 41 - Conjuntos de valores ótimos obtidos para diferentes níveis de robustez, no problema dado pela Eq. (91).



Fonte: O autor, 2019

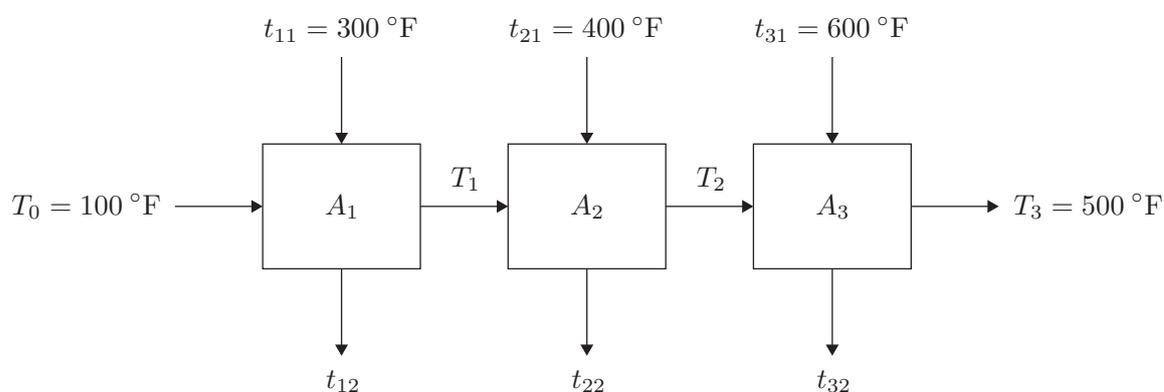
Analisando os perfis obtidos, fica clara a contribuição particular de ambas as abordagens de incerteza: maximizar a confiabilidade, em termos de β_t , representa um aumento, não necessariamente linear (como será visto mais a frente), do valor ótimo de f , à medida que o índice de confiabilidade também aumenta. No caso da análise de robustez, o perfil é deslocado como um todo, também acometendo f , em relação às curvas de Pareto obtidas quando δ é menor. Estes resultados vão ao encontro daqueles obtidos nas Figs. 39 e 40, quando os resultados calculados mostram afastamento gradual em relação ao minimizador determinístico, mas por razões distintas. Estes fatos, atrelados aos interesses econômicos do projeto, são primordiais na escolha da solução a ser implementada na prática.

7.3.2 Problemas de Engenharia Química

7.3.2.1 Projeto de Rede de Trocadores de Calor

Em um projeto de uma rede contendo três trocadores de calor posicionados em série, como mostra a Fig. 42, são conhecidas as constantes que representam as temperaturas do fluido quente que entra nos trocadores de calor, t_{11} , t_{21} e t_{31} , assim como os coeficientes gerais de transferência de calor, U_1 , U_2 e U_3 . Um fluido, com um dado fluxo volumétrico W e calor específico C_p , é aquecido da temperatura T_0 até T_3 , pela passagem através dos três trocadores de calor. T_0 é a temperatura inicial do fluido sendo aquecido, T_i é a temperatura do fluido em cada estágio de aquecimento e t_{i2} é a temperatura do fluido quente que sai dos trocadores de calor, para $i = 1, 2, 3$. Em cada trocador de calor (compreendido como um estágio do processo), a corrente fria é aquecida por um fluido quente com a mesma vazão W e calor específico C_p (AVRIEL; WILLIAMS, 1971). No contexto da otimização mono-objetivo, sem considerar incertezas, este problema foi estudado por Adjiman et al. (1998), Andrei (2013) e outros.

Figura 42 - Rede de trocadores de calor com três estágios.



Fonte: Adaptado de Avriel e Williams (1971).

A taxa de calor transferida para o fluido frio é menor ou igual à taxa de calor cedida pela corrente quente. Existem três balanços de calor que expressam este fato, dados por $WC_p(T_i - T_{i-1}) \leq WC_p(t_{i1} - t_{i2})$, para $i = 1, 2, 3$. Equivalentemente, estes balanços podem ser descritos como

$$T_i + t_{i2} \leq t_{i1} + T_{i-1} \quad (92)$$

De maneira similar, as inequações das transferências de calor em relação a cada estágio da

rede são dadas por

$$WC_p (T_i - T_{i-1}) \leq U_i A_i (t_{i2} - T_{i-1}) , \quad (93)$$

onde os coeficientes gerais de transferência de calor $U = (120, 80, 40)$ Btu/(ft² h °F) e $WC_p = 10^5$ Btu/(h °F).

Considerando as restrições dadas pelas Eqs. (92) e (93), o objetivo é minimizar a área total dos trocadores de calor, dada por $A_T(A_1, A_2, A_3) = A_1 + A_2 + A_3$. Portanto, o problema de otimização determinístico da rede de trocadores de calor é formulado como

$$\begin{aligned} & \arg \min_{A_1, A_2, A_3} A_T(A_1, A_2, A_3) \\ \text{Sujeito a} & \quad \frac{1}{400} (T_1 + t_{12}) - 1 \leq 0 \\ & \quad \frac{1}{400} (T_2 + t_{22} - T_1) - 1 \leq 0 \\ & \quad \frac{1}{100} (t_{32} - T_2) - 1 \leq 0 \\ & \quad A_1 (100 - t_{12}) + \frac{10^5}{120} T_1 - \frac{10^7}{120} \leq 0 \\ & \quad A_2 (T_1 - t_{22}) - 1250 T_1 + 1250 T_2 \leq 0 \\ & \quad A_3 (T_2 - t_{32}) - 2500 T_2 + 1250000 \leq 0 \\ & \quad 100 \leq A_1 \leq 10000 \\ & \quad 1000 \leq A_2, A_3 \leq 10000 \\ & \quad 10 \leq T_1, T_2, t_{12}, t_{22}, t_{32} \leq 1000 \end{aligned}$$

A área total mínima é $A_T(579,31, 1359,97, 5109,97) \approx 7049,25$ ft², obtida com os parâmetros $(T_1, T_2, t_{12}, t_{22}, t_{32}) = (182,02, 295,6, 217,98, 286,42, 395,60)$ (ADJIMAN et al., 1998).

Originalmente, o problema tem oito variáveis de decisão. Entretanto, Angira e Babu (2002) mostram que as restrições de desigualdade, referentes à Eq. (92), podem ser substituídas por restrições de igualdade, sem prejuízo ao modelo. Reescrevendo os termos das temperaturas do fluido quente que entra nos trocadores de calor, t_{11} , t_{21} e t_{31} , em termos das temperaturas T_1 e T_2 , e substituindo as expressões resultantes nas restrições definidas pela Eq. (93), o problema passa a ter cinco variáveis de decisão, a saber, a área dos trocadores de calor que compõem a rede, A_1 , A_2 e A_3 , além das temperaturas T_1 e T_2 . Então, as novas restrições do problema são as seguintes:

$$A_1 T_1 + \frac{10^5}{120} T_1 - 300 A_1 - \frac{10^7}{120} \leq 0 \quad (94a)$$

$$T_2 A_2 - 400 A_2 - 1250 (T_1 - T_2) \leq 0 \quad (94b)$$

$$1250000 - 2500T_2 - 100A_3 \leq 0. \quad (94c)$$

Para a otimização baseada em confiabilidade, Lobato et al. (2019) sugerem que as incertezas sejam associadas aos termos constantes das inequações, de forma que sejam manipulados como variáveis aleatórias. As inequações do problema determinístico da rede de trocadores de calor, cujas restrições equivalentes às do modelo original são dada pela Eq. (94), podem ser transformadas em restrições probabilísticas. Neste contexto, seja $\mathbf{d} = (A_1, A_2, A_3, T_1, T_2)$ o vetor de variáveis de decisão determinísticas e $\mathbf{X} = (X_1, \dots, X_8)$ o vetor de variáveis aleatórias. Para um dado índice de confiabilidade alvo β_t , o modelo de rede de trocadores de calor considerando a confiabilidade das soluções pode ser modelado como

$$\begin{aligned} & \arg \min_{d_1, d_2, d_3} A_T(d_1, d_2, d_3) \\ & \max \beta_t \\ \text{Sujeito a} & \quad P[g_i(\mathbf{d}, \mathbf{X}) \leq 0] \leq \Phi(-\beta_t) \\ & \quad 100 \leq d_1 \leq 10000 \quad , \\ & \quad 1000 \leq d_2, d_3 \leq 10000 \\ & \quad 10 \leq d_4, d_5 \leq 1000 \\ & \quad 0,1 \leq \beta_t \leq 3 \end{aligned}$$

para $i = 1, 2, 3$, onde as funções de desempenho são dadas pelas seguintes expressões:

$$g_1(\mathbf{d}, \mathbf{X}) = d_1 d_4 + X_1 d_4 - X_2 d_1 - X_3$$

$$g_2(\mathbf{d}, \mathbf{X}) = d_5 d_2 - X_4 d_2 - X_5 (d_4 - d_5)$$

$$g_3(\mathbf{d}, \mathbf{X}) = X_6 - X_7 d_5 - X_8 d_3 .$$

Neste modelo, \mathbf{X} é um vetor de variáveis aleatórias independentes, que seguem a distribuição normal, cujas médias e desvios-padrão são listados na Tabela 9.

Tabela 9 - Propriedades estatísticas das variáveis aleatórias do problema da rede de trocadores de calor.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Média	$\frac{2500}{3}$	300	$\frac{250000}{3}$	400	1250	1250000	2500	100
Desvio-padrão	$\frac{125}{3}$	15	$\frac{12500}{3}$	20	$\frac{125}{2}$	62500	125	5

Fonte: O autor, 2019.

No caso da análise de robustez, as incertezas são consideradas, neste problema, como variações geométricas referentes às áreas dos trocadores de calor (A_1 , A_2 e A_3) e diferenças relacionadas às temperaturas aferidas durante a passagem do fluido entre os trocadores de calor (T_1 e T_2). Assim, o vetor de incertezas δ determina a faixa de variação permissível de cada uma das componentes do vetor \mathbf{d} . Portanto, o problema de otimização multi-objetivo do projeto de rede de trocadores de calor considerando robustez e confiabilidade pode ser formulado como

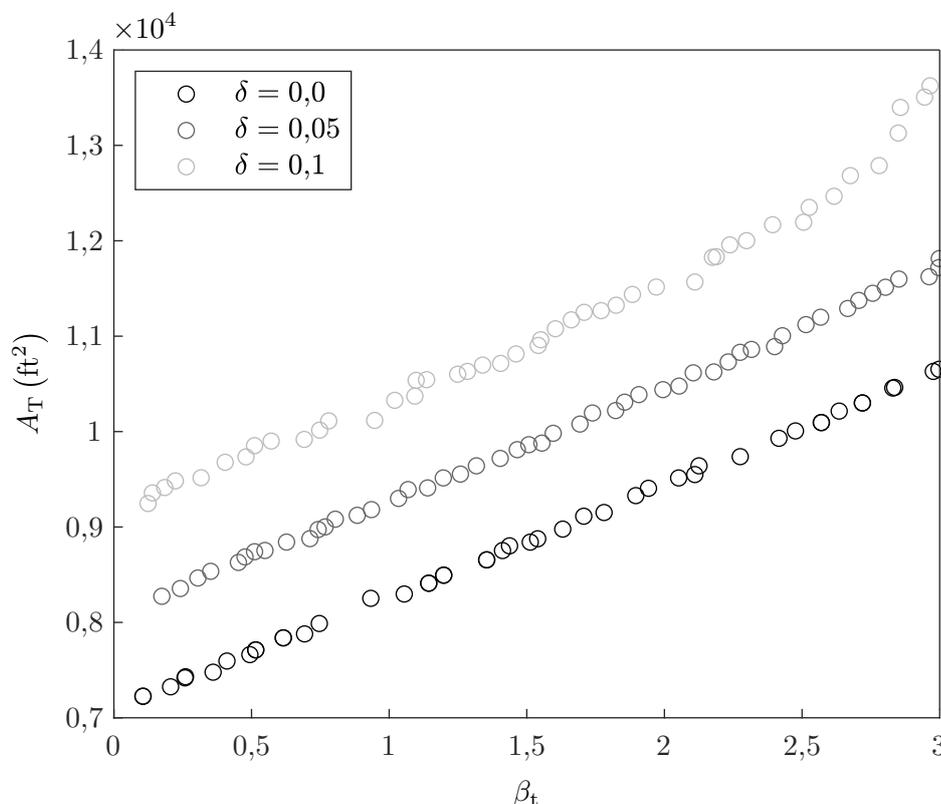
$$\begin{aligned}
 & \arg \min_{\mathbf{d}} \quad A_T(\mathbf{d}, \delta) \\
 & \max \quad \beta_t \\
 \text{Sujeito a} \quad & P[g_i(\mathbf{d}, \mathbf{X}, \delta) \leq 0] \leq \Phi(-\beta_t) \\
 & 100 \leq d_1 \leq 10000 \\
 & 1000 \leq d_2, d_3 \leq 10000 \\
 & 10 \leq d_4, d_5 \leq 1000 \\
 & 0,1 \leq \beta_t \leq 3
 \end{aligned} \tag{96}$$

Para a solução deste problema é empregado o algoritmo MODE, apresentado na Seção 2.4.2, com fator de amplificação $F = 0,5$, probabilidade de cruzamento $CR = 0,8$, parâmetro de redução $r = 0,9$ e número de pseudo-frentes $R = 10$. O algoritmo é executado por 500 gerações, com $NP = 50$ indivíduos. No caso do método ASOSL, as execuções são realizadas com $\delta_\eta = 1$, além de $\alpha_b = 10^{-4}$ e $s_b = 0,5$ para a etapa de busca local utilizando *backtracking*. Na Fig. 43 são mostradas as soluções obtidas, considerando-se diferentes níveis de robustez, os quais são calculados pela estratégia de Média Efetiva, mostrada na Seção 3.3.1.2. São empregadas $H = 50$ amostras aleatórias, geradas pela estratégia de Hipercubo Latino. Por sua vez, a variação de β_t no intervalo adotado garante resultados cuja confiabilidade variam entre $\Phi(0,1) \approx 0,5398$ e $\Phi(3) \approx 0,9987$.

Inicialmente, pode-se notar que o resultado do problema determinístico, obtido no caso em que $\delta = 0$, equivale ao mais baixo nível de confiabilidade avaliado, na formulação proposta. Em geral, problemas deste tipo exprimem a área total do trocador de calor como uma função do seu custo de desenvolvimento e operação. Sendo assim, valores ótimos cuja confiabilidade é maximizada tendem a apresentar maior custo (como consequência do inevitável aumento da área total da rede de trocadores de calor), em detrimento ao resultado determinístico. Por sua vez, quando se eleva o nível de robustez desejado para os resultados, através do aumento de δ , o conjunto de soluções dominantes é deslocado, manifestando aumento ainda mais acentuado em relação à área total.

O perfil aproximadamente linear obtido para cada conjunto de soluções dominantes, no intervalo em que β_t é avaliado, pode ser decorrente do comportamento linear do problema, tanto em relação à função-objetivo quanto às restrições probabilísticas. Entre-

Figura 43 - Soluções do problema de otimização multi-objetivo de uma rede de trocadores de calor com incertezas, definido pela Eq. (96), para diferentes níveis de robustez.



Fonte: O autor, 2019

tanto, o conjunto de soluções nunca se cruza. Este fato corrobora a solidez da formulação do problema, assim como os métodos numéricos utilizados para obtenção dos resultados, uma vez que conjuntos de soluções com diferentes níveis de robustez não devem ter predominância entre si, fazendo com que a escolha do otimizador a ser implementado na prática seja determinada considerando-se níveis de robustez e confiabilidade independentes.

A análise dos valores ótimos mostrados na Fig. 43 pode ser complementada com os resultados listados na Tabela 10, onde são mostrados os valores das variáveis de decisão referentes a três minimizadores distintos, para cada valor de δ . Os valores escolhidos são pertinentes a soluções caracterizadas pelo compromisso extremo em relação a cada um dos objetivos, isto é, as soluções em que a área total é mínima ou o índice de confiabilidade é máximo, além da solução com compromisso intermediário em relação a ambos.

Nota-se que, para um dado nível de robustez, não é possível verificar qualquer relação linear óbvia no que diz respeito ao aumento da área total da rede de trocadores de calor, em detrimento ao aumento praticamente linear observado na Fig. 43. O mesmo vale para as temperaturas T_1 e T_2 . Por outro lado, analisando cada um dos minimizadores da Tabela 10, fica claro que a maior contribuição para o aquecimento do fluido até $500\text{ }^\circ\text{F}$ é

dada no terceiro estágio, o que justifica os valores de A_3 em todos os resultados reportados, maiores do que os demais trocadores de calor da mesma rede.

Tabela 10 - Resultados ótimos do problema de rede de trocadores de calor, considerando diferentes níveis de compromisso entre os objetivos. Para cada conjunto de resultados, os valores são dispostos em relação a $\mathbf{d} = (A_1, A_2, A_3, T_1, T_2)^T$.

Robustez	Compromisso Extremo (A_T)	Compromisso Intermediário	Compromisso Extremo (β_t)
$\delta = 0,0$	725,64 ft ²	595,07 ft ²	551,11 ft ²
	1270,47 ft ²	1260,02 ft ²	1279,83 ft ²
	5229,70 ft ²	6986,06 ft ²	8822,1 ft ²
	190,93 °F	169,00 °F	153,48 °F
	295,11 °F	269,24 °F	246,37 °F
$\delta = 0,05$	467,80 ft ²	431,84 ft ²	368,17 ft ²
	1482,09 ft ²	1380,90 ft ²	1497,34 ft ²
	6323,53 ft ²	8065,84 ft ²	9948,55 ft ²
	155,83 °F	146,84 °F	129,49 °F
	270,75 °F	247,02 °F	227,35 °F
$\delta = 0,1$	141,83 ft ²	464,59 ft ²	458,75 ft ²
	1643,37 ft ²	1756,62 ft ²	3178,67 ft ²
	7461,87 ft ²	8857,41 ft ²	9989,32 ft ²
	113,95 °F	135,72 °F	113,86 °F
	242,45 °F	244,91 °F	246,04 °F

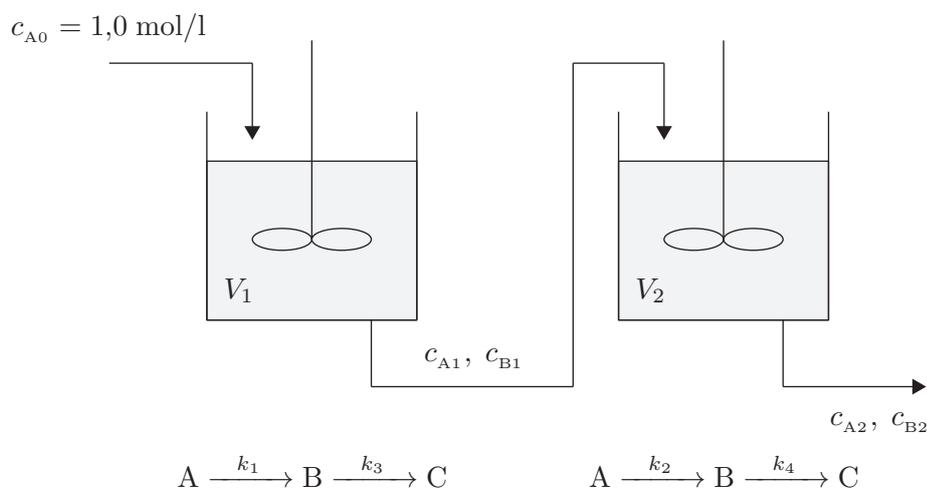
Fonte: O autor, 2019.

7.3.2.2 Projeto de Rede de Reatores

O problema do projeto de rede de reatores envolve dois reatores tanque agitado contínuo, também chamados de CSTR (*continuous stirred tank reactor*), onde ocorre a reação $A \rightarrow B \rightarrow C$, sendo A, B e C as componentes do sistema. Em ambas as reações, assume-se uma cinética de primeira ordem e os reatores são considerados perfeitamente agitados, sem variação do volume reacional. O objetivo é projetar o sistema de modo que a concentração de B na corrente de saída do segundo reator (c_{B_2}) seja maximizada, e o custo do investimento não exceda um dado limite superior (SMITH, 1996). Na Fig. 44, a estrutura essencial do projeto de rede de reatores é ilustrada.

Sejam V_1 e V_2 os tempos de residência para o primeiro e segundo reatores, respectivamente. Considere, também, $k_1 = 0,09755988 \text{ s}^{-1}$, $k_2 = 0,09658428 \text{ s}^{-1}$, $k_3 = 0,0391908 \text{ s}^{-1}$

Figura 44 - Esquema do projeto de rede de reatores.



Fonte: Adaptado de Smith (1996).

e $k_4 = 0,03527172 \text{ s}^{-1}$ as constantes das taxas de reação, cuja relação está estabelecida na Fig. 44. Neste modelo, supõe-se que o custo de capital de um reator seja proporcional à raiz quadrada de seu tempo de residência. O problema de projeto de rede de reatores, cujo propósito é maximizar $f(c_{B2}) = c_{B2}$, é formulado como

$$\begin{aligned}
 & \arg \max_{c_{B2}} f(c_{B2}) \\
 \text{Sujeito a} \quad & c_{A1} - c_{A0} + k_1 c_{A1} V_1 = 0 \\
 & c_{A2} - c_{A1} + k_2 c_{A2} V_2 = 0 \\
 & c_{B1} + c_{A1} - c_{A0} + k_3 c_{B1} V_1 = 0 \\
 & c_{B2} - c_{B1} + c_{A2} - c_{A1} + k_4 c_{B2} V_2 = 0 \\
 & \sqrt{V_1} + \sqrt{V_2} \leq 4 \\
 & 0 \leq c_{A1}, c_{A2}, c_{B1}, c_{B2} \leq 1 \\
 & 0 \leq V_1, V_2 \leq 16
 \end{aligned} \tag{97}$$

Este modelo foi proposto por Manousiouthakis e Sourlas (1992) e, no contexto da otimização mono-objetivo, também foi estudado por Ryoo e Sahinidis (1995), Angira e Babu (2002), Kheawhom (2010), Sharma et al. (2013) e Dong et al. (2014). Todos os autores relatam um ponto em comum: o problema é considerado difícil pois possui dois ótimos locais muito próximos ao ótimo global, como pode ser visto na Tabela 11. Note que as duas soluções locais utilizam apenas um dos dois reatores (V_1 e V_2 alternam-se com valores nulos em ambos os ótimos locais), enquanto a solução global utiliza ambos. Além disso, o problema contém um conjunto de restrições não-convexas, que podem ser um complicador

para alguns métodos de otimização.

Tabela 11 - Soluções ótimas do problema de projeto de rede de reatores, com a indicação do tipo de solução ótima (local ou global) de cada ponto.

c_{A1} (mol/l)	c_{A2} (mol/l)	c_{B1} (mol/l)	c_{B2} (mol/l)	V_1 (s)	V_2 (s)	f	Tipo
0,390	0,390	0,375	0,375	16	0	0,375	Local
1	0,393	0	0,388	0	16	0,388	Local
0,771	0,517	0,204	0,389	3,037	5,096	0,389	Global

Fonte: Adaptado de Ryoo e Sahinidis (1995).

No cenário da otimização com incertezas, mais especificamente na otimização baseada em confiabilidade, Lobato et al. (2019) propuseram que potenciais incertezas poderiam ser associadas às constantes k_1 a k_4 , uma vez que são estimadas usando a dependência de temperatura de Arrhenius, cujo procedimento está sujeito a variações e incertezas. Os autores sugerem uma simplificação para o problema definido na Eq. (97), através da eliminação das restrições de igualdade. Inicialmente, o termo c_{B2} é explicitado na última restrição de igualdade da Eq. (97). Em seguida, as demais restrições de igualdade são escritas apenas em função de c_{A1} e c_{A2} . As expressões resultantes são substituídas nos termos correspondentes, do lado direito da equação de c_{B2} . Com isso, a função-objetivo pode ser escrita como

$$f(\mathbf{d}, \mathbf{X}) = \frac{X_2 d_2 (X_1 d_1 (1 - d_1) - (X_1 d_1 + X_3 (1 - d_1)) (d_2 - d_1))}{(X_1 d_1 + X_3 (1 - d_1)) (X_2 d_2 + X_4 (d_1 - d_2))}. \quad (98)$$

Desta forma, o problema passa a ter uma única restrição de desigualdade que, em geral, é mais fácil de ser manipulada, e a função-objetivo passa a ser definida por duas variáveis de decisão, c_{A1} e c_{A2} .

No caso da restrição de desigualdade da Eq. (97), a simplificação proposta por Lobato et al. (2019) segue o mesmo procedimento adotado anteriormente. Explicitando V_1 e V_2 nas duas primeiras restrições de igualdade, torna-se possível escrever a restrição de desigualdade também em função de c_{A1} e c_{A2} . Assim, a restrição de desigualdade equivalente é dada por

$$\sqrt{\frac{1 - d_1}{X_1 d_1}} + \sqrt{\frac{d_1 - d_2}{X_2 d_2}} - 4 \leq 0, \quad (99)$$

onde \mathbf{d} é o vetor de variáveis de decisão determinísticas, dado por $\mathbf{d} = (c_{A1}, c_{A2})$, e o vetor $\mathbf{X} = (X_1, X_2, X_3, X_4)$ é um conjunto de variáveis aleatórias independentes, com distribuição normal. O problema de otimização multi-objetivo da rede de reatores, considerando a

confiabilidade do sistema, é definido como

$$\begin{aligned} \arg \max_{\mathbf{d}} \quad & F = (f(\mathbf{d}, \mathbf{X}), \beta_t) \\ \text{Sujeito a} \quad & P[g_1(\mathbf{d}, \mathbf{X}) \leq 0] \leq \Phi(-\beta_t), \\ & 10^{-5} \leq d_1, d_2 \leq 1 \\ & 0,1 \leq \beta_t \leq 5 \end{aligned} \quad (100)$$

em que o objetivo $f(\mathbf{d}, \mathbf{X})$ é definido pela Eq. (98) e $g_1(\mathbf{d}, \mathbf{X})$ é dado pelo lado esquerdo da Eq. (99). As médias e desvios-padrão de \mathbf{X} são mostrados na Tabela 12.

Tabela 12 - Propriedades estatísticas das variáveis aleatórias do problema da rede de reatores.

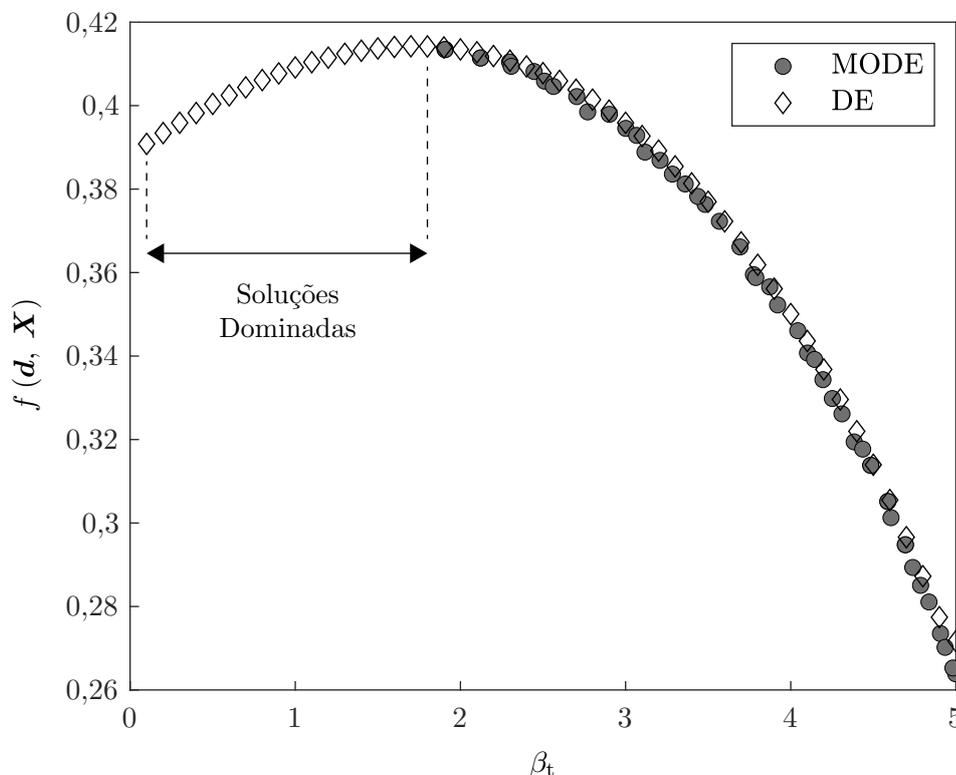
	X_1	X_2	X_3	X_4
Média	0,09755988	0,09658428	0,0391908	0,03527172
Desvio-padrão	0,01463398	0,01448764	0,00587862	0,00529076

Fonte: Reproduzido de Lobato et al. (2019).

O problema analisado é resolvido através de duas abordagens distintas. Na primeira, o problema da Eq. (100) é otimizado utilizando o MODE. Assim como nos problemas anteriores, é adotado fator de amplificação $F = 0,5$, probabilidade de cruzamento $CR = 0,8$, parâmetro de redução $r = 0,9$ e número de pseudo-frentes $R = 10$. O algoritmo também é executado por 500 gerações, com $NP = 50$ indivíduos. Na segunda abordagem, o problema da Eq. (100) é convertido em um problema mono-objetivo, cujo único objetivo é dado por $f(\mathbf{d}, \mathbf{X})$. Assim, são realizadas 50 execuções independentes do algoritmo DE, variando-se o índice de confiabilidade, o qual é definido por elementos igualmente espaçados no intervalo $0,1 \leq \beta_t \leq 5$. Em cada uma das execuções, adota-se $NP = 50$ indivíduos, fator de amplificação $F = 0,5$ e probabilidade de cruzamento $CR = 0,8$ para a DE, com número máximo de gerações igual a 100. Em ambas as abordagens, as restrições probabilísticas são resolvidas com o método ASOSL, onde $\delta_\eta = 1$, $\alpha_b = 10^{-4}$ e $s_b = 0,5$. A Fig. 45 apresenta o conjunto de resultados obtidos por ambos os métodos no mesmo conjunto de eixos.

Sob esta perspectiva, fica clara a relevância de se modelar o problema de otimização baseada em confiabilidade em um contexto multi-objetivo: no intervalo analisado para o índice de confiabilidade, a DE obtém diversas soluções dominadas, como mostra a Fig. 45, pois critérios de dominância não são levados em conta durante o seu procedimento de otimização (isso vale para a maioria das metaheurísticas no cenário mono-objetivo). Isso ocorre pois as execuções do algoritmo são independentes. Portanto, a DE não é capaz de avaliar se o índice de confiabilidade é adequado, mas apenas maximizar o valor do objetivo, para

Figura 45 - Conjuntos de resultados ótimos obtidos pelo MODE para o problema dado pela Eq. (100), e pela DE para um problema mono-objetivo equivalente, variando-se o índice de confiabilidade a cada execução.



Fonte: O autor, 2019

um dado β_t fixado. Todas as soluções obtidas pela DE quando $\beta_t < 1,8$, aproximadamente, possuem outra com objetivo praticamente equivalente, mas com índice de confiabilidade superior.

Por sua vez, a formulação multi-objetivo, como mostrada na Eq. (100), propicia que a metaheurística obtenha um conjunto de resultados que seja composto apenas por soluções dominantes, mesmo que o intervalo adotado para β_t , neste caso, possa produzir candidatos a solução cuja confiabilidade é inadequada. Naturalmente, este contraponto também se reflete no custo computacional de se obter um conjunto de soluções através de cada uma das abordagens. O problema mono-objetivo deve ser resolvido tantas vezes quanto forem a quantidade de soluções desejadas, enquanto a abordagem com múltiplos objetivos fornece um conjunto de soluções dominantes com uma única execução. Evidentemente, quantificar o custo computacional com exatidão depende da estratégia evolutiva adotada por cada técnica.

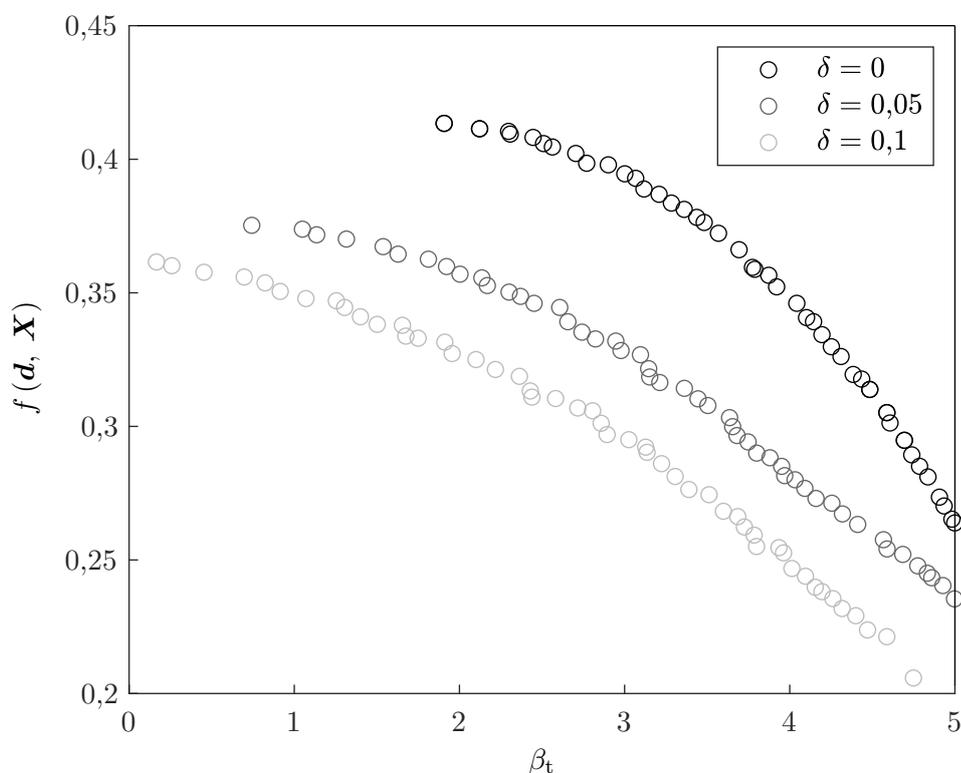
Agora, considere o problema dado pela Eq. (100) e a incorporação da análise de robustez. A formulação do problema de otimização multi-objetivo de um projeto de rede de

reatores, considerando robustez e confiabilidade, é dada por

$$\begin{aligned} \arg \max_{\mathbf{d}} \quad & F = (f(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}), \beta_t) \\ \text{Sujeito a} \quad & P[g_1(\mathbf{d}, \mathbf{X}, \boldsymbol{\delta}) \leq 0] \leq \Phi(-\beta_t) \\ & 10^{-5} \leq d_1, d_2 \leq 1 \\ & 0,1 \leq \beta_t \leq 5 \end{aligned} \quad (101)$$

Neste caso, propõe-se avaliar a robustez do problema analisando a influência da sensibilidade das correntes c_{A1} e c_{A2} na função-objetivo. Para esta tarefa, é empregada a técnica de Média Efetiva, com $H = 50$ amostras, cujas incertezas são estabelecidas em 5% e 10%. A Fig. 46 mostra os perfis obtidos para ambos os valores de δ , além do caso em que não se considera a sensibilidade do vetor \mathbf{d} , isto é, quando $\delta = 0$. Novamente, os resultados foram obtidos com os algoritmos MODE e ASOSL, utilizando os mesmos parâmetros do caso anterior.

Figura 46 - Conjunto de resultados do problema de otimização multi-objetivo de uma rede de reatores com incertezas, dado pela Eq. (101), para diferentes níveis de robustez.



Fonte: O autor, 2019

Assim como nos casos anteriores, nota-se um deslocamento dos perfis à medida que o valor de δ é aumentado. Contudo, diferentemente dos demais problemas analisados, este

modelo apresenta duas particularidades. Primeiro, considere um conjunto de curvas hipotéticas que melhor aproxime os resultados computados, para cada δ . Em cada uma delas, nota-se uma atenuação do grau de curvatura, que coincide com o aumento do nível de sensibilidade. Naturalmente, para um dado índice de confiabilidade, isso se reflete em uma redução não-linear do valor do objetivo. Como exemplo, escolha $\beta_t = 3$. O declínio do valor de f é mais acentuado entre os valores calculados quando $\delta = 0$ e $\delta = 0,05$ do que para $\delta = 0,05$ e $\delta = 0,1$. Evidentemente, este fato deve ser levado em conta durante a escolha do resultado a ser implementado.

A segunda particularidade diz respeito à influência do grau de robustez na confiabilidade dos resultados. O aumento de δ tende a fornecer valores de objetivo com β_t mais baixo. Contudo, este aumento também pode limitar o nível máximo possível de confiabilidade dos resultados. Este fato fica claro para o conjunto de valores ótimos calculados quando $\delta = 0,1$, onde o resultado cuja confiabilidade é maximizada tem $\beta_t = 4,749$. Neste caso, pode-se dizer que robustez e confiabilidade, em termos de δ e β_t , respectivamente, também são conflitantes entre si.

Outra questão importante, concernente aos resultados da Fig. 46 e também observado em outros casos, como na Fig. 43, diz respeito à “suavidade” dos conjuntos de resultados que, neste caso, refere-se à condição dos valores ótimos estarem dispostos sobre uma curva contínua hipotética, e não à característica de existência de derivadas. Esta peculiaridade pode ser mais explorada com o ajuste dos conjuntos de resultados, obtidos para cada valor de δ , por um polinômio de grau dois. Assim, é possível quantificar os desvios de cada um dos valores ótimos em relação à curva ajustada e estimar o erro cometido e a dispersão do conjunto de pontos. A Tabela 13 apresenta a soma dos quadrados dos resíduos (SQR), os coeficientes de determinação simples (R^2) e ajustado (R_a^2), e a raiz do erro médio quadrático (RMS) (para mais detalhes, veja Gujarati e Porter (2008)) das curvas ajustadas em relação aos resultados computados para cada nível de robustez.

Tabela 13 - Medidas de qualidade da curva ajustada em relação à capacidade de estimação dos valores ótimos computados para cada nível de robustez na Fig. 46.

	SQR	R^2	R_a^2	RMS
$\delta = 0$	$9,483 \times 10^{-5}$	0,9991	0,9991	$1,452 \times 10^{-3}$
$\delta = 0,05$	$2,851 \times 10^{-4}$	0,9969	0,9968	$2,463 \times 10^{-3}$
$\delta = 0,1$	$3,828 \times 10^{-3}$	0,9604	0,9596	$8,93 \times 10^{-3}$

Fonte: O autor, 2019.

Analisando as estatísticas dos ajustes, de fato há uma tendência de aumento da dispersão do conjunto de pontos quando o valor de δ é aumentado. Essa predisposição pode ser

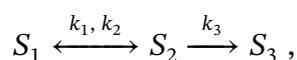
decorrente da aleatoriedade inserida pela análise de robustez, através da técnica de Média Efetiva. O aumento de δ reflete na extensão do intervalo, na vizinhança de um candidato a solução, onde as amostras aleatórias são geradas. Portanto, o valor médio do objetivo, calculado por este conjunto de amostras aleatórias, influencia na dispersão dos elementos que compõem o perfil obtido pela metaheurística. É possível que este efeito possa ser atenuado por meio do aumento da quantidade de amostras H . Entretanto, esta alternativa acarreta o aumento exponencial do custo computacional para obtenção dos resultados. Por sua vez, é improvável que este comportamento seja causado pela convergência prematura do método, uma vez que cada conjunto de resultados é produto da execução do método por 500 gerações.

Em vista desse cenário, é pertinente obter os resultados do problema da Eq. (101), utilizando as demais abordagens de análise de robustez mostradas no Cap. 3, a fim de verificar o seu comportamento em relação aos resultados da Fig. 46. A Fig. 47 apresenta os conjuntos de valores ótimos, obtidos com as técnicas de Diferença Normalizada e de Penalização, apresentadas nas Seções 3.3.1.3 e 3.3.1.4, respectivamente, para níveis de robustez definidos em 5% e 10%, além do conjunto ótimo do problema onde não é considerada a análise de robustez. Em todas as execuções, os resultados foram computados com os métodos MODE e ASOSL, utilizando os mesmos valores para parâmetros de controle adotados na obtenção dos resultados da Fig. 46.

Inicialmente, nota-se que ambas as abordagens não são capazes de alcançar índices de confiabilidade próximos a $\beta_t = 5$ quando se define o nível de robustez em 10%, assim como ocorre na estratégia de Média Efetiva. Comparando ambas, fica claro que a técnica de Penalização apresenta resultados mais conservadores, promovendo penalização mais acentuada, frente ao aumento dos níveis de robustez, em relação às demais técnicas. Por sua vez, a técnica de Diferença Normalizada apresenta resultados muito próximos àqueles alcançados utilizando a Média Efetiva. Em relação à dispersão dos resultados, fica claro que não se trata de uma particularidade da técnica de Média Efetiva, uma vez que esta característica está presente em todos os perfis da Fig. 47.

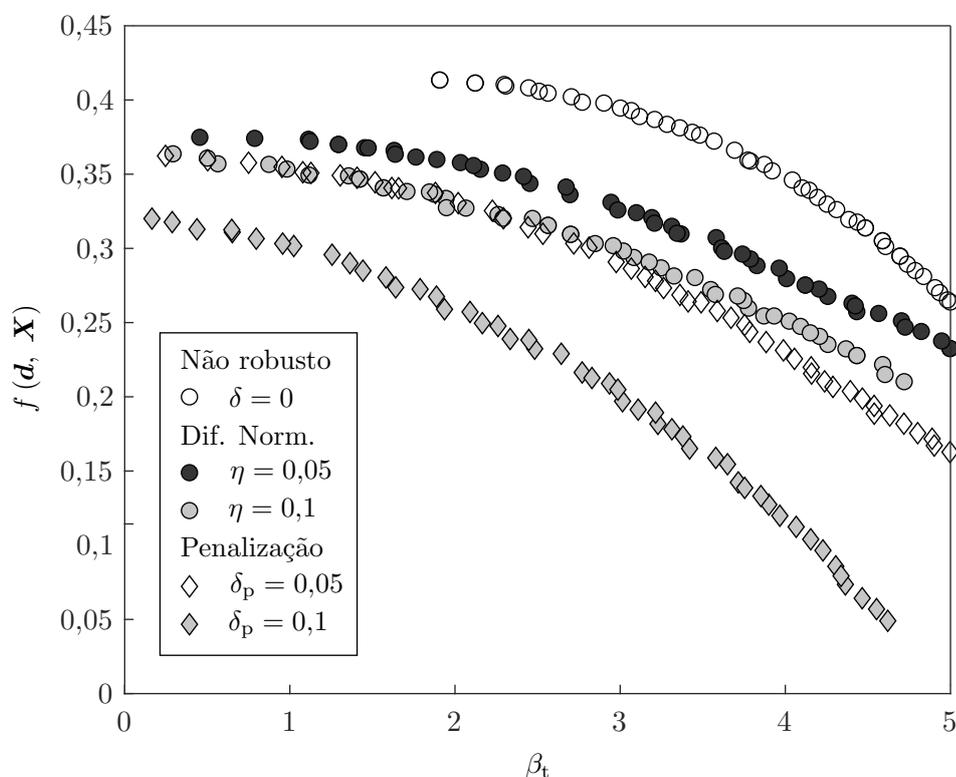
7.3.2.3 Problema de Mistura de Catalisadores

Neste problema clássico, o controle da mistura ideal de dois catalisadores ao longo do comprimento de um reator tubular deve ser determinado. O problema envolve as reações



onde k_1 e k_2 são as constantes da taxa de reação das duas primeiras reações, em um reator em que o catalisador consiste inteiramente da substância que catalisa as reações reversíveis $S_1 \rightleftharpoons S_2$, e k_3 é a constante da taxa de reação $S_2 \rightarrow S_3$. O modelo, originalmente formu-

Figura 47 - Conjunto de resultados do problema de otimização multi-objetivo de uma rede de reatores com incertezas, dado pela Eq. (101), utilizando as técnicas de Diferença Normalizada e de Penalização, com níveis de robustez em 5% e 10%.



Fonte: O autor, 2019

lado por Gunn e Thomas (1965), tem como finalidade maximizar a produção de S_3 . Ambos os catalisadores ativam as três reações, mas as taxas de ativação são variadas para cada par de reação e agentes catalisadores. Portanto, diferentes políticas de mistura de catalisadores levarão a uma ampla variedade de composições de produtos.

O modelo do problema de mistura de catalisadores é definido como

$$\begin{aligned}
 & \arg \max \quad f(t) \\
 \text{Sujeito a} \quad & \frac{dy_1(t)}{dt} = v(10y_2(t) - y_1(t)) \\
 & \frac{dy_2(t)}{dt} = v(y_1(t) - 10y_2(t)) - (1 - v)y_2(t), \\
 & y_1(0) = 1, y_2(0) = 0, t_f = 1 \\
 & 0 \leq v \leq 1
 \end{aligned} \tag{102}$$

onde y_1 e y_2 são, respectivamente, as frações molares das substâncias S_1 e S_2 , a função-

objetivo

$$f(t) = 1 - y_1(t) - y_2(t) \quad (103)$$

e t representa o tempo de permanência das substâncias no reator, a partir do instante de entrada, até o tempo final t_f . A fração de mistura do catalisador, dada por v , é a porção do catalisador formada pela substância que catalisa a reação $S_1 \leftrightarrow S_2$, a qual pode variar ao longo da posição axial do reator, e deve ser determinada pelo otimizador (DADEBO; MCAULLEY, 1995; LOBATO et al., 2019). No contexto da otimização mono-objetivo, este problema também foi estudado por Vassiliadis (1996), Tanartkit e Biegler (1997), Bell e Sargent (2000), Irizarry (2005), Liu et al. (2013) e Liu et al. (2015).

A solução deste problema é obtida discretizando o elemento de controle (v) para transformar o problema original com flutuação do índice diferencial em um problema de otimização não-linear, com índice diferencial constante (igual a 1). Com este propósito, o tempo é discretizado e, para cada subintervalo dado por (t_k, t_{k+1}) , uma fração de mistura do catalisador é aproximada, passando a ser representada por v_k , para $k = 0, \dots, n_t - 1$, onde n_t representa o número de subintervalos de tempo. Portanto, $\mathbf{d} = (v_0, \dots, v_{n_t-1}, t_1, \dots, t_{n_t-1})$ é o vetor de variáveis determinísticas que deve ser calculado pelo otimizador.

Especificamente para o caso em que $n_t = 3$, o algoritmo de otimização deve determinar os elementos de controle v_k , para $k = 0, 1, 2$, além dos instantes de tempo intermediários, t_1 e t_2 , uma vez que $t_0 = 0$ e $t_3 = t_f = 1$ são conhecidos. Utilizando o algoritmo DE (STORN; PRICE, 1997), com $NP = 50$ indivíduos na população, fator de amplificação $F = 0,5$ e probabilidade de cruzamento $CR = 0,8$, são obtidos, após 100 gerações, os seguintes resultados: $v_0 = 1$ para $t \leq 0,1338$, $v_1 = 0,2248$ para $0,1338 < t \leq 0,7237$ e $v_2 = 0$ para $0,7237 < t \leq t_f$. Nestas condições, o valor ótimo da função-objetivo é $f = 0,048065$. Este resultado concorda com aqueles obtidos por Gunn e Thomas (1965), Vassiliadis (1996) e Lobato et al. (2011).

No contexto de incertezas, Lobato et al. (2019) sugerem que o problema multi-objetivo pode ser formulado considerando possíveis desvios em termos dos elementos de controle discretizados. Neste caso, seja $\mathbf{X} = (X_0, X_1, X_2)$ um conjunto de variáveis aleatórias independentes, onde $\mathbf{X} \sim \mathcal{N}(\mu_x, \sigma_x^2)$. Para este problema, o desvio-padrão de \mathbf{X} é definido em relação ao coeficiente de variação, $\sigma_x = 0,1\mu_x$, onde μ_x representa a média da variável aleatória. No caso da análise de robustez, δ representa o vetor de ruídos, relacionados aos elementos de controle v_k , para $k = 0, 1, 2$. Desta maneira, o vetor de variáveis determinísticas passa a ser composto apenas pelos instantes de tempo, isto é, $\mathbf{d} = (t_1, t_2)$. Seja $\mathbf{F} = (f(\mathbf{d}, \mu_x, \delta), \beta_t)$, o problema definido na Eq. (102), mas formulado como problema

de otimização multi-objetivo com incertezas, é definido como

$$\begin{aligned}
 & \arg \max_{\mathbf{d}, \mu_X} F \\
 \text{Sujeito a } & \frac{dy_1(t)}{dt} = X(10y_2(t) - y_1(t)) \\
 & \frac{dy_2(t)}{dt} = X(y_1(t) - 10y_2(t)) - (1 - X)y_2(t) \\
 & y_1(0) = 1, y_2(0) = 0 \\
 & 0 \leq d_1 \leq 0,5 \\
 & 0,5 < d_2 \leq 1 \\
 & P[g_i(\mathbf{X}, \boldsymbol{\delta}) \leq 0] \leq \Phi(-\beta_t) \\
 & 0 < \mu_{x_i} \leq 1 \\
 & 0,1 \leq \beta_t \leq 5
 \end{aligned} \tag{104}$$

onde $g_i(\mathbf{X}, \boldsymbol{\delta}) = X_i$ para $i = 1, 2, 3$.

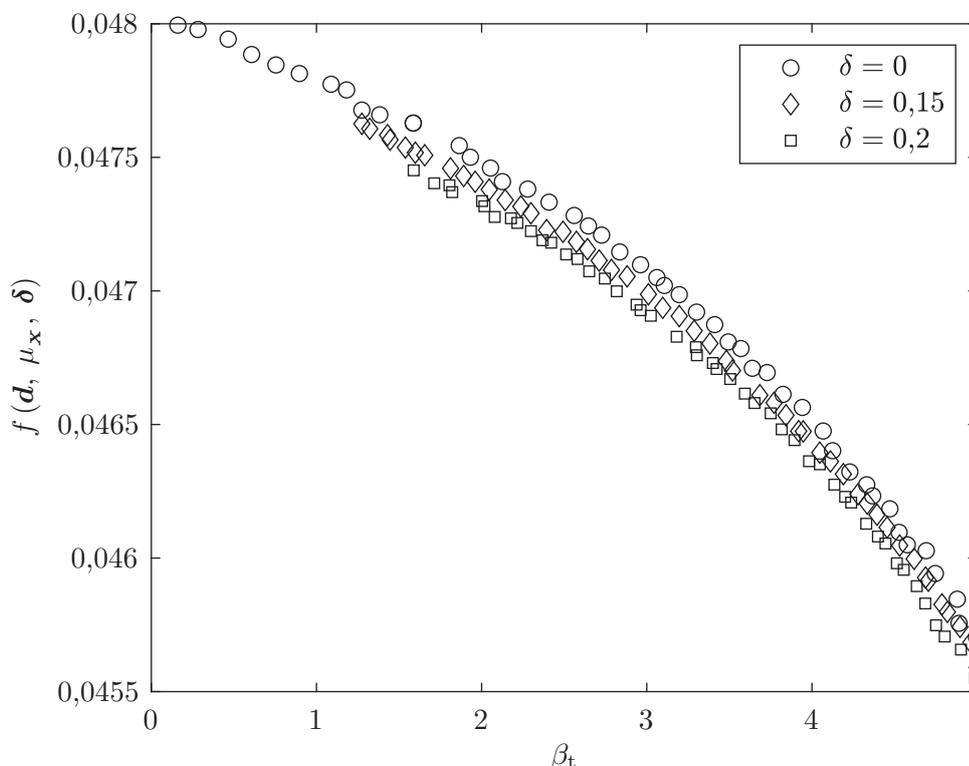
O conjunto de soluções ótimas deste problema é obtido através das etapas descritas a seguir. Inicialmente, a metaheurística gera uma população inicial de potenciais otimizadores. A cada avaliação da função-objetivo, primeiro o vetor de variáveis de decisão é perturbado, a fim de que a sua robustez seja avaliada. Para este problema, a perturbação ocorre apenas nas coordenadas referentes aos elementos de controle discretizados (neste contexto, o tempo e o índice de confiabilidade não são passíveis de perturbação). Após a perturbação, cada indivíduo da população dá origem a H novos candidatos a solução, localizados na vizinhança do vetor perturbado, os quais devem ser avaliados pela função-objetivo.

No procedimento de avaliação da função-objetivo, primeiro são analisadas as restrições do problema. No caso das restrições probabilísticas, um método de análise de confiabilidade inversa fornece os valores ótimos dos elementos de controle, assim como os respectivos valores da função de desempenho, de acordo com a distribuição de probabilidades das variáveis aleatórias. Estes dados, assim como os instantes de tempo discretizados, são usados para obter as soluções de ambas as equações diferenciais. Então, as frações molares $y_1(t)$ e $y_2(t)$ são utilizadas para avaliação do objetivo dado pela Eq. (103).

O método MODE é empregado para obtenção do conjunto de soluções deste problema, as restrições probabilísticas são avaliadas com o ASOSL e a análise de robustez é realizada pela estratégia de Média Efetiva, todos utilizando os mesmos valores de parâmetros adotados nos problemas anteriores. A Fig. 48 apresenta os conjuntos de resultados ótimos, calculados para $\delta = 0,15$ e $\delta = 0,2$, além do caso onde não se considera a robustez das variáveis de projeto.

Diferentemente das demais análises, neste caso são adotados valores elevados para os níveis de robustez. Estas escolhas são motivadas, puramente, por questões didáticas,

Figura 48 - Conjunto de resultados do problema de otimização multi-objetivo da mistura de catalisadores com incertezas, dado pela Eq. (104), para diferentes níveis de robustez.



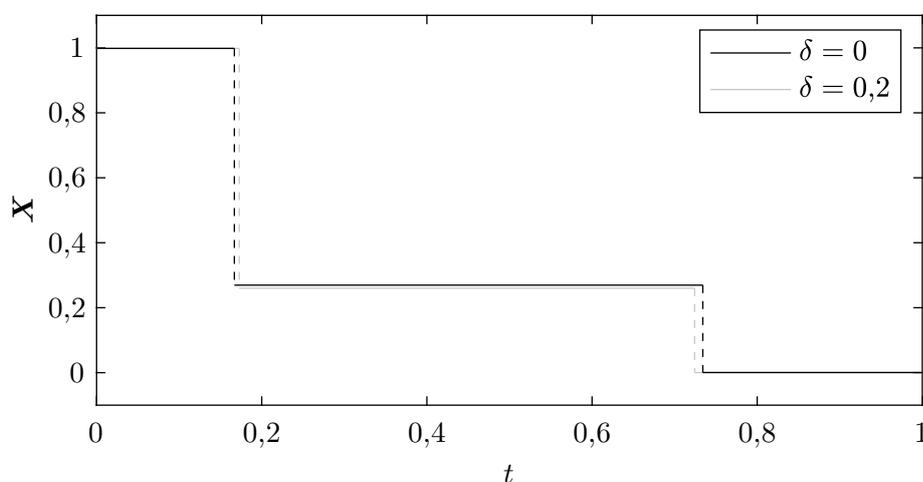
Fonte: O autor, 2019

uma vez que possibilitam verificar o comportamento dos perfis obtidos, frente ao conjunto de resultados em que não são consideradas perturbações externas. Em geral, problemas de engenharia são analisados levando em conta níveis de ruído inferiores. Contudo, quando este problema de otimização é avaliado para valores de δ mais próximos a zero, a curva de Pareto associada fica praticamente sobreposta àquela em que $\delta = 0$.

Este comportamento pode ser compreendido pela análise da Fig. 49, onde são mostrados os perfis de controle do problema dado pela Eq. (104) quando $\delta = 0$ e $\delta = 0,2$, com índices de confiabilidade aproximadamente iguais ($\beta_t \approx 1,6$). No caso em que $\delta = 0$, tem-se $x_0 = 0,9984$ para $t \leq 0,1669$, $x_1 = 0,2695$ para $0,1669 < t \leq 0,7341$ e $x_2 = 0,0004$ para $0,7341 < t \leq t_f$. Quando $\delta = 0,2$, os resultados obtidos são $x_0 = 0,9982$ para $t \leq 0,1728$, $x_1 = 0,2601$ para $0,1728 < t \leq 0,7242$ e $x_2 = 0,0002$ para $0,7242 < t \leq t_f$. Para estas variáveis de decisão, os respectivos objetivos valem $f = 0,0476$ e $f = 0,0474$.

Em ambos os casos, tem-se que $x_0 \approx 1$ e $x_2 \approx 0$. Assim, o valor do objetivo é praticamente determinado por x_1 e pelos tempos de ocorrência, t_1 e t_2 . Com isso, um desvio considerável do parâmetro de controle x_1 , em relação aos casos em que $\delta = 0$ e $\delta = 0,2$, poderia representar grande diferença entre seus objetivos, haja vista que os tempos t_1 e t_2

Figura 49 - Perfis de controle obtidos para dois níveis de robustez distintos, quando $\beta_t \approx 1,6$.



Fonte: O autor, 2019

também são muito próximos, em ambos os casos. Contudo, note que os perfis de controle obtidos nos casos analisados na Fig. 49 são muito semelhantes. Inclusive, os resultados obtidos assemelham-se, até mesmo, à solução do problema determinístico, mostrada anteriormente. Isso significa que o problema, formulado como apresentado na Eq. (104), possui soluções em regiões pouco sensíveis a perturbações externas, o que justifica a ligeira diferença das curvas de Pareto obtidas na Fig. 48, mesmo com excessiva variação do nível de robustez.

7.3.2.4 Custo Anualizado Total de um Trocador de Calor de Casco e Tubo

Os trocadores de calor de casco e tubo consistem em um conjunto de tubos redondos montados em um invólucro cilíndrico. Um fluido escoia dentro dos tubos e o outro através e junto com os tubos, de modo a fornecer ou absorver o calor necessário do fluido sendo aquecido ou resfriado, respectivamente. A ampla aplicabilidade deste tipo de trocador de calor é devido à sua eficácia e capacidade de operar sob uma ampla gama de pressões, vazão e temperaturas do fluido. Além disso, são menos suscetíveis a queda de pressão, em comparação com outros tipos de trocador de calor. Por outro lado, a expansão térmica dos tubos pode ser problemática, se o fluido estiver a uma temperatura mais alta. Estas estruturas são geralmente usadas nas indústrias de geração de energia, refrigeração, refinarias e petroquímica, indústrias de aquecimento e ar condicionado, e indústrias farmacêuticas (PATEL et al., 2019).

O projeto de um trocador de calor é um processo gradual, em que uma configuração geométrica de referência do equipamento é escolhida primeiro e um valor permitido de

queda de pressão é fixo. Em seguida, os valores das variáveis de projeto são definidos com base nas especificações de projeto e na suposição de vários parâmetros mecânicos e termodinâmicos, a fim de obter um coeficiente de transferência de calor satisfatório. Estas escolhas são verificadas com base em procedimentos iterativos que envolvem muitas tentativas até que um projeto razoável seja obtido e atenda às especificações com um compromisso satisfatório entre quedas de pressão e desempenho de troca térmica. Este tipo de abordagem, baseada em tentativa e erro, consome tempo e pode não levar a projetos econômicos, pois nenhum critério de custo é explicitamente levado em conta. Em particular, a minimização de gastos relacionados à energia é crítica, na óptica financeira e de conservação de recursos. Considerando a importância funcional, um projeto de custo mínimo é, portanto, um objetivo importante (CAPUTO et al., 2008).

Devido à transferência de calor desejada no trocador de calor ocorrer na superfície do tubo, a escolha das variáveis geométricas do tubo é importante do ponto de vista do desempenho térmico. O tamanho do tubo é especificado pelo diâmetro externo e espessura da parede. Tubos de menor diâmetro produzem coeficientes de transferência de calor mais altos e, portanto, resultam em um trocador de calor compacto. No entanto, tubos de diâmetro maior são mais fáceis de limpar, mais resistentes e necessários quando a pressão lateral permitida no tubo é pequena. Para uma determinada área de superfície, um trocador mais econômico é possível com um diâmetro pequeno do casco e tubos longos, consistentes com o espaço e a disponibilidade das instalações de manuseio. O número de tubos depende da vazão do fluido e da queda de pressão disponível (ASADI et al., 2014).

A seguir, é apresentado o modelo capaz de estimar o custo de operação de um trocador de calor de casco e tubo. O modelo é baseado nas especificações de construção e capacidade de operação do trocador de calor. Primeiro, considere a modelagem do trocador de calor do lado do tubo. A velocidade do fluido que passa através dos tubos (v_t), necessária para o cálculo do coeficiente de transferência de calor, depende do número de tubos (N_t) e da quantidade de passes de tubos (N_p), e é calculada como (KERN, 1950)

$$v_t = \frac{m_t}{A_t \rho_t} \left(\frac{N_p}{N_t} \right), \quad (105)$$

onde m_t é a taxa de fluxo de massa, ρ_t é a densidade do fluido e a área da seção transversal do tubo é dada por

$$A_t = \frac{\pi}{4} d_i^2,$$

com d_i sendo o diâmetro interno do tubo. Neste modelo, assume-se que $d_i = 0,8d_o$, onde d_o é o diâmetro externo do tubo. Para N_p passes de tubos, o número de tubos pode ser

aproximado por

$$N_t = C \left(\frac{D_s}{d_o} \right)^{n_1}, \quad (106)$$

o qual depende do diâmetro interno do casco (D_s), do diâmetro externo do tubo (d_o) e de duas constantes, C e n_1 , cujos valores variam conforme o número de passes e o arranjo dos tubos. A descrição detalhada de ambas as constantes é mostrada em Sinnott (1993).

Utilizando o valor da velocidade do fluido que passa através dos tubos, calculada na Eq. (105), torna-se possível obter o número de Reynolds, dado por

$$Re_t = \frac{\rho_t v_t d_i}{\mu_t},$$

onde μ_t representa a viscosidade dinâmica do fluido. Por sua vez, considerando o calor específico (C_{pt}) e a condutividade térmica (k_t), número de Prandtl é calculado por

$$Pr_t = \frac{\mu_t C_{pt}}{k_t}.$$

Ambos os valores obtidos para Re_t e Pr_t são empregados no cálculo do coeficiente de transferência de calor. Aproximações suficientemente precisas deste coeficiente são determinadas por faixas de valor, que dependem do número de Reynolds, dadas por (KERN, 1950; LEE et al., 2005)

$$h_t = \begin{cases} \frac{k_t}{d_i} \left(3,657 + \frac{0,0677 \left(Re_t Pr_t \frac{d_i}{L} \right)^{1,33}}{1 + 0,1 Pr_t \left(Re_t \frac{d_i}{L} \right)^{0,3}} \right) & , \text{ se } Re_t < 2300 \\ \frac{k_t}{d_i} \left(\frac{\frac{f_t}{8} (Re_t - 1000) Pr_t}{1 + 12,7 \sqrt{\frac{f_t}{8}} \left(\sqrt[3]{Pr_t^2} - 1 \right)} \left(1 + \frac{d_i}{L} \right)^{0,67} \right) & , \text{ se } 2300 \leq Re_t < 10000 \\ 0,027 \frac{k_t}{d_o} Re_t^{0,8} \sqrt[3]{Pr_t} \left(\frac{\mu_t}{\mu_{wt}} \right)^{0,14} & , \text{ se } Re_t \geq 10000 \end{cases}.$$

Adicionalmente, o cálculo de h_t depende do valor do comprimento dos tubos (L), da viscosidade na parede do tubo, dada por $\mu_{wt} = 0,75\mu_t$, e do fator de fricção de Darcy (HEWITT, 2008), obtido por

$$f_t = (1,82 \log_{10} Re_t - 1,64)^{-2}.$$

Agora, considere a modelagem do trocador de calor do lado do casco. A velocidade

do fluxo, neste caso, depende da taxa de fluxo de massa (m_s), da densidade do fluido (ρ_s) e é dada por (KERN, 1950)

$$v_s = \frac{m_s}{\rho_s A_s},$$

obtida levando-se em consideração a área da seção transversal normal à direção do fluxo,

$$A_s = D_s B \left(1 - \frac{d_o}{P_t} \right),$$

onde B é o espaçamento de defletores e $P_t = 1,25d_o$, definida como a distância entre diâmetros de tubos adjacentes.

Em relação ao casco, o número de Reynolds é dado por

$$Re_s = \frac{\rho_s v_s D_e}{\mu_s} = \frac{m_s D_e}{\mu_s A_s}, \quad (107)$$

onde μ_s é a viscosidade dinâmica e D_e é o diâmetro hidráulico externo do casco, cujo valor depende do tipo de arranjo dos tubos. Das diversas formas que podem ser adotadas, destacam-se os arranjos quadrado e triangular. Nestas condições, o diâmetro hidráulico é dado por (KERN, 1950; SINNOTT, 1993)

$$D_e = \begin{cases} \frac{4 \left(P_t^2 - \left(\frac{\pi}{4} d_o^2 \right) \right)}{\pi d_o} & , \text{ arranjo quadrado} \\ \frac{4 \left(0,43 P_t^2 - 0,5 \left(\frac{\pi}{4} d_o^2 \right) \right)}{0,5 \pi d_o} & , \text{ arranjo triangular} \end{cases} .$$

Por sua vez, o número de Prandtl é calculado por

$$Pr_s = \frac{\mu_s C_{ps}}{k_s}, \quad (108)$$

onde C_{ps} é o calor específico e k_s é a condutividade térmica.

Analogamente ao efetuado para o lado do casco, os números de Reynolds e Prandtl, dados pelas Eqs. (107) e (108), também são empregados no cálculo do coeficiente de transferência de calor do lado do caso (KERN, 1950), obtido por

$$h_s = 0,36 \frac{k_t}{D_e} Re_s^{0,55} Pr_s^{1/3} \left(\frac{\mu_s}{\mu_{ws}} \right)^{0,14} .$$

Neste caso, $\mu_{ws} = 0,75\mu_s$ e representa a viscosidade da parede do casco. Após calcular os coeficientes de transferência de calor no lado casco e do tubo, o coeficiente geral de transferência de calor é determinado levando em consideração a resistência à incrustação nos dois

lados. As resistências de incrustação são atribuídas a partir dos dados da literatura, correspondentes a tipos de fluidos e temperaturas de operação. Portanto, considere as constantes R_{ft} e R_{fs} , que representam o fator de incrustação dos lados do tubo e do casco, respectivamente. O coeficiente geral de transferência de calor é dado por (KERN, 1950; SINNOTT, 1993)

$$U = \frac{1}{\frac{1}{h_s} + R_{fs} + \left(\frac{d_o}{d_i}\right) \left(R_{ft} + \frac{1}{h_t}\right)} .$$

Considerando o fluxo cruzado entre dois defletores adjacentes, a diferença de temperatura média logarítmica é determinada como

$$\Delta T_{LM} = \frac{(T_{hi} - T_{co}) - (T_{ho} - T_{ci})}{\ln\left(\frac{T_{hi} - T_{co}}{T_{ho} - T_{ci}}\right)} .$$

Este termo refere-se à diferença de temperatura média logarítmica, que representa o perfil de temperaturas ao longo do interior de um tubo, com temperatura da superfície constante. Dependendo da configuração do fluxo e do tipo de invólucro, um fator de correção é introduzido para calcular a área de superfície do trocador de calor, requerido quando há múltiplos passes, dado por

$$F = \sqrt{\frac{R^2 + 1}{R^2 - 1}} ,$$

onde R é o coeficiente de correção de temperatura, definido como

$$R = \frac{T_{hi} - T_{ho}}{T_{co} - T_{ci}} ,$$

em que T_{ci} e T_{hi} referem-se às temperaturas de entrada dos fluidos frio e quente, respectivamente, e T_{co} e T_{ho} referem-se às temperaturas de saída dos mesmos fluidos.

Por fim, para transferência de calor sensível, a taxa de transferência de calor é dada por

$$Q = m_s C_{ps} (T_{hi} - T_{ho}) = m_t C_{pt} (T_{co} - T_{ci}) ,$$

a área da superfície do trocador de calor é calculada por

$$A = \frac{Q}{U \Delta T_{LM} F} \tag{109}$$

e, portanto, o comprimento dos tubos pode ser estimado, baseado na área de superfície do

trocador de calor, calculada na Eq. (109), através de

$$L = \frac{A}{\pi d_o N_t} .$$

A queda de pressão em um trocador de calor é a pressão estática do fluido que pode ser gasta para conduzir o fluido através do trocador. Do ponto de vista físico e econômico, existe uma estreita relação entre o coeficiente de transferência de calor e a queda de pressão em todos os trocadores de calor. Para um trocador de calor com capacidade de calor constante, o aumento na velocidade do fluxo causa uma melhoria no coeficiente de transferência de calor, o que resulta em um projeto compacto do trocador e em um menor custo de investimento. Por outro lado, o aumento da velocidade do fluxo causa maior queda de pressão nos trocadores de calor, o que resulta em um custo adicional de operação. Devido a este fato, ao projetar uma troca de calor, a queda de pressão deve ser considerada com a transferência de calor e uma solução ideal para o sistema deve ser encontrada, a fim de alcançar o custo total mínimo, que inclui investimentos e custos operacionais (PATEL; RAO, 2010).

A queda de pressão do lado do tubo para o trocador de calor é calculada como os resultados combinados da queda de pressão distribuída ao longo do comprimento do tubo e das perdas de pressão concentradas nos bicos de entrada e saída. Assim,

$$\Delta P_t = \frac{\rho_t v_t^2}{2} \left(\frac{L}{d_i} f_t + p \right) N_p , \quad (110)$$

onde p é uma constante empírica. Similarmente, a queda de pressão no lado do casco é calculada como

$$\Delta P_s = f_s \left(\frac{\rho_s v_s^2}{2} \right) \left(\frac{L}{B} \right) \left(\frac{D_s}{D_e} \right) .$$

Neste caso, o fator de atrito é calculado por $f_s = 1,44 Re_s^{-0,15}$, válido quando $Re_s \leq 40000$ (PETERS et al., 2003).

O custo total (C_T) de operação de um trocador de calor de casco e tubo envolve os custos de investimento (C_{in}), energético (C_{en}) e operacional total descontado (C_d) (SINNOTT, 1993; TAAL et al., 2003). O investimento de capital é calculado em função da área de superfície do trocador de calor, isto é, $C_{in} = a_1 + a_2 A^{a_3}$. Para um trocador de calor fabricado com aço inoxidável no casco e nos tubos, $a_1 = 8000$, $a_2 = 259,2$ e $a_3 = 0,91$. O custo operacional total descontado, relacionado à energia de bombeamento para superar as perdas por atrito por n_a anos, é calculado por

$$C_d = \sum_{k=1}^{n_a} \frac{C_o}{(1+r)^k} ,$$

onde r é a porcentagem da taxa de desconto (taxa de juros), considerando que o valor futuro do fluxo de caixa líquido em um dado ano é calculado por $C_o = PC_{en}H$, onde $C_{en} = 0,00012$ \$/kW h, o número de horas de operação por ano é $H = 7000$ h e o poder de bombeamento é obtido com

$$P = \frac{1}{E_{eff}} \left(\frac{m_t}{\rho_t} \Delta P_t + \frac{m_s}{\rho_s} \Delta P_s \right),$$

onde E_{eff} representa a eficiência de bombeamento, em termos percentuais. Portanto, a estimativa do custo total de um trocador de calor de casco e tubo é dada por

$$C_T = C_{in} + C_d. \quad (111)$$

O problema de otimização determinístico do custo de operação de um trocador de calor de casco e tubo, cujas variáveis são o diâmetro interno do casco (D_s), o diâmetro externo do tubo (d_o) e o espaçamento de defletores (B), é dado por

$$\begin{aligned} \arg \max_{D_s, d_o, B} \quad & C_T(D_s, d_o, B) \\ \text{Sujeito a} \quad & 0,1 \leq D_s \leq 1,5, \\ & 0,015 \leq d_o \leq 0,051, \\ & 0,05 \leq B \leq 0,5 \end{aligned} \quad (112)$$

cujo objetivo C_T é definido pela Eq. (111). São analisados dois estudos de caso: no primeiro, retirado de Sinnott (1993), considera-se um trocador de calor de metanol e água salobra, com carga térmica $Q = 4,34$ MW. No segundo estudo de caso, apresentado por Kern (1950), o trocador de calor é composto por querosene e petróleo bruto, com carga térmica $Q = 1,44$ MW. Para ambos os casos, é adotado $p = 4$ para a constante empírica relacionada à Eq. (110), assim como realizado por Kern (1950). As propriedades físicas para ambos os estudos de caso são apresentadas na Tabela 14.

Com base na literatura, no primeiro estudo de caso é adotada uma configuração de tubos com arranjo triangular, contendo $N_p = 2$ passes de tubos. Nestas condições, o número de tubos, aproximado pela Eq. (106), é calculado com $C = 0,249$ e $n_1 = 2,207$. Por sua vez, o arranjo quadrático é adotado no segundo estudo de caso, contendo $N_p = 4$ passes de tubos, onde $C = 0,158$ e $n_1 = 2,263$. Nestas condições, vários autores resolveram o problema dado pela Eq. (112) utilizando diversas técnicas, dentre os quais pode-se destacar Caputo et al. (2008), Patel e Rao (2010), Hadidi e Nazari (2013), Hadidi et al. (2013), Asadi et al. (2014), Mohanty (2016) e Libotte et al. (2019b). Os melhores resultados reportados para o custo total, que se tem acesso, são $C_T = \$ 45782$ e $C_T = \$ 19433$, para o primeiro e segundo estudos de caso, respectivamente.

Nenhum dos estudos citados anteriormente considera possíveis incertezas no mo-

Tabela 14 - Valores das propriedades físicas para ambos os estudos de caso do problema de custo de operação de um trocador de calor de casco e tubo.

	Estudo de Caso 1		Estudo de Caso 2	
	Casco Metanol	Tubo Água Salobra	Casco Querosene	Tubo Petróleo Bruto
m (kg/s)	27,8	68,9	5,52	18,8
T_i (°C)	95,0	25,0	199,0	37,8
T_o (°C)	40,0	40,0	93,3	76,7
ρ (kg/m ³)	750,0	995,0	850,0	995,0
C_p (kJ/kg K)	2,84	4,2	2,47	2,05
μ (Pa s)	0,00034	0,0008	0,0004	0,00358
μ_w (Pa s)	0,00038	0,00052	0,00036	0,00213
k (W/m K)	0,19	0,59	0,13	0,13
R_f (m ² K/W)	0,00033	0,0002	0,00061	0,00061

Fonte: Reproduzido de Asadi et al. (2014).

delo definido pela Eq. (112). A seguir, é proposta uma formulação para o problema do custo de operação de um trocador de calor de casco e tubo, que leva em conta a robustez e confiabilidade de soluções. Nesta formulação, as variáveis do problema são tratadas como variáveis aleatórias, isto é, $\mathbf{X} = (D_s, d_o, B)$ é um conjunto de variáveis aleatórias independentes, que seguem a distribuição normal, com desvio-padrão $\sigma_x = 0,25\mu_x$, onde μ_x representa a média. Seja $\mathbf{F} = (C_T(\mu_x, \delta), \beta_t)$, onde o custo total de operação é definido pela Eq. (111), e δ é o vetor de ruído. O problema de otimização multi-objetivo com incertezas é dado por

$$\begin{aligned}
 & \arg \max_{\mu_x} \mathbf{F} \\
 \text{Sujeito a } & P[g_i(\mathbf{X}, \delta) \leq 0] \leq \Phi(-\beta_t) \\
 & 0,1 \leq \mu_{x_1} \leq 1,5 \\
 & 0,015 \leq \mu_{x_2} \leq 0,051 \\
 & 0,05 \leq \mu_{x_3} \leq 0,5 \\
 & 1 \leq \beta_t \leq 3
 \end{aligned} \tag{113}$$

onde $i = 1, \dots, 6$. As restrições de desigualdade são dadas por $g_1(\mathbf{X}, \delta) = 0,1 - X_1$, $g_2(\mathbf{X}, \delta) = X_1 - 1,5$, $g_3(\mathbf{X}, \delta) = 0,015 - X_2$, $g_4(\mathbf{X}, \delta) = X_2 - 0,051$, $g_5(\mathbf{X}, \delta) = 0,05 - X_3$ e $g_6(\mathbf{X}, \delta) = X_3 - 0,5$.

Considere o primeiro estudo de caso, quando se dispõe de metanol e água salobra dos lados do casco e do tubo, respectivamente, de um trocador de calor. Assim como nos casos

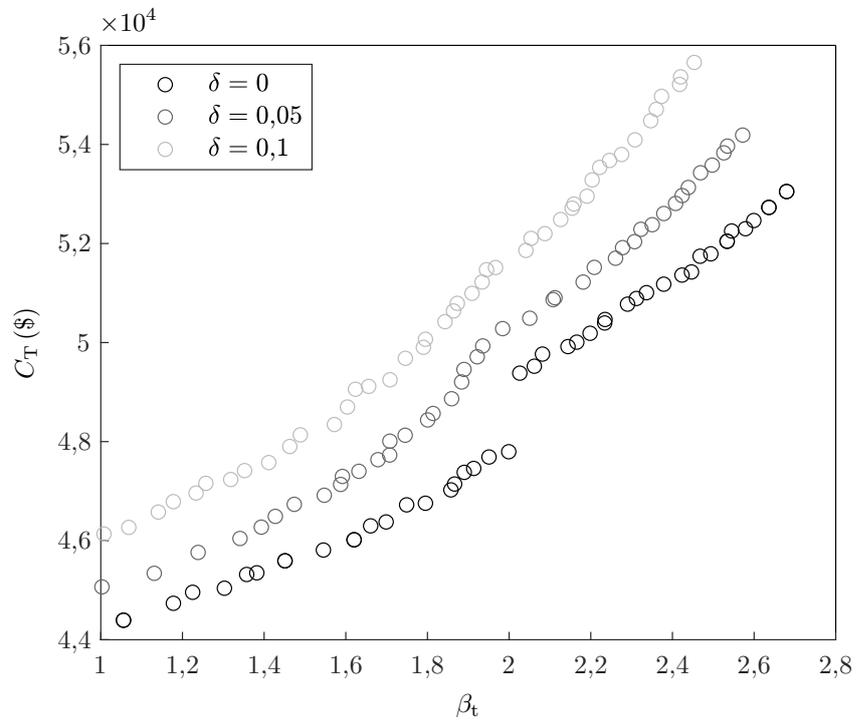
anteriores, o algoritmo MODE, em conjunto com o ASOSL, é empregado na solução deste problema, utilizando o mesmo conjunto de parâmetros de controle já citados. A Fig. 50a mostra os conjuntos de soluções ótimas, obtidas para três diferentes níveis de robustez, em termos dos valores de δ . Novamente, as curvas de Pareto revelam que as soluções dominantes não alcançam o limite superior do índice de confiabilidade e ficam cada vez mais limitadas a isto, à medida que um nível de robustez mais elevado é exigido.

Este modelo também apresenta uma particularidade frente aos demais: quando $\delta = 0$, a curva de Pareto é não-conexa. A descontinuidade ocorre quando $\beta_t \approx 2$, o que equivale à probabilidade de falha $p_f \approx 2,28\%$. Nesta região, os valores ótimos apresentam diferença de custo total de, aproximadamente, \$ 1580. Naturalmente, esta característica representa um fator essencial na escolha da solução ótima, uma vez que resultados com $\beta_t > 2$ podem significar aumento expressivo de C_T , sem que a confiabilidade do modelo seja igualmente elevada. A Fig. 50b apresenta a relação entre a probabilidade de falha p_f , em termos percentuais, e o custo total do sistema. Note que o aumento mais expressivo de C_T ocorre quando a probabilidade de falha tende a zero, concordando com o comportamento da função de distribuição acumulada da distribuição normal.

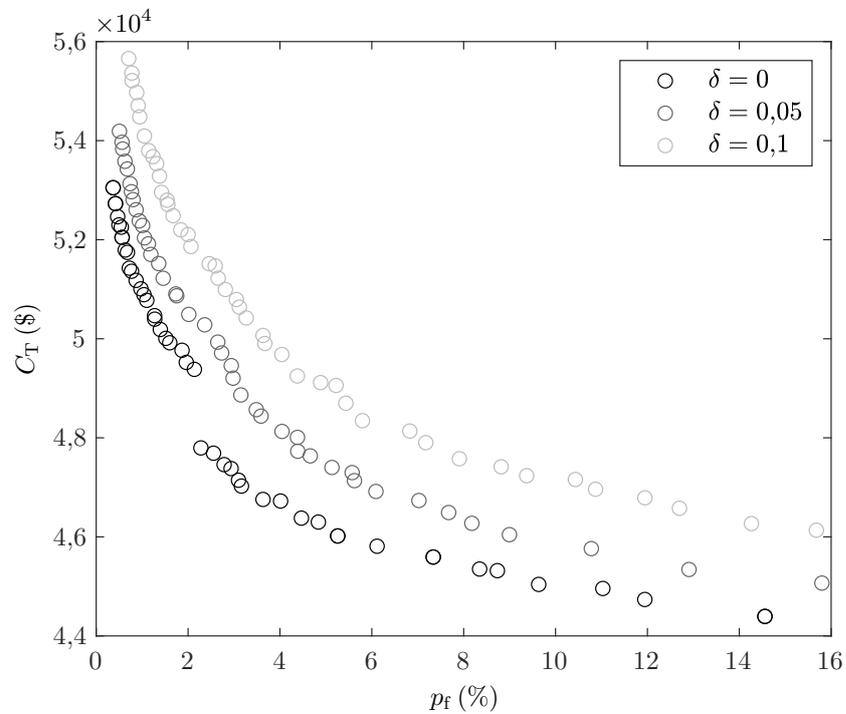
Agora, considere o segundo estudo de caso, para um trocador de calor com querosene e petróleo bruto. A Fig. 51a apresenta as curvas de Pareto obtidas para o problema da Eq. (113), considerando três casos relacionados à robustez, em termos do vetor de incertezas δ . Nesta configuração, os conjuntos de valores ótimos apresentam descontinuidades ainda mais acentuadas do que no caso anterior, como pode ser visto comparando-se os perfis mostrados nas Figs. 50a e 51a. Para determinados níveis de confiabilidade, as curvas apresentam aumento contundente em relação ao custo total do trocador de calor. Note, ainda, que o aumento do nível de robustez faz com esta transição ocorra para valores do índice de confiabilidade cada vez menores, de forma que priorizar soluções com margem de incerteza muito reduzida pode significar um aumento dos custos muito relevante, até mesmo inviabilizando o projeto.

A Fig. 51b mostra a probabilidade de falha do sistema, isto é, a probabilidade de se obter uma solução cujo valor de pelo menos uma das variáveis de decisão não esteja dentro dos limites de viabilidade do modelo físico, em relação ao custo total. Pode-se destacar duas faixas de custo: na primeira, quando $C_T \leq 20350$ aproximadamente, as variáveis de decisão apresentam baixa sensibilidade, de forma que é possível obter soluções com custos semelhantes, mesmo com o aumento do nível de robustez, estendendo moderadamente a probabilidade de falha (que passa de 0,91% para 3,49%, na faixa de robustez avaliada). A segunda faixa de custos, que se referem àqueles quando $C_T > 20350$, o seu aumento ocorre subitamente, para ligeiras mudanças no vetor de variáveis de projeto, evidenciando a sensibilidade das variáveis nesta região.

Figura 50 - Conjunto de resultados do primeiro estudo de caso do problema dado pela Eq. (113), para diferentes níveis de robustez.



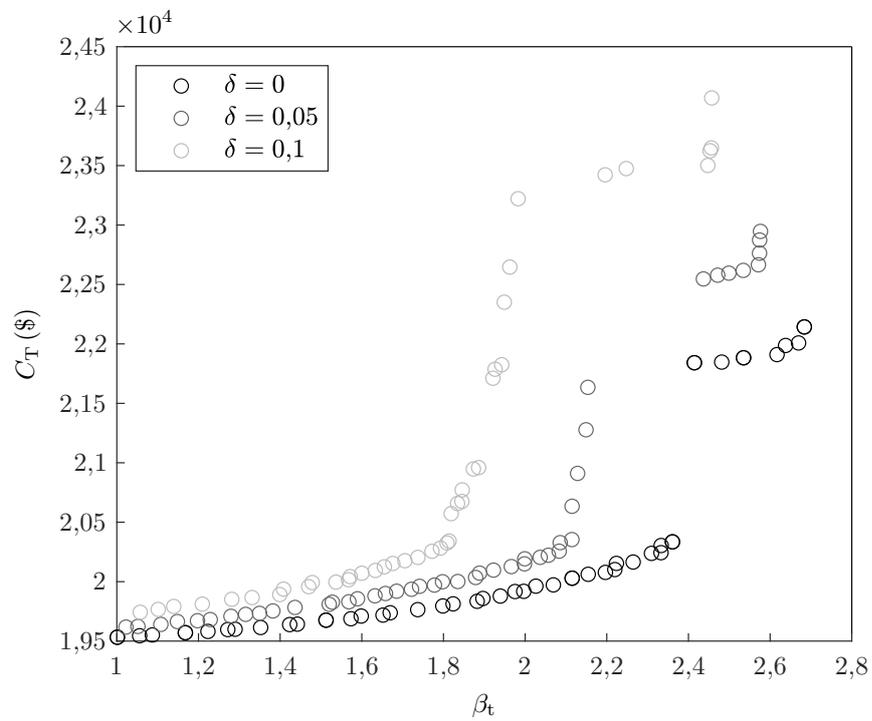
(a) Relação entre o índice de confiabilidade e o custo total.



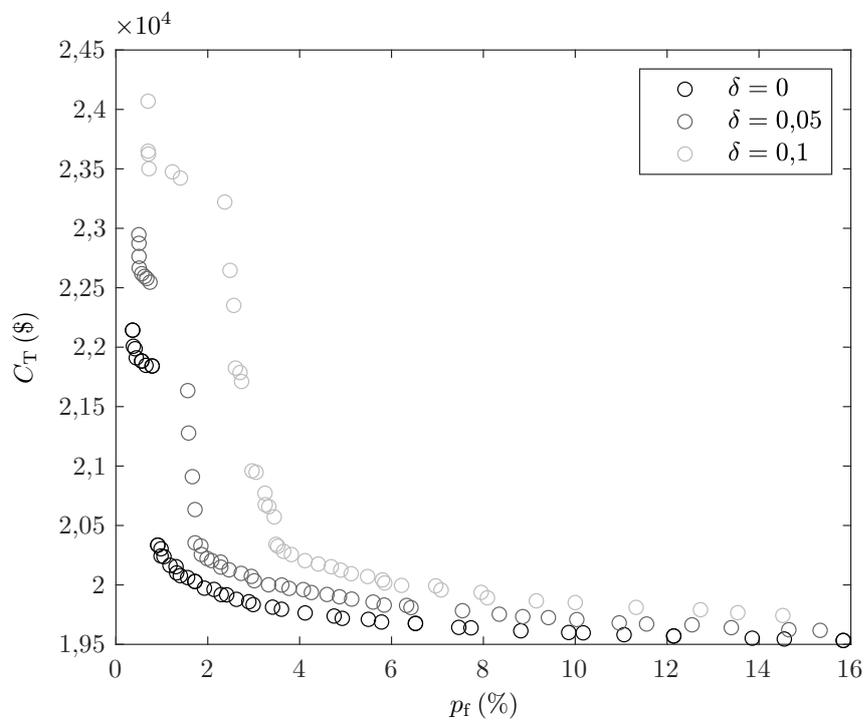
(b) Relação entre a probabilidade de falha e o custo total.

Fonte: O autor, 2019.

Figura 51 - Conjunto de resultados do segundo estudo de caso do problema dado pela Eq. (113), para diferentes níveis de robustez.



(a) Relação entre o índice de confiabilidade e o custo total.



(b) Relação entre a probabilidade de falha e o custo total.

Fonte: O autor, 2019.

CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS

No método proposto para análise de confiabilidade inversa, os testes apresentados deixam claro que assumir um passo constante, como na estratégia adotada pelo SLA, em favor da redução do custo computacional, não se traduz necessariamente em uma redução no número de iterações, tampouco no número de avaliações da função de desempenho. A técnica de obtenção do ponto mais provável de falha proposto no ASOSL para análise de confiabilidade inversa, por outro lado, apresenta uma melhora no desempenho computacional em relação ao SLA, uma vez que a estratégia utilizada para estimar o limite superior do intervalo na busca linear, que fornece uma estimativa inicial para o procedimento de *backtracking*, diminui consideravelmente o número de operações necessárias para se obter o tamanho de passo ótimo.

Observa-se que nos métodos em que o tamanho do passo é calculado usando o valor de aproximações subsequentes ao ponto mais provável de falha, o comportamento oscilatório tende a ocorrer com mais frequência, principalmente se o método demonstrar sensibilidade a determinados parâmetros. Várias escolhas podem produzir valores praticamente constantes para o tamanho de passo durante as iterações, diminuindo a convergência em direção ao ponto ótimo. O ASOSL não apresenta, em nenhum dos testes realizados, a ocorrência de comportamento oscilatório. Este fato também se deve ao cálculo do tamanho de passo ótimo a cada iteração.

Nos testes realizados, o ASOSL esteve entre as metodologias que apresentam menor sensibilidade à escolha dos parâmetros de controle. Essa é uma grande vantagem nos casos em que o método deve ser aplicado a problemas sem conhecimento prévio. Mesmo com poucas opções de parâmetros, o processo iterativo converge para pontos próximos à solução ótima e, em termos gerais, representa a maior taxa de convergência entre as técnicas comparadas na análise proposta. Pode-se concluir que o ASOSL é uma boa alternativa para análise de confiabilidade inversa, principalmente devido às características de obter iterativamente o tamanho de passo ótimo, com possibilidade de aumentar e reduzir seu comprimento, baixo custo computacional, fraca sensibilidade à escolha dos parâmetros de controle e capacidade de evitar comportamentos oscilatórios.

Por sua vez, a formulação proposta neste trabalho, acerca de problemas de otimização multi-objetivo considerando a robustez e confiabilidade dos resultados, revela uma maneira consistente de se considerar incertezas durante o procedimento de otimização de modelos computacionais. Fica clara a contribuição de cada uma destas abordagens nos resultados obtidos, o que torna capaz compreender a dinâmica e finalidade de ambas, e modelar o problema de acordo com as necessidades do projeto. A formulação também apresenta versatilidade, uma vez que permite considerar incertezas de diferentes tipos em quaisquer das variáveis de decisão e restrições de desigualdade do problema, proporcionando auto-

mia necessária para formular o problema que atenda a necessidades específicas.

Apesar de confiabilidade e robustez mostrarem-se conflitantes entre si, em termos de β_t e δ , a análise do problema bidimensional, tratando o índice de confiabilidade como um dos objetivos e fixando um vetor de ruídos, propicia um diagnóstico mais claro dos perfis obtidos. Isto faz com que a escolha do minimizador a ser implementado na prática seja facilitada, uma vez que a inserção de um terceiro objetivo na formulação do problema significaria a obtenção de um conjunto de soluções, no espaço objetivo, representado por uma superfície. Tal escolha também é motivada por outro fator importante: naturalmente, computar uma superfície de soluções dominantes no espaço requer mais recursos computacionais do que para uma curva no plano, quando mantidos os mesmos critérios de refinamento, além do fato de que isso não representa uma vantagem, do ponto de vista da análise dos resultados.

Em todos os problemas estudados, foi observada certa dispersão dos valores ótimos quando os níveis de ruído das variáveis de decisão são elevados. Esta tendência não está ligada, exatamente, ao método de análise de robustez empregado, tampouco à formulação do problema, mas à estratégia utilizada para avaliar a sensibilidade de um dado candidato a solução, levando em conta um conjunto de amostras aleatórias dispostas ao seu redor. A metaheurística avalia a dominância de cada indivíduo da população, a fim de definir quais avançam para gerações subsequentes, considerando a contribuição média destes conjuntos de amostras aleatórias. Conseqüentemente, a sensibilidade da função-objetivo em um dado ponto afeta os operadores genéticos do algoritmo de otimização, fazendo com que os perfis obtidos apresentem tal dispersão.

De forma geral, os resultados são capazes de manifestar a importância de se considerar incertezas durante o procedimento de otimização. Nota-se que soluções robustas e confiáveis tendem a se afastar dos valores ótimos determinísticos, o que evidencia o fato de que minimizadores determinísticos dificilmente podem ser implementados na prática, tendo em vista que a chance de ocorrência de falhas não pode ser negligenciada. Usualmente, priorizar soluções robustas e confiáveis introduz uma questão importante no pós-processamento dos resultados, que diz respeito à escolha da melhor solução a ser utilizada, do ponto de vista econômico. O equilíbrio entre o interesse financeiro e a segurança do sistema é determinante para esta escolha.

Tendo em vista o aumento do custo computacional necessário para se calcular soluções por meio da formulação proposta, é imprescindível que, futuramente, sejam propostas metodologias capazes de reduzir o custo computacional necessário para o cômputo de problemas de otimização multi-objetivo com incerteza. Uma possibilidade para efetivar esta redução é simples: considere a modificação da função-objetivo, dada pela Eq. (87). A penalização de um dado candidato a solução depende da violação de, pelo menos, uma restrição. Isto significa que para $p \geq 2$, onde p é o número de restrições probabilísticas, quando uma destas restrições não é satisfeita, as demais não precisam ser verificadas. Isso pode significar

uma redução de até $(p - 1) \times n_G$ avaliações da função de desempenho, para cada amostra aleatória. Em vista do custo computacional total, essa redução pode representar um ganho considerável de tempo, principalmente para problemas que têm muitas restrições probabilísticas.

Ainda acerca da questão do custo computacional, futuramente pretende-se propor uma nova técnica de análise de confiabilidade inversa, baseada no método do gradiente conjugado. Assim, seria possível obter um ganho ainda maior do que o que foi obtido pelo método ASOSL, pela redução do custo de se avaliar se um dado vetor de projeto satisfaz às restrições probabilísticas. Outra proposta é realizar um estudo acerca da quantidade ótima de amostras aleatórias que devem ser geradas ao redor de um candidato a solução, a fim de se obter um valor ótimo que pondere a redução do custo computacional (em termos do valor de H) e a redução da dispersão do conjunto de valores ótimos no espaço objetivo. Uma terceira possibilidade, mais voltada à otimização robusta, diz respeito à aplicação de metamodelos, que podem ser capazes de reduzir o custo de se avaliar a sensibilidade das variáveis de controle.

REFERÊNCIAS

- ADJIMAN, C.; ANDROULAKIS, I.; FLOUDAS, C. A Global Optimization Method, α BB, for General Twice-differentiable Constrained NLPs — II Implementation and Computational Results. *Computers & Chemical Engineering*, v. 22, n. 9, p. 1159–1179, 1998. ISSN 0098-1354. Citado 2 vezes nas páginas 154 e 155.
- AGGARWAL, K. K. *Reliability Engineering*. 1. ed. Dordrecht: Springer Netherlands, 1993. ISBN 978-94-011-1928-3. Citado 2 vezes nas páginas 203 e 205.
- AIZAWA, A. N.; WAH, B. W. Scheduling of Genetic Algorithms in a Noisy Environment. *Evolutionary Computation*, v. 2, n. 2, p. 97–122, 1994. ISSN 1063-6560. Citado na página 67.
- ANDREI, N. *Relaxed Gradient Descent and a New Gradient Descent Methods for Unconstrained Optimization*. Bucharest, 2004. Research Institute for Informatics (ICI) Technical Notes. Citado 3 vezes nas páginas 119, 122 e 125.
- ANDREI, N. *Nonlinear Optimization Applications Using the GAMS Technology*. Boston: Springer US, 2013. 161–222 p. ISBN 978-1-4614-6797-7. Citado na página 154.
- ANGIRA, R.; BABU, B. V. *Evolutionary Computation for Global Optimization of Non-linear Chemical Engineering Processes*. Pilani, 2002. Citado 2 vezes nas páginas 155 e 160.
- ANKENMAN, B. E.; DEAN, A. M. Quality Improvement and Robustness Via Design of Experiments. In: KHATTREE, R.; RAO, C. R. (Ed.). *Statistics in Industry*. 1. ed. Amsterdam: Elsevier, 2003, (Handbook of Statistics, v. 22). cap. 8, p. 263–317. Citado na página 65.
- ASADI, M.; SONG, Y.; SUNDEN, B.; XIE, G. Economic Optimization Design of Shell-and-tube Heat Exchangers by a Cuckoo Search Algorithm. *Applied Thermal Engineering*, v. 73, n. 1, p. 1032–1040, 2014. ISSN 1359-4311. Citado 3 vezes nas páginas 172, 177 e 178.
- AVRIEL, M.; WILLIAMS, A. C. An Extension of Geometric Programming with Applications in Engineering Optimization. *Journal of Engineering Mathematics*, v. 5, n. 2, p. 187–194, 1971. ISSN 1573-2703. Citado na página 154.
- BABU, B. V.; CHAKOLE, P. G.; MUBEEN, J. H. S. Multiobjective Differential Evolution (MODE) for Optimization of Adiabatic Styrene Reactor. *Chemical Engineering Science*, v. 60, n. 17, p. 4822–4837, 2005. ISSN 0009-2509. Citado 2 vezes nas páginas 20 e 27.
- BÄCK, T.; HAMMEL, U. Evolution Strategies Applied to Perturbed Objective Functions. In: *Proceedings of the First IEEE Conference on Evolutionary Computation*. Orlando: IEEE, 1994. v. 1, p. 40–45. ISBN 0-7803-1899-4. Citado na página 67.
- BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. M. *Nonlinear Programming: Theory and Algorithms*. 3. ed. Hoboken: John Wiley & Sons, Inc., 2006. v. 1. ISBN 978-0471486008. Citado na página 47.
- BELL, M.; SARGENT, R. Optimal Control of Inequality Constrained DAE Systems. *Computers & Chemical Engineering*, v. 24, n. 11, p. 2385–2404, 2000. ISSN 0098-1354. Citado na página 168.

BENAYOUN, R.; MONTGOLFIER, J.; TERGNY, J.; LARITCHEV, O. Linear Programming with Multiple Objective Functions: Step Method (STEM). *Mathematical Programming*, v. 1, n. 1, p. 366–375, 1971. ISSN 1436-4646. Citado na página 46.

BEYER, H.-G.; DEB, K. On Self-adaptive Features in Real-parameter Evolutionary Algorithms. *IEEE Transactions on Evolutionary Computation*, v. 5, n. 3, p. 250–270, 2001. ISSN 1089-778X. Citado na página 54.

BEYER, H.-G.; SENDHOFF, B. Robust Optimization — A Comprehensive Survey. *Computer Methods in Applied Mechanics and Engineering*, v. 196, n. 33, p. 3190–3218, 2007. ISSN 0045-7825. Citado 3 vezes nas páginas 62, 63 e 66.

BICHON, B.; ELDRED, M.; MAHADEVAN, L. S. S.; MCFARLAND, J. Multimodal Reliability Assessment for Complex Engineering Applications using Efficient Global Optimization. In: STRUCTURES, STRUCTURAL DYNAMICS, AND MATERIALS AND CO-LOCATED CONFERENCES. *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. Hunter Mill: American Institute of Aeronautics and Astronautics, 2007. p. 13. Citado na página 84.

BICHON, B. J.; ELDRED, M. S.; SWILER, L. P.; MAHADEVAN, S.; MCFARLAND, J. M. Efficient Global Reliability Analysis for Nonlinear Implicit Performance Functions. *AIAA Journal*, American Institute of Aeronautics and Astronautics, v. 46, n. 10, p. 2459–2468, 2008. ISSN 0001-1452. Citado na página 89.

BIGLARI, M.; LANGSTAFF, A.; ELKAMEL, A. The Application of Response Surface Methodology for the Optimization of an Extractive Distillation Process. *Petroleum Science and Technology*, Taylor & Francis, v. 28, n. 17, p. 1788–1798, 2010. ISSN 1532-2459. Citado na página 65.

BIRJANDI, N.; YOUNESI, H.; BAHRAMIFAR, N.; GHAFARI, S.; ZINATIZADEH, A. A.; SETHUPATHI, S. Optimization of Coagulation-flocculation Treatment on Paper-recycling Wastewater: Application of Response Surface Methodology. *Journal of Environmental Science and Health, Part A*, Taylor & Francis, v. 48, n. 12, p. 1573–1582, 2013. ISSN 1532-4117. Citado na página 65.

BJERAGER, P. Methods for Structural Reliability Computations. In: CASCIATI, F.; ROBERTS, J. B. (Ed.). *Reliability Problems: General Principles and Applications in Mechanics of Solids and Structures*. Vienna: Springer Vienna, 1991. p. 89–135. ISBN 978-3-7091-2616-5. Citado 4 vezes nas páginas 86, 94, 95 e 96.

BOWMAN, V. J. On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-criteria Objectives. In: THIRIEZ, H.; ZIONTS, S. (Ed.). *Multiple Criteria Decision Making*. Berlin: Springer Berlin Heidelberg, 1976. p. 76–86. ISBN 978-3-642-87563-2. Citado na página 41.

BOX, G.; SHOEMAKER, A. C.; TSUI, K.-L.; LEÓN, R. V.; PARR, W. C.; NAIR, V. N.; PREGIBON, D.; CARROLL, R. J.; RUPPERT, D.; GUNTER, B.; ULLMAN, N. R. Signal-to-noise Ratios, Performance Criteria, and Transformations. *Technometrics*, American Society for Quality Control/American Statistical Association, Alexandria, v. 30, n. 1, p. 1–40, 1988. ISSN 0040-1706. Citado na página 64.

- BOX, G. E. P.; WILSON, K. B. On the Experimental Attainment of Optimum Conditions. *Journal of the Royal Statistical Society*, Royal Statistical Society/Wiley, v. 13, n. 1, p. 1–45, 1951. ISSN 00359246. Citado na página 64.
- BOX, J. F. R. A. Fisher and the Design of Experiments, 1922-1926. *The American Statistician*, American Statistical Association, v. 34, n. 1, p. 1–7, 1980. ISSN 0003-1305. Citado na página 62.
- BRANKE, J. Creating Robust Solutions by Means of Evolutionary Algorithms. In: EIBEN, A. E.; BÄCK, T.; SCHOENAUER, M.; SCHWEFEL, H.-P. (Ed.). *Parallel Problem Solving from Nature — PPSN V*. Heidelberg: Springer Berlin Heidelberg, 1998. p. 119–128. ISBN 978-3-540-49672-4. Citado 2 vezes nas páginas 66 e 68.
- BRANKE, J. Efficient Evolutionary Algorithms for Searching Robust Solutions. In: PARMEE, I. C. (Ed.). *Evolutionary Design and Manufacture*. London: Springer London, 2000. p. 275–285. ISBN 978-1-4471-0519-0. Citado 2 vezes nas páginas 66 e 68.
- BREITUNG, K. Asymptotic Approximations for Multinormal Integrals. *Journal of Engineering Mechanics*, v. 110, n. 3, p. 357–366, 1984. ISSN 1943-7889. Citado 2 vezes nas páginas 104 e 108.
- BREITUNG, K. Asymptotic Approximations for Probability Integrals. *Probabilistic Engineering Mechanics*, v. 4, n. 4, p. 187–190, 1989. ISSN 0266-8920. Citado na página 104.
- BREITUNG, K.; HOHENBICHLER, M. Asymptotic Approximations for Multivariate Integrals with an Application to Multinormal Probabilities. *Journal of Multivariate Analysis*, v. 30, n. 1, p. 80–97, 1989. ISSN 0047-259X. Citado na página 104.
- BUCHER, C. G.; BOURGUND, U. A Fast and Efficient Response Surface Approach for Structural Reliability Problems. *Structural Safety*, v. 7, n. 1, p. 57–66, 1990. ISSN 0167-4730. Citado na página 86.
- CAPUTO, A. C.; PELAGAGGE, P. M.; SALINI, P. Heat Exchanger Design Based on Economic Optimisation. *Applied Thermal Engineering*, v. 28, n. 10, p. 1151–1159, 2008. ISSN 1359-4311. Citado 2 vezes nas páginas 172 e 177.
- CASTILLO, E. D.; FAN, S.-K.; SEMPLÉ, J. The Computation of Global Optima in Dual Response Systems. *Journal of Quality Technology*, Taylor & Francis, v. 29, n. 3, p. 347–353, 1997. ISSN 2575-6230. Citado na página 65.
- CASTILLO, E. D.; MONTGOMERY, D. C. A Nonlinear Programming Solution to the Dual Response Problem. *Journal of Quality Technology*, Taylor & Francis, v. 25, n. 3, p. 199–204, 1993. ISSN 2575-6230. Citado na página 65.
- CAVALINI JR, A. A.; LOBATO, F. S.; KOROISHI, E. H.; STEFFEN JR, V. Model Updating of a Rotating Machine Using the Self-adaptive Differential Evolution Algorithm. *Inverse Problems in Science and Engineering*, Taylor & Francis, v. 24, n. 3, p. 504–523, 2016. ISSN 1741-5985. Citado 2 vezes nas páginas 57 e 59.
- CAVAZZUTI, M. *Optimization Methods: From Theory to Design Scientific and Technological Aspects in Mechanics*. Heidelberg: Springer Berlin Heidelberg, 2013. ISBN 978-3-642-31187-1. Citado na página 24.

CHARNES, A.; COOPER, W. W. *Management Models and Industrial Applications of Linear Programming*. 1. ed. New York: John Wiley & Sons, 1961. v. 1. ISBN 978-0471148500. Citado na página 43.

CHARNES, A.; COOPER, W. W.; FERGUSON, R. O. Optimal Estimation of Executive Compensation by Linear Programming. *Management Science*, INFORMS, v. 1, n. 2, p. 138–151, 1955. ISSN 00251909. Citado na página 43.

CHEN, W.; ALLEN, J. K.; TSUI, K.-L.; MISTREE, F. A Procedure for Robust Design: Minimizing Variations Caused by Noise Factors and Control Factors. *Journal of Mechanical Design*, v. 118, n. 4, p. 478–485, 1996. ISSN 1528-9001. Citado na página 65.

CHOI, S.-K.; CANFIELD, R. A.; GRANDHI, R. V. *Reliability-based Structural Design*. 1. ed. London: Springer London, 2007. ISBN 978-1-84628-445-8. Citado 6 vezes nas páginas 91, 92, 204, 207, 212 e 216.

COELLO COELLO, C. A. A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques. *Knowledge and Information Systems*, v. 1, n. 3, p. 269–308, 1999. ISSN 0219-3116. Citado na página 37.

COELLO COELLO, C. A. An Updated Survey of GA-based Multiobjective Optimization Techniques. *ACM Comput. Surv.*, ACM, New York, v. 32, n. 2, p. 109–143, 2000. ISSN 0360-0300. Citado 2 vezes nas páginas 42 e 44.

COELLO COELLO, C. A.; LAMONT, G. B.; VELDHUIZEN, D. A. van. *Evolutionary Algorithms for Solving Multi-Objective Problems*. 2. ed. New York: Springer, 2007. 800 p. ISBN 978-0-387-36797-2. Citado na página 20.

COLLETTE, Y.; SIARRY, P. *Multiobjective Optimization: Principles and Case Studies*. 2. ed. New York: Springer-Verlag Berlin Heidelberg, 2004. v. 1. 293 p. (Decision Engineering, v. 1). ISBN 978-3-662-08883-8. Citado 9 vezes nas páginas 20, 34, 36, 38, 43, 44, 45, 46 e 47.

COPELAND, K. A. F.; NELSON, P. R. Dual Response Optimization via Direct Function Minimization. *Journal of Quality Technology*, Taylor & Francis, v. 28, n. 3, p. 331–336, 1996. ISSN 2575-6230. Citado na página 65.

CORMEN, T. T.; LEISERSON, C. E.; RIVEST, R. L. *Introduction to Algorithms*. Cambridge: MIT Press, 1990. ISBN 0-262-03141-8. Citado na página 31.

CORNELL, C. A. A Probability-Based Structural Code. *American Concrete Institute Journal*, v. 66, n. 12, p. 974–985, 1969. Citado na página 91.

CURSI, E. S.; SAMPAIO, R. Reliability-Based Optimization. In: CURSI, E. S.; SAMPAIO, R. (Ed.). *Uncertainty Quantification and Stochastic Modeling with Matlab*. 1. ed. Amsterdam: Elsevier, 2015. p. 421–433. ISBN 978-1-78548-005-8. Citado 2 vezes nas páginas 18 e 19.

DADEBO, S.; MCAULEY, K. Dynamic Optimization of Constrained Chemical Engineering Problems Using Dynamic Programming. *Computers & Chemical Engineering*, v. 19, n. 5, p. 513–525, 1995. ISSN 0098-1354. Citado na página 168.

DAS, I.; DENNIS, J. E. A Closer Look at Drawbacks of Minimizing Weighted Sums of Objectives for Pareto Set Generation in Multicriteria Optimization Problems. *Structural Optimization*, v. 14, n. 1, p. 63–69, 1997. ISSN 1615-1488. Citado na página 38.

DEB, K. *Multi-Objective Optimization Using Evolutionary Algorithms*. New York: John Wiley & Sons, 2001. ISBN 0-471-87339-X. Citado 11 vezes nas páginas 31, 32, 36, 37, 38, 42, 43, 44, 48, 49 e 54.

DEB, K.; AGRAWAL, R. B. Simulated Binary Crossover for Continuous Search Space. *Complex Systems*, v. 9, n. 2, p. 115–148, 1995. ISSN 0891-2513. Citado na página 54.

DEB, K.; AGRAWAL, S.; PRATAP, A.; MEYARIVAN, T. A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II. In: SCHOENAUER, M.; DEB, K.; RUDOLPH, G.; YAO, X.; LUTTON, E.; MERELO, J. J.; SCHWEFEL, H.-P. (Ed.). *Parallel Problem Solving from Nature — PPSN VI*. Heidelberg: Springer Berlin Heidelberg, 2000. p. 849–858. ISBN 978-3-540-45356-7. Citado na página 49.

DEB, K.; GUPTA, H. Searching for Robust Pareto-Optimal Solutions in Multi-objective Optimization. In: COELLO COELLO, C. A.; AGUIRRE, A. H.; ZITZLER, E. (Ed.). *Evolutionary Multi-Criterion Optimization*. Heidelberg: Springer Berlin Heidelberg, 2005. p. 150–164. ISBN 978-3-540-31880-4. Citado 3 vezes nas páginas 68, 73 e 74.

DEB, K.; GUPTA, H. Introducing Robustness in Multi-Objective Optimization. *Evolutionary Computation*, v. 14, n. 4, p. 463–494, 2006. ISSN 1063-6560. Citado 2 vezes nas páginas 27 e 68.

DEB, K.; PRATAP, A.; AGARWAL, S.; MEYARIVAN, T. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, v. 6, n. 2, p. 182–197, 2002. ISSN 1089-778X. Citado 3 vezes nas páginas 49, 51 e 53.

DELLINO, G.; KLEIJNEN, J. P.; MELONI, C. Robust Optimization in Simulation: Taguchi and Response Surface Methodology. *International Journal of Production Economics*, v. 125, n. 1, p. 52–59, 2010. ISSN 0925-5273. Citado na página 65.

DER KIUREGHIAN, A. First- and Second-Order Reliability Methods. In: NIKOLAIDIS, E.; GHIOCEL, D. M.; SINGHAL, S. (Ed.). *Engineering Design Reliability Handbook*. Boca Raton: CRC Press, 2004. p. 14–1–14–24. ISBN 9780849311802. Citado 3 vezes nas páginas 106, 107 e 108.

DER KIUREGHIAN, A.; ZHANG, Y.; LI, C. Inverse Reliability Problem. *Journal of Engineering Mechanics*, v. 120, n. 5, p. 1154–1159, 1994. ISSN 1943-7889. Citado na página 109.

DONG, M.; WANG, N.; CHENG, X.; JIANG, C. Composite Differential Evolution with Modified Oracle Penalty Method for Constrained Optimization Problems. *Mathematical Problems in Engineering*, Hindawi Publishing Corporation, v. 2014, p. 1–15, 2014. ISSN 1024-123X. Article ID 617905. Citado na página 160.

DU, K.-L.; SWAMY, M. N. S. *Search and Optimization by Metaheuristics: Techniques and Algorithms Inspired by Nature*. 1. ed. Basel: Birkhäuser Basel, 2016. v. 1. ISBN 978-3-319-41192-7. Citado na página 48.

DU, X. *Probabilistic Engineering Design*. 1. ed. Rolla: University of Missouri, 2005. Citado na página 95.

- ECHARD, B.; GAYTON, N.; LEMAIRE, M. AK-MCS: An Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation. *Structural Safety*, v. 33, n. 2, p. 145–154, 2011. ISSN 0167-4730. Citado na página 89.
- EL HAMI, A.; RADI, B.; HUANG, C. Overview of Structural Reliability Analysis Methods — Part I: Local Reliability Methods. *Uncertainties and Reliability of Multiphysical Systems*, v. 1, p. 10, 2017. ISSN 2514-569X. Citado na página 84.
- EL HAMI, A.; RADI, B.; HUANG, C. Overview of Structural Reliability Analysis Methods — Part II: Sampling Methods. *Uncertainties and Reliability of Multiphysical Systems*, v. 1, p. 10, 2017. ISSN 2514-569X. Citado 2 vezes nas páginas 84 e 87.
- EL HAMI, A.; RADI, B.; HUANG, C. Overview of Structural Reliability Analysis Methods — Part III: Global Reliability Methods. *Uncertainties and Reliability of Multiphysical Systems*, v. 1, p. 8, 2017. ISSN 2514-569X. Citado 3 vezes nas páginas 84, 89 e 90.
- ESCHENAUER, H.; KOSKI, J.; OSYCZKA, A. *Multicriteria Design Optimization: Procedures and Applications*. 1. ed. Berlin: Springer-Verlag Berlin Heidelberg, 1990. v. 1. ISBN 978-3-642-48697-5. Citado 3 vezes nas páginas 28, 44 e 45.
- FISSLER, B.; RACKWITZ, R.; NEUMANN, H.-J. Quadratic Limit States in Structural Reliability. *Journal of the Engineering Mechanics Division*, v. 105, n. 4, p. 661–676, 1979. ISSN 0044-7951. Citado na página 104.
- FITZPATRICK, J. M.; GREFENSTETTE, J. J. Genetic Algorithms in Noisy Environments. *Machine Learning*, v. 3, n. 2, p. 101–120, 1988. ISSN 1573-0565. Citado na página 67.
- FLETCHER, R. *Practical Methods of Optimization*. 2. ed. Chichester: John Wiley & Sons, 2013. ISBN 9781118723203. Citado na página 24.
- FONSECA, C. M.; FLEMING, P. J. An Overview of Evolutionary Algorithms in Multiobjective Optimization. *Evolutionary Computation*, v. 3, n. 1, p. 1–16, 1995. ISSN 1063-6560. Citado na página 36.
- GAO, X. K.; LOW, T. S.; LIU, Z. J.; CHEN, S. X. Robust Design for Torque Optimization Using Response Surface Methodology. *IEEE Transactions on Magnetics*, v. 38, n. 2, p. 1141–1144, 2002. ISSN 0018-9464. Citado na página 65.
- GASPAR-CUNHA, A.; COVAS, J. A. Robustness in Multi-objective Optimization Using Evolutionary Algorithms. *Computational Optimization and Applications*, v. 39, n. 1, p. 75–96, 2008. ISSN 1573-2894. Citado na página 20.
- GASS, S.; SAATY, T. The Computational Algorithm for the Parametric Objective Function. *Naval Research Logistics Quarterly*, v. 2, n. 1–2, p. 39–45, 1955. Citado na página 37.
- GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. 1. ed. Boston: Addison-Wesley Longman Publishing, 1989. 432 p. ISBN 0201157675. Citado 3 vezes nas páginas 66, 69 e 71.
- GOLDBERG, D. E. Zen and the Art of Genetic Algorithms. In: SCHAFFER, J. D. (Ed.). *Proceedings of the Third International Conference on Genetic Algorithms*. San Francisco: Morgan Kaufmann, 1989. p. 80–85. ISBN 1-55860-006-3. Citado na página 71.

GUJARATI, D. N.; PORTER, D. C. *Basic Econometrics*. 5. ed. New York: McGraw-Hill Irwin, 2008. ISBN 978-0-07-337577-9. Citado na página 165.

GUNN, D.; THOMAS, W. Mass Transport and Chemical Reaction in Multifunctional Catalyst Systems. *Chemical Engineering Science*, v. 20, n. 2, p. 89–100, 1965. ISSN 0009-2509. Citado 2 vezes nas páginas 167 e 168.

HADIDI, A.; HADIDI, M.; NAZARI, A. A New Design Approach for Shell-and-Tube Heat Exchangers Using Imperialist Competitive Algorithm (ICA) from Economic Point of View. *Energy Conversion and Management*, v. 67, p. 66–74, 2013. ISSN 0196-8904. Citado na página 177.

HADIDI, A.; NAZARI, A. Design and Economic Optimization of Shell-and-Tube Heat Exchangers Using Biogeography-based (BBO) Algorithm. *Applied Thermal Engineering*, v. 51, n. 1, p. 1263–1272, 2013. ISSN 1359-4311. Citado na página 177.

HAIMES, Y. Y.; LASDON, L. S.; WISMER, D. A. On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-1, n. 3, p. 296–297, 1971. ISSN 0018-9472. Citado na página 39.

HALDAR, A.; MAHADEVAN, S. First-Order and Second-Order Reliability Methods. In: SUNDARARAJAN, C. (Ed.). *Probabilistic Structural Mechanics Handbook: Theory and Industrial Applications*. Boston: Springer US, 1995. p. 27–52. ISBN 978-1-4615-1771-9. Citado 2 vezes nas páginas 92 e 99.

HALDAR, A.; MAHADEVAN, S. *Probability, Reliability, and Statistical Methods in Engineering Design*. 1. ed. Chichester: John Wiley & Sons, 2000. v. 1. ISBN 0-471-33119-8. Citado 12 vezes nas páginas 80, 81, 82, 83, 88, 106, 204, 205, 209, 211, 217 e 218.

HAMMEL, U.; BÄCK, T. Evolution Strategies on Noisy Functions How to Improve Convergence Properties. In: DAVIDOR, Y.; SCHWEFEL, H.-P.; MÄNNER, R. (Ed.). *Parallel Problem Solving from Nature — PPSN III*. Heidelberg: Springer Berlin Heidelberg, 1994. p. 159–168. ISBN 978-3-540-49001-2. Citado na página 67.

HAO, P.; WANG, Y.; LIU, C.; WANG, B.; WU, H. A Novel Non-probabilistic Reliability-based Design Optimization Algorithm Using Enhanced Chaos Control Method. *Computer Methods in Applied Mechanics and Engineering*, v. 318, p. 572–593, 2017. ISSN 0045-7825. Citado na página 115.

HARBITZ, A. An Efficient Sampling Method for Probability of Failure Calculation. *Structural Safety*, v. 3, n. 2, p. 109–115, 1986. ISSN 0167-4730. Citado na página 87.

HASOFER, A. M.; LIND, N. C. Exact and Invariant Second-Moment Code Format. *Journal of the Engineering Mechanics Division*, v. 100, n. 1, p. 111–121, 1974. ISSN 0044-7951. Citado 3 vezes nas páginas 93, 100 e 101.

HEWITT, G. *Heat Exchanger Design Handbook*. New York: Begell House, 2008. ISBN 9781567002591. Citado na página 173.

HILL, W. J.; HUNTER, W. G. A Review of Response Surface Methodology: A Literature Survey. *Technometrics*, Taylor & Francis/American Statistical Association/American Society for Quality, v. 8, n. 4, p. 571–590, 1966. ISSN 0040-1706. Citado na página 65.

HOLLAND, J. Genetic Algorithms and the Optimal Allocation of Trials. *SIAM Journal on Computing*, v. 2, n. 2, p. 88–105, 1973. ISSN 1095-7111. Citado na página 67.

HOLLAND, J. H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. 1. ed. Ann Arbor: University of Michigan Press, 1975. Citado na página 69.

HU, X.; COELLO COELLO, C. A.; HUANG, Z. A New Multi-objective Evolutionary Algorithm: Neighbourhood Exploring Evolution Strategy. *Engineering Optimization*, Taylor & Francis, v. 37, n. 4, p. 351–379, 2005. ISSN 1029-0273. Citado 2 vezes nas páginas 56 e 57.

HUGHES, E. J. Evolutionary Multi-objective Ranking with Uncertainty and Noise. In: ZITZLER, E.; THIELE, L.; DEB, K.; COELLO COELLO, C. A.; CORNE, D. (Ed.). *Evolutionary Multi-Criterion Optimization*. Heidelberg: Springer Berlin Heidelberg, 2001. p. 329–343. ISBN 978-3-540-44719-1. Citado na página 68.

HWANG, K.-H.; LEE, K.-W.; PARK, G.-J. Robust Optimization of an Automobile Rearview Mirror for Vibration Reduction. *Structural and Multidisciplinary Optimization*, v. 21, n. 4, p. 300–308, 2001. ISSN 1615-1488. Citado na página 63.

IGUSA, T.; DER KIUREGHIAN, A. Dynamic Characterization of Two-Degree-of-Freedom Equipment-Structure Systems. *Journal of Engineering Mechanics*, v. 111, n. 1, p. 1–19, 1985. ISSN 1943-7889. Citado 2 vezes nas páginas 136 e 137.

IRIZARRY, R. A Generalized Framework for Solving Dynamic Optimization Problems Using the Artificial Chemical Process Paradigm: Applications to Particulate Processes and Discrete Dynamic Systems. *Chemical Engineering Science*, v. 60, n. 21, p. 5663–5681, 2005. ISSN 0009-2509. Citado na página 168.

JIN, Y.; BRANKE, J. Evolutionary Optimization in Uncertain Environments — A Survey. *IEEE Transactions on Evolutionary Computation*, v. 9, n. 3, p. 303–317, 2005. ISSN 1089-778X. Citado na página 69.

JIN, Y.; SENDHOFF, B. Trade-off Between Performance and Robustness: an Evolutionary Multiobjective Approach. In: FONSECA, C. M.; FLEMING, P. J.; ZITZLER, E.; THIELE, L.; DEB, K. (Ed.). *Evolutionary Multi-Criterion Optimization*. Heidelberg: Springer Berlin Heidelberg, 2003. p. 237–251. ISBN 978-3-540-36970-7. Citado 2 vezes nas páginas 20 e 68.

JONG, K. A. D. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. Tese (Doutorado) — University of Michigan, Ann Arbor, 1975. Citado na página 67.

JURECKA, F. *Robust Design Optimization Based on Metamodeling Techniques*. 2007. 184 f. Tese (Doutorado) — Technical University of Munich, Munich, 2007. Citado 2 vezes nas páginas 18 e 62.

KARAMCHANDANI, A.; BJERAGER, P.; CORNELL, C. A. Adaptive Importance Sampling. In: ANG, A. H.-S.; SHINOZUKA, M.; SCHUËLLER, G. I. (Ed.). *Proceedings of 5th International Conference on Structural Safety and Reliability (ICOSSAR '89)*. San Francisco: American Society of Civil Engineers, 1989. p. 855–862. ISBN 0872627438. Citado na página 88.

KARR, A. F. *Probability*. 1. ed. New York: Springer New York, 1993. ISBN 978-1-4612-0891-4. Citado na página 212.

KECECIOGLU, D. *Reliability Engineering Handbook*. 1. ed. Lancaster: Prentice Hall, 1991. v. 1. ISBN 978-0137722945. Citado na página 216.

KERN, D. *Process Heat Transfer*. Tokyo: McGraw-Hill, 1950. ISBN 0-07-085353-3. Citado 5 vezes nas páginas 172, 173, 174, 175 e 177.

KESHTEGAR, B. A Modified Mean Value of Performance Measure Approach for Reliability-Based Design Optimization. *Arabian Journal for Science and Engineering*, v. 42, n. 3, p. 1093–1101, 2017. ISSN 2191-4281. Citado na página 115.

KESHTEGAR, B.; BAHAROM, S.; EL-SHAFIE, A. Self-adaptive Conjugate Method for a Robust and Efficient Performance Measure Approach for Reliability-based Design Optimization. *Engineering with Computers*, v. 34, n. 1, p. 187–202, 2018. ISSN 1435-5663. Citado 2 vezes nas páginas 19 e 116.

KESHTEGAR, B.; CHAKRABORTY, S. Dynamical Accelerated Performance Measure Approach for Efficient Reliability-based Design Optimization with Highly Nonlinear Probabilistic Constraints. *Reliability Engineering & System Safety*, v. 178, p. 69–83, 2018. ISSN 0951-8320. Citado 4 vezes nas páginas 116, 136, 137 e 139.

KESHTEGAR, B.; HAO, P. A Hybrid Self-adjusted Mean Value Method for Reliability-based Design Optimization Using Sufficient Descent Condition. *Applied Mathematical Modelling*, v. 41, p. 257–270, 2017. ISSN 0307-904X. Citado 3 vezes nas páginas 115, 136 e 137.

KESHTEGAR, B.; HAO, P. A Hybrid Descent Mean Value for Accurate and Efficient Performance Measure Approach of Reliability-based Design Optimization. *Computer Methods in Applied Mechanics and Engineering*, v. 336, p. 237–259, 2018. ISSN 0045-7825. Citado na página 116.

KESHTEGAR, B.; HAO, P. Enriched Self-adjusted Performance Measure Approach for Reliability-based Design Optimization of Complex Engineering Problems. *Applied Mathematical Modelling*, v. 57, p. 37–51, 2018. ISSN 0307-904X. Citado na página 116.

KESHTEGAR, B.; HAO, P.; MENG, Z. A Self-adaptive Modified Chaos Control Method for Reliability-based Design Optimization. *Structural and Multidisciplinary Optimization*, v. 55, n. 1, p. 63–75, 2017. ISSN 1615-1488. Citado na página 115.

KESHTEGAR, B.; LEE, I. Relaxed Performance Measure Approach for Reliability-based Design Optimization. *Structural and Multidisciplinary Optimization*, v. 54, n. 6, p. 1439–1454, 2016. ISSN 1615-1488. Citado na página 115.

KHATAEE, A. R.; KASIRI, M. B.; ALIDOKHT, L. Application of Response Surface Methodology in the Optimization of Photocatalytic Removal of Environmental Pollutants Using Nanocatalysts. *Environmental Technology*, Taylor & Francis, v. 32, n. 15, p. 1669–1684, 2011. ISSN 1479-487X. Citado na página 65.

KHEAWHOM, S. Efficient Constraint Handling Scheme for Differential Evolutionary Algorithm in Solving Chemical Engineering Optimization Problem. *Journal of Industrial and Engineering Chemistry*, v. 16, n. 4, p. 620–628, 2010. ISSN 1226-086X. Citado na página 160.

- KIM, K.-J.; LIN, D. K. J. Dual Response Surface Optimization: A Fuzzy Modeling Approach. *Journal of Quality Technology*, Taylor & Francis, v. 30, n. 1, p. 1–10, 1998. ISSN 2575-6230. Citado na página 65.
- KIM, Y.-K.; HONG, J.-P.; LEE, G.-H.; JO, Y.-S. Application of Response Surface Methodology to Robust Design for Racetrack Type High Temperature Superconducting Magnet. *IEEE Transactions on Applied Superconductivity*, v. 12, n. 1, p. 1434–1437, 2002. ISSN 1051-8223. Citado na página 65.
- KIUREGHIAN, A. D.; DE STEFANO, M. Efficient Algorithm for Second-Order Reliability Analysis. *Journal of Engineering Mechanics*, v. 117, n. 12, p. 2904–2923, 1991. ISSN 1943-7889. Citado na página 137.
- KU, K. J.; RAO, S. S.; CHEN, L. Taguchi-aided Search Method for Design Optimization of Engineering Systems. *Engineering Optimization*, Taylor & Francis, v. 30, n. 1, p. 1–23, 1998. Citado na página 63.
- LEE, P.-S.; GARIMELLA, S. V.; LIU, D. Investigation of Heat Transfer in Rectangular Microchannels. *International Journal of Heat and Mass Transfer*, v. 48, n. 9, p. 1688–1704, 2005. ISSN 0017-9310. Citado na página 173.
- LEMAIRE, M.; CHATEAUNEUF, A.; MITTEAU, J. *Structural Reliability*. Hoboken: John Wiley & Sons, 2010. ISBN 9780470611708. Citado na página 101.
- LI, G.; MENG, Z.; HU, H. An Adaptive Hybrid Approach for Reliability-based Design Optimization. *Structural and Multidisciplinary Optimization*, v. 51, n. 5, p. 1051–1065, 2015. ISSN 1615-1488. Citado na página 115.
- LIBOTTE, G. B.; LOBATO, F. S.; MOURA NETO, F. D.; PLATT, G. M. Aplicação do Método da Penalização Robusta para Análise de Problemas de Otimização Multi-objetivo. In: SANTOS, C. C. (Ed.). *Estudos Interdisciplinares nas Ciências Exatas e da Terra e Engenharias*. Ponta Grossa: Atena Editora, 2019. v. 4, cap. 26, p. 277–288. ISBN 978-85-7247-621-8. Citado na página 76.
- LIBOTTE, G. B.; LOBATO, F. S.; PLATT, G. M.; MOURA NETO, F. D. Projeto de Trocadores de Calor de Casco e Tubo Usando o Algoritmo de Polinização de Flores e Modelos Caóticos. *Cereus*, v. 11, n. 1, p. 117–139, 2019. ISSN 2175-7275. Citado na página 177.
- LIBOTTE, G. B.; MOURA NETO, F. D.; LOBATO, F. S.; PLATT, G. M. Second-Order Inverse Reliability Analysis: a New Methodology to the Treatment of Reliability in Engineering System. In: RODRIGUES, H. C.; HERSKOVITS, J.; SOARES, C. M. M.; ARAÚJO, A. L.; GUEDES, J. M.; FOLGADO, J. O.; MOLEIRO, F.; MADEIRA, J. F. A. (Ed.). *EngOpt 2018 Proceedings of the 6th International Conference on Engineering Optimization*. Cham: Springer International Publishing, 2019. p. 855–866. ISBN 978-3-319-97773-7. Citado na página 117.
- LIN, D. K. J.; TU, W. Dual Response Surface Optimization. *Journal of Quality Technology*, Taylor & Francis, v. 27, n. 1, p. 34–39, 1995. ISSN 2575-6230. Citado na página 65.
- LIU, P.; LI, G.; LIU, X. Fast Engineering Optimization: a Novel Highly Effective Control Parameterization Approach for Industrial Dynamic Processes. *ISA Transactions*, v. 58, p. 248–254, 2015. ISSN 0019-0578. Citado na página 168.

LIU, X.; CHEN, L.; HU, Y. Solution of Chemical Dynamic Optimization Using the Simultaneous Strategies. *Chinese Journal of Chemical Engineering*, v. 21, n. 1, p. 55–63, 2013. ISSN 1004-9541. Citado na página 168.

LOBATO, F. S.; SILVA, M. A.; CAVALINI JR., A. A.; STEFFEN JR., V. Reliability-based Multi-objective Optimization Applied to Chemical Engineering Design. *Brazilian Journal of Chemical Engineering*, Scielo, v. 36, p. 317–333, 2019. ISSN 0104-6632. Citado 4 vezes nas páginas 156, 161, 162 e 168.

LOBATO, F. S.; STEFFEN JR, V. A New Multi-objective Optimization Algorithm Based on Differential Evolution and Neighborhood Exploring Evolution Strategy. *Journal of Artificial Intelligence and Soft Computing Research*, Polish Neural Network Society, v. 1, n. 4, p. 259–267, 2011. ISSN 2449-6499. Citado na página 57.

LOBATO, F. S.; STEFFEN JR, V. *Multi-Objective Optimization Problems: Concepts and Self-Adaptive Parameters with Mathematical and Engineering Applications*. 1. ed. Cham: Springer International Publishing, 2017. v. 1. ISBN 978-3-319-58565-9. Citado 6 vezes nas páginas 37, 48, 55, 56, 57 e 58.

LOBATO, F. S.; STEFFEN JR, V.; SILVA NETO, A. J. Solution of Singular Optimal Control Problems Using the Improved Differential Evolution Algorithm. *Journal of Artificial Intelligence and Soft Computing Research*, v. 1, n. 3, p. 195–206, 2011. ISSN 2449-6499. Citado na página 168.

LOCHNER, R. H. Pros and Cons of Taguchi. *Quality Engineering*, Taylor & Francis, v. 3, n. 4, p. 537–549, 1991. ISSN 1532-4222. Citado na página 63.

MANOUSIOUTHAKIS, V.; SOURLAS, D. A Global Optimization Approach to Rationally Constrained Rational Programming. *Chemical Engineering Communications*, Taylor & Francis, v. 115, n. 1, p. 127–147, 1992. Citado na página 160.

MELCHERS, R. E. Improved Importance Sampling Methods for Structural System Reliability Calculation. In: ANG, A. H.-S.; SHINOZUKA, M.; SCHUËLLER, G. I. (Ed.). *Proceedings of 5th International Conference on Structural Safety and Reliability (ICOSSAR '89)*. San Francisco: American Society of Civil Engineers, 1989. p. 1185–1192. ISBN 0872627438. Citado na página 88.

MELCHERS, R. E. *Structural Reliability Analysis and Prediction*. 2. ed. Chichester: Wiley, 1999. ISBN 978-0471987710. Citado na página 85.

MELCHERS, R. E.; BECK, A. T. *Structural Reliability Analysis and Prediction*. 3. ed. Hoboken: John Wiley & Sons, 2018. ISBN 9781119266105. Citado 2 vezes nas páginas 111 e 112.

MENG, Z.; KESHTEGAR, B. Adaptive Conjugate Single-loop Method for Efficient Reliability-based Design and Topology Optimization. *Computer Methods in Applied Mechanics and Engineering*, v. 344, p. 95–119, 2019. ISSN 0045-7825. Citado na página 116.

MENG, Z.; LI, G.; WANG, B. P.; HAO, P. A Hybrid Chaos Control Approach of the Performance Measure Functions for Reliability-based Design Optimization. *Computers & Structures*, v. 146, p. 32–43, 2015. ISSN 0045-7949. Citado na página 115.

- MEYER, P. L. *Probabilidade: Aplicações à Estatística*. 2. ed. Rio de Janeiro: LTC, 1987. ISBN 978-8521602941. Citado 5 vezes nas páginas 92, 203, 208, 209 e 214.
- MIETTINEN, K. *Nonlinear Multiobjective Optimization*. 1. ed. New York: Springer, 1998. 298 p. ISBN 978-1-4615-5563-6. Citado 11 vezes nas páginas 20, 33, 34, 36, 37, 38, 39, 40, 41, 43 e 44.
- MIRJALILI, S.; LEWIS, A. Obstacles and Difficulties for Robust Benchmark Problems: A Novel Penalty-based Robust Optimisation Method. *Information Sciences*, v. 328, p. 485–509, 2016. ISSN 0020-0255. Citado 2 vezes nas páginas 75 e 76.
- MOHANTY, D. K. Application of Firefly Algorithm for Design Optimization of a Shell and Tube Heat Exchanger from Economic Point of View. *International Journal of Thermal Sciences*, v. 102, p. 228–238, 2016. ISSN 1290-0729. Citado na página 177.
- MYERS, R. H.; CARTER, W. H. Response Surface Techniques for Dual Response Systems. *Technometrics*, Taylor & Francis/American Statistical Association/American Society for Quality, v. 15, n. 2, p. 301–317, 1973. ISSN 0040-1706. Citado 2 vezes nas páginas 64 e 66.
- MYERS, R. H.; KHURI, A. I.; CARTER, W. H. Response Surface Methodology: 1966–1988. *Technometrics*, Taylor & Francis/American Statistical Association/American Society for Quality, v. 31, n. 2, p. 137–157, 1989. ISSN 00401706. Citado na página 66.
- MYERS, R. H.; MONTGOMERY, D. C. *Response Surface Methodology: Process and Product in Optimization Using Designed Experiments*. 1. ed. New York: John Wiley & Sons, Inc., 1995. 700 p. ISBN 0471581003. Citado na página 65.
- MYERS, R. H.; MONTGOMERY, D. C.; ANDERSON-COOK, C. M. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 4. ed. Hoboken: John Wiley & Sons, 2016. (Wiley Series in Probability and Statistics). ISBN 978-1-118-91601-8. Citado na página 85.
- MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G.; BORROR, C. M.; KOWALSKI, S. M. Response Surface Methodology: A Retrospective and Literature Survey. *Journal of Quality Technology*, Taylor & Francis, v. 36, n. 1, p. 53–77, 2004. ISSN 2575-6230. Citado na página 66.
- NAIR, V. N.; PREGIBON, D. Analyzing Dispersion Effects from Replicated Factorial Experiments. *Technometrics*, American Society for Quality Control/American Statistical Association, Alexandria, v. 30, n. 3, p. 247–257, 1988. ISSN 0040-1706. Citado na página 64.
- NELDER, J. A.; MEAD, R. A Simplex Method for Function Minimization. *The Computer Journal*, v. 7, n. 4, p. 308–313, 1965. ISSN 1460-2067. Citado na página 65.
- NIKOLAIDIS, E.; BURDISSO, R. Reliability Based Optimization: a Safety Index Approach. *Computers & Structures*, v. 28, n. 6, p. 781–788, 1988. ISSN 0045-7949. Citado na página 96.
- NOCEDAL, J.; WRIGHT, S. *Numerical Optimization*. New York: Springer New York, 2006. ISBN 978-0-387-40065-5. Citado 2 vezes nas páginas 25 e 202.
- O'HAGAN, A. *Probability: Methods and Measurement*. 1. ed. Dordrecht: Springer Netherlands, 1988. ISBN 978-94-009-1211-3. Citado na página 212.

OLSSON, A.; SANDBERG, G.; DAHLBLOM, O. On Latin Hypercube Sampling for Structural Reliability Analysis. *Structural Safety*, v. 25, n. 1, p. 47–68, 2003. ISSN 0167-4730. Citado na página 88.

OSY CZKA, A. *Multicriterion Optimization in Engineering with FORTRAN Programs*. 1. ed. Chichester: Ellis Horwood Limited, 1984. v. 1. ISBN 0-85312-481-7. Citado na página 38.

PARDALOS, P. M.; ŽILINSKAS, A.; ŽILINSKAS, J. *Non-Convex Multi-Objective Optimization*. 1. ed. Cham: Springer International Publishing, 2017. v. 123. ISBN 978-3-319-61005-4. Citado na página 41.

PARK, G.-J.; HWANG, W.-J.; LEE, W.-I. Structural Optimization Post-process Using Taguchi Method. *JSME international journal. Ser. A, Mechanics and material engineering*, v. 37, n. 2, p. 166–172, 1994. ISSN 0914-8809. Citado na página 63.

PARK, G.-J.; LEE, T.-H.; LEE, K. H.; HWANG, K.-H. Robust Design: An Overview. *AIAA Journal*, v. 44, n. 1, p. 181–191, 2006. ISSN 1533-385X. Citado na página 62.

PATEL, V.; RAO, R. Design Optimization of Shell-and-Tube Heat Exchanger Using Particle Swarm Optimization Technique. *Applied Thermal Engineering*, v. 30, n. 11, p. 1417–1425, 2010. ISSN 1359-4311. Citado 2 vezes nas páginas 176 e 177.

PATEL, V. K.; SAVSANI, V. J.; TAWHID, M. A. *Thermal System Optimization: A Population-Based Metaheuristic Approach*. Cham: Springer International Publishing, 2019. ISBN 978-3-030-10477-1. Citado na página 171.

PETERS, M.; TIMMERHAUS, K.; WEST, R. *Plant Design and Economics for Chemical Engineers*. New York: McGraw-Hill Education, 2003. ISBN 9780072392661. Citado na página 176.

RACKWITZ, R. Practical Probabilistic Approach to Design First Order Reliability Concepts for Design Codes. In: *Bulletin d'information 112: First Order Reliability Concepts for Design Codes*. Paris: European Committee for Concrete, 1976. p. 13–89. Citado 2 vezes nas páginas 99 e 101.

RACKWITZ, R.; FIESSLER, B. Structural Reliability Under Combined Random Load Sequences. *Computers & Structures*, v. 9, n. 5, p. 489–494, 1978. ISSN 0045-7949. Citado 2 vezes nas páginas 100 e 102.

RAVINDRAN, A.; RAGSDELL, K. M.; REKLAITIS, G. V. *Engineering Optimization: Methods and Applications*. 2. ed. Hoboken: John Wiley & Sons, 2006. ISBN 978-0-471-55814-9. Citado na página 24.

ROBINSON, T. J.; BORROR, C. M.; MYERS, R. H. Robust Parameter Design: A Review. *Quality and Reliability Engineering International*, v. 20, n. 1, p. 81–101, 2003. ISSN 1099-1638. Citado 2 vezes nas páginas 64 e 66.

ROSENBLATT, M. Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics*, The Institute of Mathematical Statistics, v. 23, n. 3, p. 470–472, 1952. Citado na página 94.

ROSHANIAN, J.; EBRAHIMI, M. Latin Hypercube Sampling Applied to Reliability-based Multidisciplinary Design Optimization of a Launch Vehicle. *Aerospace Science and Technology*, v. 28, n. 1, p. 297–304, 2013. ISSN 1270-9638. Citado na página 88.

ROY, R. K. *A Primer on the Taguchi Method*. 2. ed. Dearborn: Society of Manufacturing Engineers, 1990. 300 p. ISBN 9780872634688. Citado na página 62.

RYOO, H.; SAHINIDIS, N. Global Optimization of Nonconvex NLPs and MINLPs with Applications in Process Design. *Computers & Chemical Engineering*, v. 19, n. 5, p. 551–566, 1995. ISSN 0098-1354. Citado 2 vezes nas páginas 160 e 161.

SHARMA, T. K.; PANT, M.; ABRAHAM, A. Blend of Local and Global Variant of PSO in ABC. In: *2013 World Congress on Nature and Biologically Inspired Computing*. Fargo: Institute of Electrical and Electronics Engineers (IEEE), 2013. p. 113–119. Citado na página 160.

SHINOZUKA, M. Basic Analysis of Structural Safety. *Journal of Structural Engineering*, v. 109, n. 3, p. 721–740, 1983. ISSN 1943-541X. Citado na página 97.

SIARRY, P. *Metaheuristics*. 1. ed. Cham: Springer International Publishing, 2016. v. 1. ISBN 978-3-319-45403-0. Citado na página 48.

SINNOTT, R. *Coulson and Richardson's Chemical Engineering*. 2. ed. Amsterdam: Pergamon, 1993. (Chemical Engineering Technical Series). ISBN 978-0-08-041865-0. Citado 5 vezes nas páginas 173, 174, 175, 176 e 177.

SMITH, E. M. de B. *On the Optimal Design of Continuous Processes*. Tese (Doutorado) — Imperial College of Science, Technology and Medicine, South Kensington, 1996. Citado 2 vezes nas páginas 159 e 160.

SRINIVAS, N.; DEB, K. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, v. 2, n. 3, p. 221–248, 1994. ISSN 1063-6560. Citado na página 49.

STORN, R.; PRICE, K. Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, Kluwer Academic Publishers, Hingham, v. 11, n. 4, p. 341–359, 1997. ISSN 0925-5001. Citado 4 vezes nas páginas 55, 61, 149 e 168.

SUNDARARAJAN, C.; WITT, F. J. Stress-Strength Interference Method. In: SUNDARARAJAN, C. (Ed.). *Probabilistic Structural Mechanics Handbook: Theory and Industrial Applications*. Boston: Springer US, 1995. p. 8–26. ISBN 978-1-4615-1771-9. Citado 2 vezes nas páginas 80 e 82.

TAAL, M.; BULATOV, I.; KLEMEŠ, J.; STEHLÍK, P. Cost Estimation and Energy Price Forecasts for Economic Evaluation of Retrofit Projects. *Applied Thermal Engineering*, v. 23, n. 14, p. 1819–1835, 2003. ISSN 1359-4311. Citado na página 176.

TAGUCHI, G. *Introduction to Quality Engineering: Designing Quality Into Products and Processes*. 1. ed. California: Asian Productivity Organization, 1986. 191 p. ISBN 9789283310846. Citado na página 62.

- TAGUCHI, G. *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. 1. ed. White Plains: Unipub/Kraus International, 1987. Citado na página 63.
- TAGUCHI, G.; CHOWDHURY, S.; TAGUCHI, S. *Robust Engineering: Learn How to Boost Quality While Reducing Costs & Time to Market*. 1. ed. New York: McGraw-Hill, 1999. 241 p. ISBN 0-07-134782-8. Citado 2 vezes nas páginas 19 e 63.
- TAN, X.-h.; SHEN, M.-f.; HOU, X.-l.; LI, D.; HU, N. Response Surface Method of Reliability Analysis and its Application in Slope Stability Analysis. *Geotechnical and Geological Engineering*, v. 31, n. 4, p. 1011–1025, 2013. ISSN 1573-1529. Citado 2 vezes nas páginas 85 e 86.
- TANARTKIT, P.; BIEGLER, L. A Nested, Simultaneous Approach for Dynamic Optimization Problems — II: the Outer Problem. *Computers & Chemical Engineering*, v. 21, n. 12, p. 1365–1388, 1997. ISSN 0098-1354. Citado na página 168.
- TEICH, J. Pareto-Front Exploration with Uncertain Objectives. In: ZITZLER, E.; THIELE, L.; DEB, K.; COELLO COELLO, C. A.; CORNE, D. (Ed.). *Evolutionary Multi-Criterion Optimization*. Heidelberg: Springer Berlin Heidelberg, 2001. p. 314–328. ISBN 978-3-540-44719-1. Citado na página 68.
- THOFT-CRISTENSEN, P.; BAKER, M. J. *Structural Reliability Theory and Its Applications*. 1. ed. Berlin: Springer-Verlag Berlin Heidelberg, 1982. 268 p. ISBN 978-3-642-68697-9. Citado na página 93.
- TROSSET, M. W. *Taguchi and Robust Optimization*. Houston, 1996. Citado na página 63.
- TSUTSUI, S.; GHOSH, A. Genetic Algorithms With a Robust Solution Searching Scheme. *IEEE Transactions on Evolutionary Computation*, v. 1, n. 3, p. 201–208, 1997. ISSN 1089-778X. Citado 4 vezes nas páginas 66, 68, 72 e 73.
- TSUTSUI, S.; GHOSH, A.; FUJIMOTO, Y. A Robust Solution Searching Scheme in Genetic Search. In: VOIGT, H.-M.; EBELING, W.; RECHENBERG, I.; SCHWEFEL, H.-P. (Ed.). *Parallel Problem Solving from Nature — PPSN IV*. Heidelberg: Springer Berlin Heidelberg, 1996. p. 543–552. ISBN 978-3-540-70668-7. Citado 3 vezes nas páginas 66, 67 e 68.
- TU, J. *Design Potential Concept for Reliability-based Design Optimization*. Tese (Doutorado) — The University of Iowa, Iowa City, 1999. Citado na página 151.
- TU, J.; CHOI, K. K.; PARK, Y. H. A New Study on Reliability-Based Design Optimization. *Journal of Mechanical Design*, ASME, v. 121, n. 4, p. 557–564, 1999. ISSN 1050-0472. Citado na página 109.
- VAN VELDHUIZEN, D. A.; LAMONT, G. B. Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art. *Evol. Comput.*, MIT Press, Cambridge, v. 8, n. 2, p. 125–147, 2000. ISSN 1063-6560. Citado na página 36.
- VASSILIADIS, V. *Computational Solution of Dynamic Optimization Problems with General Differential-Algebraic Constraints*. Tese (Doutorado) — Imperial College of Science, Technology and Medicine, South Kensington, 1996. Citado na página 168.

- VERMA, A. K.; SRIVIDYA, A.; KARANKI, D. R. *Reliability and Safety Engineering*. 1. ed. London: Springer London, 2010. ISBN 978-1-84996-232-2. Citado 2 vezes nas páginas 82 e 203.
- VINING, G. G.; MYERS, R. H. Combining Taguchi and Response Surface Philosophies: A Dual Response Approach. *Journal of Quality Technology*, Taylor & Francis, v. 22, n. 1, p. 38–45, 1990. ISSN 2575-6230. Citado 2 vezes nas páginas 64 e 65.
- WOLPERT, D. H.; MACREADY, W. G. No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation*, v. 1, n. 1, p. 67–82, 1997. ISSN 1089-778X. Citado na página 26.
- WU, Y.-T. Computational Methods for Efficient Structural Reliability and Reliability Sensitivity Analysis. *AIAA Journal*, American Institute of Aeronautics and Astronautics, v. 32, n. 8, p. 1717–1723, 1994. ISSN 0001-1452. Citado na página 112.
- WU, Y.-T.; MILLWATER, H. R.; CRUSE, T. A. Advanced Probabilistic Structural Analysis Method for Implicit Performance Functions. *AIAA Journal*, American Institute of Aeronautics and Astronautics, v. 28, n. 9, p. 1663–1669, 1990. ISSN 0001-1452. Citado na página 112.
- YANG, D. Chaos Control for Numerical Instability of First Order Reliability Method. *Communications in Nonlinear Science and Numerical Simulation*, v. 15, n. 10, p. 3131–3141, 2010. ISSN 1007-5704. Citado na página 115.
- YANG, N.; DAS, P. K.; YAO, X. Application of Response Surface Method for Reliability Analysis of Stiffened Laminated Plates. *Ships and Offshore Structures*, Taylor & Francis, v. 10, n. 6, p. 653–659, 2015. Citado na página 85.
- YASEEN, Z. M.; KESHTEGAR, B. Limited Descent-based Mean Value Method for Inverse Reliability Analysis. *Engineering with Computers*, p. 1–13, 2018. ISSN 1435-5663. Citado 4 vezes nas páginas 116, 136, 137 e 139.
- YATES, F. Sir Ronald Fisher and the Design of Experiments. *Biometrics*, International Biometric Society, v. 20, n. 2, p. 307–321, 1964. ISSN 1541-0420. Citado na página 62.
- YI, P.; ZHU, Z. Step Length Adjustment Iterative Algorithm for Inverse Reliability Analysis. *Structural and Multidisciplinary Optimization*, v. 54, n. 4, p. 999–1009, 2016. ISSN 1615-1488. Citado 6 vezes nas páginas 115, 117, 118, 119, 128 e 131.
- YOUN, B.; CHOI, K.; DU, L. Adaptive Probability Analysis Using an Enhanced Hybrid Mean Value Method. *Structural and Multidisciplinary Optimization*, v. 29, n. 2, p. 134–148, 2005. ISSN 1615-1488. Citado na página 115.
- YOUN, B. D.; CHOI, K. K.; DU, L. Enriched Performance Measure Approach for Reliability-based Design Optimization. *AIAA Journal*, American Institute of Aeronautics and Astronautics, v. 43, n. 4, p. 874–884, 2005. ISSN 0001-1452. Citado na página 115.
- YOUN, B. D.; CHOI, K. K.; PARK, Y. H. Hybrid Analysis Method for Reliability-Based Design Optimization. *Journal of Mechanical Design*, ASME, v. 125, n. 2, p. 221–232, 2003. ISSN 1050-0472. Citado 3 vezes nas páginas 112, 113 e 117.

ZADEH, L. Optimality and Non-scalar-valued Performance Criteria. *IEEE Transactions on Automatic Control*, v. 8, n. 1, p. 59–60, 1963. ISSN 0018-9286. Citado na página 37.

ZAHARIE, D. Control of Population Diversity and Adaptation in Differential Evolution Algorithms. In: MATOUŠEK, R.; OŠMERA, P. (Ed.). *Proceedings of the 9th International Conference on Soft Computing*. Brno: University of Technology, 2003. p. 41–46. Citado 2 vezes nas páginas 59 e 60.

ZENTNER, J. M. *A Design Space Exploration Process for Large Scale, Multi-objective Computer Simulations*. Tese (Doutorado) — Georgia Institute of Technology, Atlanta, 2006. Citado na página 88.

ZITZLER, E. *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. Tese (Doutorado) — Swiss Federal Institute of Technology Zurich, Zurich, 1999. Citado na página 36.

ZITZLER, E.; LAUMANN, M.; BLEULER, S. A Tutorial on Evolutionary Multiobjective Optimization. In: GANDIBLEUX, X.; SEVAUX, M.; SÖRENSEN, K.; T’KINDT, V. (Ed.). *Metaheuristics for Multiobjective Optimisation*. Heidelberg: Springer Berlin Heidelberg, 2004. p. 3–37. ISBN 978-3-642-17144-4. Citado na página 21.

APÊNDICE A – Técnica de *Backtracking*

Na otimização sem restrições, a busca linear por *backtracking* é um método unidimensional para escolher, de forma adaptativa, o tamanho do passo. Ele determina a quantidade máxima a ser movida ao longo de uma determinada direção de busca. Dado um ponto $\mathbf{u}^{(k)}$ e uma direção de busca $\mathbf{d}^{(k)}$, o procedimento de *backtracking* reduz iterativamente o tamanho do passo t , começando de $\bar{t}^{(k)}$, até que seja assegurada a condição de descida suficiente (para mais detalhes, consulte Nocedal e Wright (2006)), de acordo com o procedimento dado por:

1. escolha um coeficiente de deslocamento $\alpha_b > 0$, um fator de contração $s_b \in (0, 1)$ e defina $t = \bar{t}^{(k)}$;
2. enquanto $G(\mathbf{u}^{(k)} - t\mathbf{d}^{(k)}) > G(\mathbf{u}^{(k)}) - \alpha_b t (\mathbf{d}^{(k)})^T \mathbf{d}^{(k)}$, faça $t = s_b t$;
3. defina $\tau^{(k)} = t$.

APÊNDICE B – Fundamentos da Teoria de Probabilidades

B.1 Variáveis Aleatórias e Medidas de Dispersão

Um valor variável que denota o resultado de uma dada experiência aleatória é chamado de *variável aleatória*, ou também de *variável estocástica*. Uma variável aleatória, em geral, assume valores distintos em diferentes execuções de um experimento estocástico, uma vez que o seu valor depende de fatores que, em geral, não podem ser previstos (AGGARWAL, 1993). Mais precisamente, Meyer (1987) apresenta a definição abaixo.

Definição 2.1.1. Seja \mathcal{E} um experimento e \mathcal{S} um espaço amostral associado ao experimento, isto é, o conjunto de todos os potenciais resultados do experimento. Uma função X que associe a cada elemento $s \in \mathcal{S}$ um número real, $X(s)$, é denominada *variável aleatória*.

Entre as variáveis aleatórias, há duas categorias mais comumente utilizadas, as variáveis aleatórias discretas e contínuas. Se os elementos do espaço de amostra forem infinitos em número e o espaço de amostra for contínuo, a variável aleatória definida em tal espaço é conhecida como uma *variável aleatória contínua*, que se caracteriza pela existência de uma função densidade de probabilidade (VERMA et al., 2010). Esta categoria de variáveis aleatórias é a utilizada nesta tese e, portanto, será apresentada mais detalhadamente no decorrer deste capítulo.

Existem algumas maneiras de se quantificar a localização e a dispersão de um conjunto de dados observados. Essas medidas podem proporcionar uma visão mais favorável de um conjunto de valores amostrais. Suponha que X é uma variável aleatória e n amostras de X , denotadas por x_i , estão disponíveis. A *média* amostral de X , que de certa forma inclui a tendência principal dos dados, denotada por \bar{x} , é calculada através de

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i .$$

A *variância* amostral de X , que reflete a dispersão dos dados em torno da média amostral, denotada como $\text{Var}[X]$, é uma média do desvio quadrático das amostras aleatórias em relação à sua média amostral. Portanto, a variância de uma variável aleatória X pode ser calculada por meio de

$$\text{Var}[X] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 . \quad (114)$$

Um fato importante sobre a variância amostral é que o seu resultado não é expresso na mesma unidade dos dados amostrais e de sua média. Essa diferença pode ser evitada, calculando-se a raiz quadrada da variância amostral. Esta medida, conhecida como *desvio-*

padrão e denotada por σ_x , é calculada por

$$\sigma_x = \sqrt{\text{Var}[X]} . \quad (115)$$

Como a média e o desvio-padrão amostrais são expressos na mesma unidade, um termo não-dimensional pode ser introduzido, através do cálculo da razão entre ambos. Essa medida é chamada de *coeficiente de variação* e é calculada através de

$$\text{CV}[X] = \frac{\sigma_x}{\bar{x}} .$$

O seu resultado exprime a extensão da variabilidade dos dados em relação à média. Para uma variável determinística, o coeficiente de variação é sempre igual a zero. No caso de variáveis aleatórias, valores reduzidos de $\text{CV}[X]$ caracterizam uma variável com pequena quantidade de incerteza relacionada (HALDAR; MAHADEVAN, 2000).

B.2 Variáveis Aleatórias Contínuas

Suponha que exista um grande número de observações ou registros de dados relacionados a uma dada experiência aleatória. Através desses dados, é possível elaborar um *histograma*, que consiste em dividir o intervalo no qual os dados estão contidos em subintervalos de tamanhos aproximadamente semelhantes e, em seguida, construindo um retângulo sobre cada intervalo, com uma área proporcional ao número de observações relacionadas a cada um dos subintervalos. Se as áreas retangulares forem normalizadas para que a soma total de suas áreas seja unitária, o histograma representaria a distribuição de probabilidade do espaço amostral. Assim, a probabilidade de uma amostra escolhida aleatoriamente ser atribuída a um determinado subintervalo pode ser calculada somando-se a área total dentro desse intervalo.

Há um número infinito de valores que uma variável aleatória contínua pode tomar dentro de um intervalo. Como indicado por Choi et al. (2007), é possível notar que se o histograma fosse construído com um número muito grande de observações e os intervalos se tornassem infinitesimalmente pequenos, à medida que o número de observações crescesse, a distribuição de probabilidade se tornaria uma curva contínua. A função matemática que descreve a distribuição sobre o espaço amostral de uma variável aleatória contínua X , é chamada de função de densidade de probabilidade.

Definição 2.2.2. Diz-se que X é uma variável aleatória contínua, se existir uma função $f_x(x)$, denominada *função densidade de probabilidade* de X , que satisfaça às seguintes condições:

1. A função $f_x(x) \geq 0$, para todo x .

2. A integral $\int_{-\infty}^{\infty} f_x(x) dx = 1$.

3. Para quaisquer valores de $a, b \in (-\infty, \infty)$, tem-se que a probabilidade da variável aleatória X assumir um valor entre a e b é dada por $P[a \leq X \leq b] = \int_a^b f_x(x) dx$.

Quando X é uma variável aleatória contínua, a probabilidade de ocorrência de qualquer valor é nula, isto é, $P[X = x] = 0$. A função densidade de probabilidade não fornece informações diretas sobre a probabilidade da variável aleatória X assumir um determinado valor. Isto claro, não significa que o evento $[X = x]$ seja impossível. Uma maneira alternativa de se descrever a distribuição de probabilidade de variáveis aleatórias contínuas é através do cálculo da função de distribuição acumulada, como definida abaixo.

Definição 2.2.3. Seja X uma variável aleatória contínua. A função $F_x(x)$, chamada de *função de distribuição acumulada* da variável aleatória X , é definida por $F_x(x) = P[X \leq x]$. Assim, se $f_x(x)$ denota a função densidade de probabilidade da variável aleatória X , então

$$F_x(x) = \int_{-\infty}^x f_x(\theta) d\theta .$$

Obviamente, a função densidade de probabilidade equivale à primeira derivada da função de distribuição acumulada, uma vez que, pelo Teorema Fundamental do Cálculo (AGGARWAL, 1993)

$$\frac{dF_x(x)}{dx} = f_x(x) . \quad (116)$$

Além desta característica, é importante observar algumas propriedades relevantes de funções de distribuição acumulada (HALDAR; MAHADEVAN, 2000):

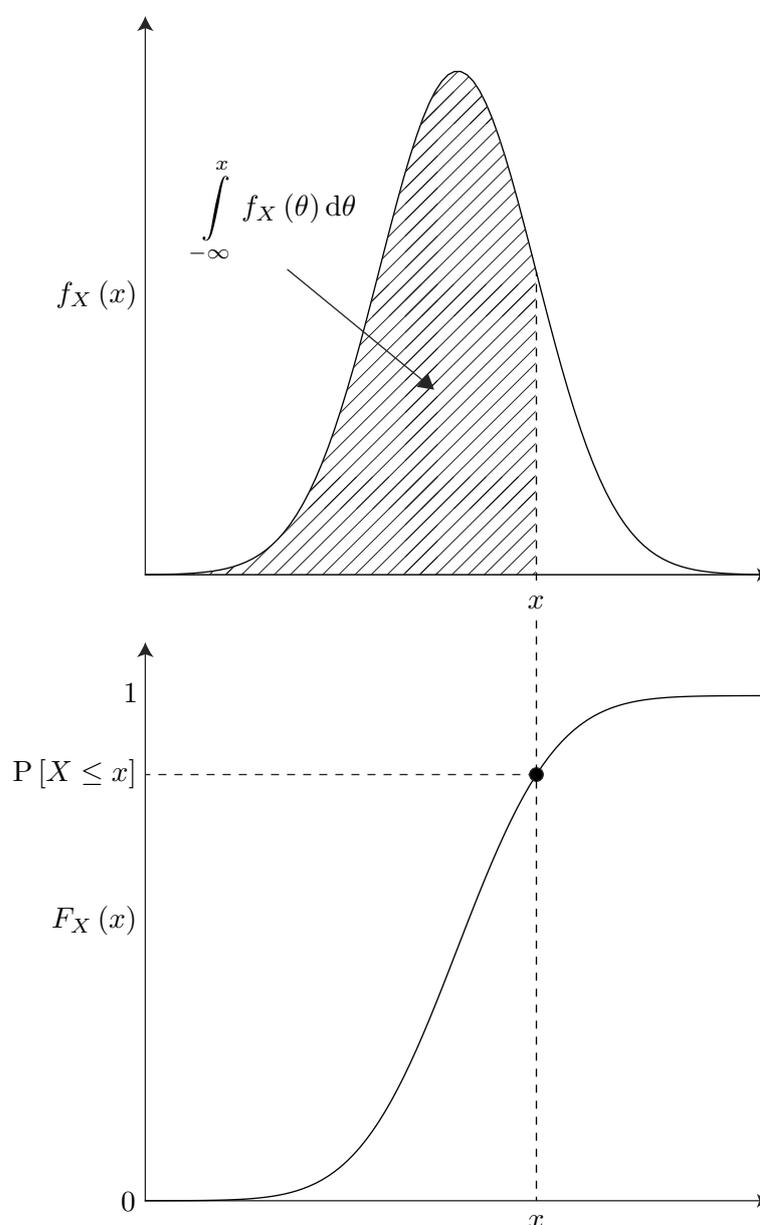
- Tem-se que $\lim_{x \rightarrow -\infty} F_x(x) = 0$ e $\lim_{x \rightarrow \infty} F_x(x) = 1$.
- A função de distribuição acumulada é sempre maior ou igual a zero e apresenta um comportamento não-decrescente, ou seja, se $x_1 \leq x_2$, então $F_x(x_1) \leq F_x(x_2)$.

A relação entre a função densidade de probabilidade e a função de distribuição acumulada é mostrada conceitualmente na Fig. 52.

B.3 Definições Gerais para Descrição de Incertezas

Conhecidas as representações analíticas capazes de descrever a probabilidade de uma variável aleatória tomar um dado valor, torna-se possível especificar as medidas de posição e de dispersão populacional, apresentadas no caso amostral na Seção B.1, em termos

Figura 52 - Relação entre a função densidade de probabilidade e a função de distribuição acumulada associada.



Fonte: O autor, 2019.

das funções densidade de probabilidade. A média ou esperança de uma variável aleatória está relacionada à “tendência central” (ou localização) desta variável, isto é, pode ser compreendida como a média ponderada dos valores que X pode assumir, com pesos fornecidos pela função densidade de probabilidade. Esta interpretação é garantida pela segunda condição da Def. 2.2.2. Portanto, a *média* ou *valor esperado* de uma variável aleatória contínua

X , em relação à função densidade de probabilidade, denotada como $f_x(x)$, é definida por

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_x(x) dx . \quad (117)$$

Neste contexto, o valor esperado também pode ser denotado como μ_x (CHOI et al., 2007).

Como já foi definido, uma variável aleatória X mapeia o espaço amostral S no conjunto dos números reais, isto é, $X : S \rightarrow \mathbb{R}$. Qualquer função de uma variável aleatória também representa uma variável aleatória. Portanto, seja $g(x)$ uma função arbitrária de x . Pode-se verificar que o valor esperado de $g(X)$ é dado por

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_x(x) dx . \quad (118)$$

Por sua vez, a variância de uma variável aleatória contínua é definida de forma similar à variância amostral, apresentada na Eq. (114). O somatório da referida equação é transformado na integração da variável aleatória. Portanto, a variância correspondente a X , denotada como $\text{Var}[X]$, é definida como

$$\text{Var}[X] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f_x(x) dx . \quad (119)$$

Se os termos polinomiais forem expandidos, a Eq. (119) assume a forma

$$\text{Var}[X] = \int_{-\infty}^{\infty} x^2 f_x(x) dx - 2\mathbb{E}[X] \int_{-\infty}^{\infty} x f_x(x) dx + (\mathbb{E}[X])^2 \int_{-\infty}^{\infty} f_x(x) dx .$$

Através da Eq. (117) e da segunda condição imposta pela Def. 2.2.2, fica claro que a expressão que define a variância de uma variável aleatória reduz-se a

$$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 . \quad (120)$$

Neste caso, o desvio-padrão é definido analogamente à Eq. (115), isto é,

$$\sigma_x = \sqrt{\text{Var}[X]} . \quad (121)$$

B.4 Variáveis Aleatórias Múltiplas

De um ponto de vista prático, pode ser necessário considerar mais de uma variável aleatória na formulação de determinado problema. Há casos em que existe o interesse em se observar dois ou mais eventos simultaneamente. Alguns exemplos deste caso são mencionados por Meyer (1987), como o experimento que relaciona dureza e tensão de ruptura de uma peça manufaturada, o estudo referente à estatura e peso de uma população e o histórico de ocorrência do volume de chuva e da temperatura de uma certa região.

Nestas e em diversas outras observações de experimentos estocásticos deste tipo, os espaços amostrais podem ser constituídos de maneira independente. Contudo, quando os eventos são relacionados, essa conduta pode causar a perda de informações estatísticas relevantes. Por esse motivo, é prudente que os dados sejam extraídos de forma conjunta. A partir disso, se faz necessário ampliar a discussão acerca das variáveis aleatórias contínuas para o contexto multivariável. Por uma questão de simplicidade, essas ideias serão apresentadas para o caso de duas variáveis aleatórias, uma vez que a extensão da maioria dos conceitos e definições ocorre de maneira natural. Inicialmente, considere as definições relacionadas às variáveis aleatórias bidimensionais, apresentada abaixo.

Definição 2.4.4. Sejam \mathcal{E} um experimento estocástico e \mathcal{S} um espaço amostral associado a \mathcal{E} . Sejam $X = X(s)$ e $Y = Y(s)$ duas funções, cada uma associando um número real a cada resultado $s \in \mathcal{S}$. Denomina-se (X, Y) uma *variável aleatória bidimensional*.

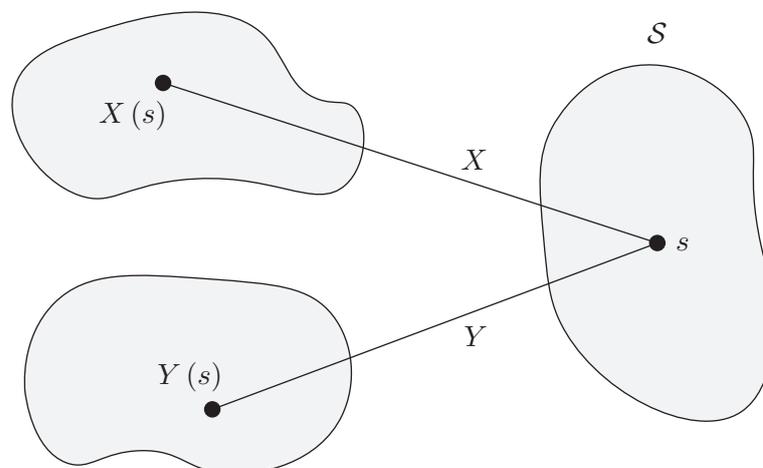
A Fig. 53 ilustra o conceito de variável aleatória bidimensional. Analogamente ao caso unidimensional, é possível definir-se variáveis aleatórias bidimensionais contínuas, as quais dependem da existência de uma função densidade de probabilidade, que é utilizada para o cálculo da probabilidade envolvendo (X, Y) .

Definição 2.4.5. Seja (X, Y) uma variável aleatória contínua tomando todos os valores em alguma região R do plano euclidiano. Uma *função densidade de probabilidade conjunta* $f_{x,y}(x, y)$ é uma função que satisfaz às seguintes condições:

1. A função $f_{x,y}(x, y) \geq 0$, para todo $(x, y) \in \mathbb{R}$.
2. A integral $\iint_R f_{x,y}(x, y) dx dy = 1$.

A intuição por trás da função densidade de probabilidade conjunta é semelhante àquela relacionada a uma variável aleatória unidimensional. Uma probabilidade conjunta representa a chance de dois ou mais eventos aleatórios ocorrerem simultaneamente. Como já foi abordado anteriormente, a área abaixo da curva que descreve a distribuição sobre o espaço amostral de uma variável aleatória contínua unidimensional, calculada dentro de um intervalo especificado, representa a probabilidade dessa variável aleatória assumir um determinado valor arbitrário. A curva que representa a função densidade de probabilidade

Figura 53 - Mapeamento de experimento estocástico com variável aleatória bidimensional.



Fonte: Adaptado de Meyer (1987).

no caso de uma variável aleatória unidimensional passa a ser uma superfície no caso bidimensional. Em contextos mais amplos, essa representação passa a ser uma hipersuperfície e a generalização dos conceitos é imediata.

Definição 2.4.6. A variável aleatória (X, Y) será uma *variável aleatória bidimensional contínua* se existir uma função densidade de probabilidade, $f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$, dado um evento aleatório $A \subset \mathbb{R}^2$, tal que $P[A] = \iint_A f_{X,Y}(x, y) dx dy$.

Dessa definição, segue o conceito da função de distribuição acumulada conjunta (MEYER, 1987). A Fig. 54 apresenta uma representação conceitual da relação entre funções densidade de probabilidade e funções de distribuição acumulada conjuntas.

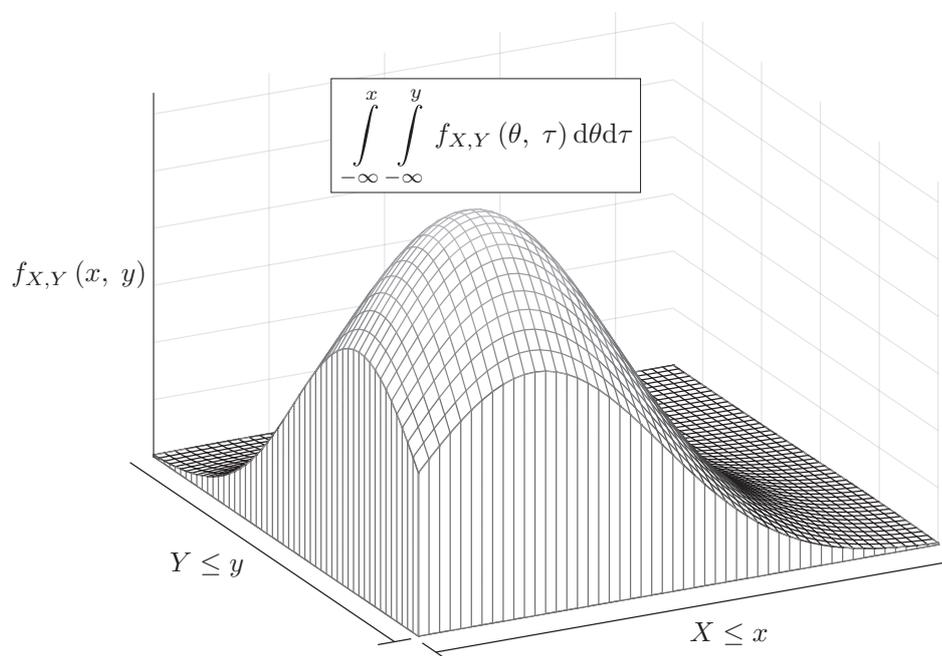
Definição 2.4.7. Seja (X, Y) uma variável aleatória contínua bidimensional. A *função de distribuição acumulada conjunta* $F_{X,Y}(x, y)$ dessa variável aleatória é dada por $F_{X,Y}(x, y) = P[X \leq x, Y \leq y]$, tal que, se $f_{X,Y}(x, y)$ representa a função densidade de probabilidade conjunta da variável aleatória (X, Y) , então

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(\theta, \tau) d\theta d\tau.$$

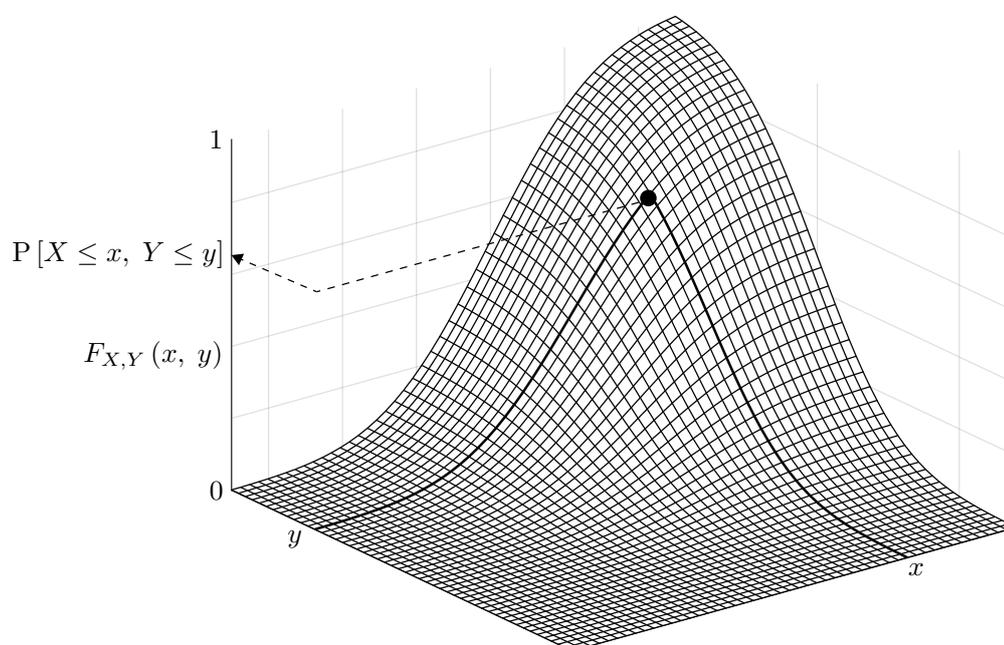
Assim como as propriedades apresentadas na Seção B.2, as funções conjuntas também possuem características similares (HALDAR; MAHADEVAN, 2000):

- Tem-se que $\lim_{x \rightarrow -\infty} F_{X,Y}(x, y) = 0$ e $\lim_{x \rightarrow \infty} F_{X,Y}(x, y) = F_Y(y)$. De forma análoga para y , $\lim_{y \rightarrow -\infty} F_{X,Y}(x, y) = 0$ e $\lim_{y \rightarrow \infty} F_{X,Y}(x, y) = F_X(x)$. Além disso, $\lim_{x,y \rightarrow -\infty} F_{X,Y}(x, y) = 0$ e $\lim_{x,y \rightarrow \infty} F_{X,Y}(x, y) = 1$.

Figura 54 - Área sob a superfície da função densidade de probabilidade conjunta e a representação da respectiva probabilidade $P[X \leq x, Y \leq y]$ na função de distribuição acumulada conjunta.



(a) Função densidade de probabilidade relativa a uma variável aleatória bidimensional.



(b) Função de distribuição acumulada para o espaço amostral de uma variável aleatória bidimensional.

Fonte: O autor, 2019.

- A função de distribuição acumulada conjunta sempre assume valor maior ou igual a zero e exibe comportamento não-decrescente.

Em determinadas aplicações práticas, pode ser conveniente utilizar informações acerca da dependência e independência entre variáveis aleatórias para extrair tantas informações do espaço amostral quanto for possível. No caso de variáveis aleatórias multidimensionais, é preciso determinar uma maneira de medir o nível de interferência entre elas, isto é, estabelecer um parâmetro capaz de avaliar o grau de associação entre as variáveis aleatórias. Isso pode ser feito por meio da análise da covariância e correlação.

A *covariância* de uma variável aleatória bidimensional (X, Y) , a qual é denotada por $\text{Cov}[X, Y]$, indica o nível de relação linear entre os seus componentes. Esse conceito, que é análogo à variância de uma variável aleatória unidimensional (mostrada na Eq. (119)), é definida, para uma variável aleatória bidimensional, por

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] , \quad (122)$$

ou seja, a covariância representa o valor esperado do produto entre as diferenças de cada uma das componentes da variável aleatória e sua respectiva média.

Se X e Y forem estatisticamente independentes, isto é, se o conhecimento do resultado de X não afetar a distribuição de Y , o cálculo da covariância irá resultar em $\text{Cov}[X, Y] = 0$, o que significa que os dados não têm relação linear entre si. Os valores calculados pela Eq. (122) têm unidade de medida igual ao quadrado da unidade do valor esperado. A obtenção de resultados adimensionais é realizada por meio do cômputo do *coeficiente de correlação*, representado por $\rho_{x,y}$ e calculado como

$$\rho_{x,y} = \frac{\text{Cov}[X, Y]}{\sigma_x \sigma_y} ,$$

onde σ_x e σ_y representam o desvio-padrão de cada componente da variável aleatória bidimensional, como mostrado na Eq. (121).

Os valores do coeficiente de correlação variam entre $-1 \leq \rho_{x,y} \leq 1$. Em um espaço amostral típico, dificilmente o coeficiente de correlação assumirá os valores extremos desse intervalo, uma vez que essa ocorrência só se dá quando uma reta passa por todos os pontos do domínio amostral, com o sinal do resultado dependente da inclinação da reta. No caso em que $\rho_{x,y}$ é próximo de zero, as componentes X e Y da variável aleatória têm pouco relação linear entre si e, portanto, são chamadas de não-correlacionadas, como apontado por Haldar e Mahadevan (2000).

B.5 Distribuições de Probabilidade

Uma distribuição de probabilidade é uma função matemática que fornece as chances de ocorrência de diferentes resultados possíveis em um experimento estocástico ou, em outras palavras, descreve o comportamento aleatório de um conjunto de dados, cujos resultados dependem do acaso. O processo de se definir uma distribuição de probabilidade para uma variável aleatória contínua está diretamente relacionado à avaliação da sua função densidade de probabilidade. Contudo, é inviável que esse desenvolvimento ocorra através da avaliação dos infinitos pontos para os quais a função densidade de probabilidade está definida. Em vista dessa necessidade, são definidas algumas funções de distribuição de referência (O'HAGAN, 1988).

A seleção de uma função de distribuição adequada é uma parte essencial da obtenção de características probabilísticas de um sistema na presença de incertezas. A escolha de um determinado tipo de distribuição depende, dentre outros fatores, da natureza do problema, das suposições associadas à distribuição (formato e comportamento) e da conveniência e simplicidade proporcionada pela distribuição em cálculos subsequentes (CHOI et al., 2007). Existem algumas distribuições de probabilidade mais difundidas, como a normal, log-normal, uniforme, exponencial, gama, beta e Weibull. A seguir, as duas primeiras da lista serão detalhadas, uma vez que descrevem o comportamento aleatório da maioria das variáveis com incerteza que serão abordadas nessa tese.

B.5.1 Distribuição Normal

A distribuição normal é uma das mais comumente utilizadas em problemas de engenharia. Essa distribuição geralmente ocorre em aplicações práticas, por conta do seu papel no *teorema central do limite*, que afirma que a soma de variáveis aleatórias de distribuição arbitrária segue assintoticamente uma distribuição normal quando o tamanho da amostra se torna suficientemente grande (KARR, 1993). A função densidade de probabilidade da distribuição normal é definida a seguir.

Definição 2.5.8. A variável aleatória X , que tome todos os valores reais $-\infty < x < \infty$, tem uma *distribuição normal* (ou *distribuição Gaussiana*) se a sua função densidade de probabilidade for da forma

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_x}{\sigma_x}\right)^2\right), \quad (123)$$

onde μ_x , $\sigma_x > 0$ são os parâmetros de posição e dispersão da variável aleatória.

Considerando a Def. 2.2.3, a função de distribuição acumulada correspondente à

distribuição normal é dada por

$$F_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right) dx .$$

Para investigar o comportamento da distribuição normal em relação às medidas de posição e dispersão probabilísticas, primeiro considere a Eq. (117) e, portanto, calcular o valor esperado de uma variável aleatória X com distribuição normal significa obter o resultado de

$$\mathbb{E}[X] = \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} x \exp\left(-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right) dx . \quad (124)$$

Fazendo a transformação linear estabelecida por $Z = (X - \mu_x) / \sigma_x$ e observando que $dx = \sigma_x dz$, a expressão anterior toma a forma

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma_x z + \mu_x) e^{-z^2/2} dz \\ &= \sigma_x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2} dz + \mu_x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz . \end{aligned}$$

Nessa soma de integrais, o termo da esquerda possui integrando $g(z)$ que satisfaz a propriedade $g(z) = -g(-z)$, isto é, $g(z)$ é uma função ímpar e o cálculo da sua integral resulta em zero. A segunda integral é igual a um, pois representa a área total abaixo da curva que descreve a função densidade de probabilidade, como estabelecida na segunda propriedade da Def. 2.2.2, sem considerar o termo μ_x . Portanto, a Eq. (124) equivale a $\mathbb{E}[X] = \mu_x$.

Agora, considere

$$\mathbb{E}[X^2] = \frac{1}{\sqrt{2\pi}\sigma_x} \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right) dx .$$

Por meio da mesma transformações realizada no caso anterior, a expressão de $\mathbb{E}[X^2]$ assume

a forma

$$\begin{aligned}\mathbb{E}[X^2] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma_x z + \mu_x)^2 e^{-z^2/2} dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma_x^2 z^2 e^{-z^2/2} dz + 2\mu_x \sigma_x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2} dz + \mu_x^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz.\end{aligned}$$

Nesse caso, para calcular a primeira integral, é possível proceder com a técnica de integração por partes, por meio das escolhas de $u = z$ e $dv = ze^{-z^2/2} dz$. Dessa forma, o seu resultado (cujo desenvolvimento não será detalhado aqui) será igual a um. No caso da resolução da segunda e terceira integrais desse somatório, pode-se notar que os seus resultados já foram encontrados durante a análise de $\mathbb{E}[X]$, realizada anteriormente. Utilizando os mesmos argumentos já citados, essas integrais valem zero e um, respectivamente.

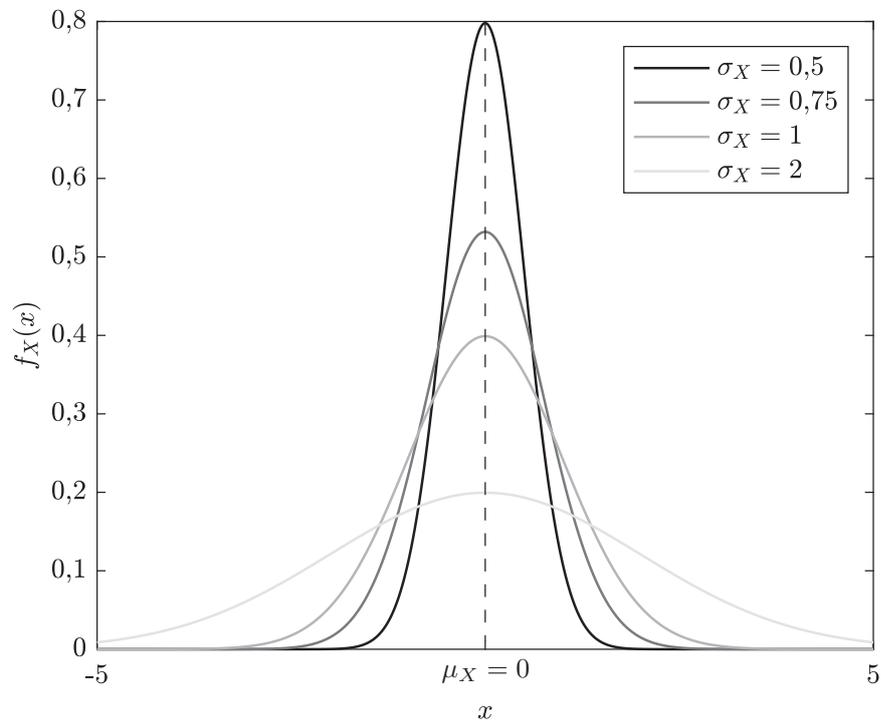
Logo, $\mathbb{E}[X^2] = \sigma_x^2 + \mu_x^2$. Recuperando a formulação da variância, apresentada na Eq. (120), e comparando-a com os resultados obtidos nessa análise, é possível concluir que $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma_x^2 + \mu_x^2 - \mu_x^2 = \sigma_x^2$. Deste modo, verifica-se que os dois parâmetros, μ_x e σ_x^2 , que caracterizam a distribuição normal, são a média e a variância da variável aleatória X , respectivamente (MEYER, 1987).

Portanto, diz-se que uma variável aleatória X tem distribuição normal, denotada como $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$, com média $\mu_x \in \mathbb{R}$, tal que $-\infty < \mu_x < \infty$, e variância $\sigma_x^2 \in \mathbb{R}_+^*$, se a sua densidade de probabilidades for dada pela Eq. (123). Em virtude da dependência de x ocorrer apenas no termo $(x - \mu_x)^2$ da distribuição, fica evidente que a disposição da função densidade de probabilidade se dá de forma simétrica em relação à média da variável aleatória.

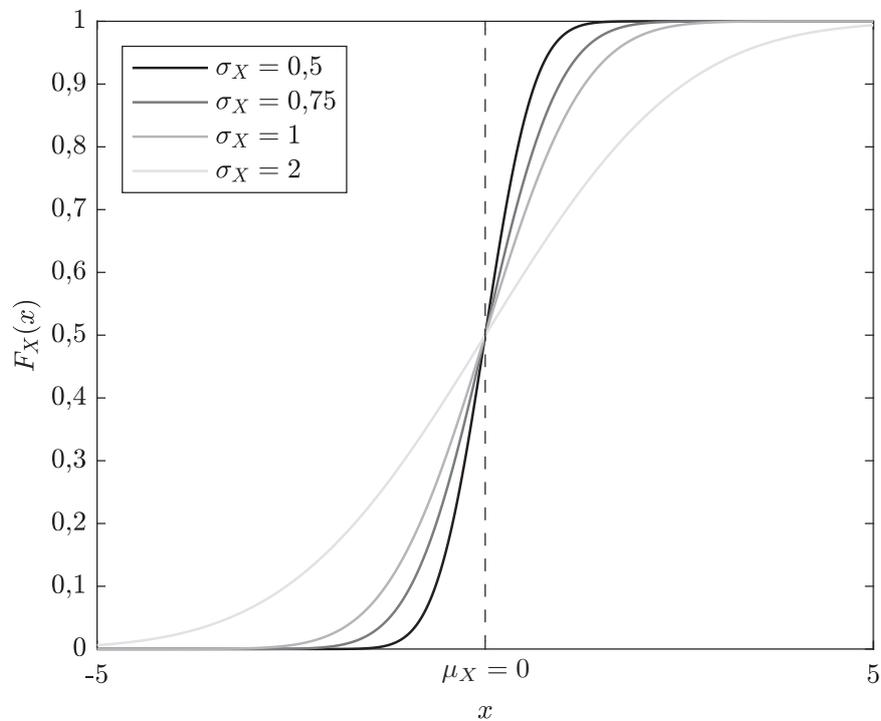
A interpretação geométrica do parâmetro σ tem relação com a mudança de concavidade do gráfico de $f_x(x)$. Como a curva que descreve a função é centrada em $x = \mu_x$, se o seu comportamento for avaliado na vizinhança dessa região, é possível notar que os valores da função são decrescentes e a curva apresenta concavidade para baixo. Quando $x \rightarrow \pm\infty$, $f_x(x) \rightarrow 0$ assintoticamente e o seu gráfico apresenta mudança na concavidade. Esses pontos de inflexão, que podem ser obtidos por meio da resolução de $d^2 f_x(x)/dx^2 = 0$, ocorrem quando $x = \mu_x \pm \sigma_x$. A Fig. 55 mostra quatro casos de funções densidade de probabilidade de uma variável aleatória X com distribuição normal, centrada em $\mu_x = 0$ e com diferentes valores de σ_x , seguida das respectivas curvas que descrevem o comportamento da função de distribuição acumulada associada.

A distribuição Gaussiana pode ser normalizada através da transformação da variável $Z = (X - \mu_x)/\sigma_x$, utilizada anteriormente. Assumindo uma nova variável normal padrão Z , com média igual a zero e desvio-padrão igual a um, a função densidade de probabilidade

Figura 55 - Função densidade de probabilidade e função de distribuição acumulada de uma variável aleatória X com distribuição normal.



(a) Função densidade de probabilidade para $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$.



(b) função de distribuição acumulada para $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$.

Fonte: O autor, 2019.

da Eq. (123) assume a forma

$$\phi_z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right). \quad (125)$$

Naturalmente, a função de distribuição acumulada também está sujeita à transformação na variável aleatória, passando a ser representada como

$$\Phi_z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{1}{2}\theta^2\right) d\theta.$$

Especificamente nesse caso em que $Z \sim \mathcal{N}(0, 1)$, é dito que Z possui distribuição normal padrão. Se $\Phi_z(z) = p_z$ é conhecida, isto é, a probabilidade p_z da variável aleatória assumir um valor tal que $Z \leq z$ é dada, é possível recuperar o valor de z , por meio de $z = \Phi_z^{-1}(p_z)$, cuja função é conhecida como função inversa da função de distribuição acumulada da distribuição normal padrão e não pode ser expressa utilizando funções elementares (CHOI et al., 2007). Em virtude da simetria da distribuição normal padrão, é válido que $\Phi_z(-z) = 1 - \Phi_z(z) = p_z$ e $-z = \Phi_z^{-1}(p_z) = -\Phi_z^{-1}(1 - p_z)$.

B.5.2 Distribuição Log-normal

Por conta de determinados aspectos físicos em problemas de engenharia, algumas variáveis aleatórias não podem assumir valores negativos. Em situações deste tipo, considerar o logaritmo natural das variáveis aleatórias pode ser uma alternativa para a eliminação da possibilidade dessas variáveis assumirem valores menores ou iguais a zero. Quando $Y = \ln X$, a variável aleatória X é dita seguir a distribuição log-normal se Y segue a distribuição normal (KECECIOGLU, 1991). Portanto, diz-se que uma variável aleatória X tem distribuição log-normal se $\ln X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ ou, por conveniência, esta representação também pode ser dada por $X \sim \mathcal{LN}(\mu_x, \sigma_x^2)$. Note que, nesta notação, μ_x e σ_x^2 não representam a média e variância de X , respectivamente, mas de $\ln X$.

Sejam $\phi_x(x)$ e $\Phi_x(x)$ a função densidade de probabilidade e a função de distribuição acumulada de uma variável aleatória X , respectivamente, com $X \sim \mathcal{LN}(0, 1)$. Pela Def. 2.2.3, pode-se inferir que $\Phi_x(x) = P[\ln X \leq \ln x]$. Consequentemente, é possível obter a função densidade de probabilidade da distribuição log-normal, por meio de

$$f_x(x) = \frac{d}{dx} \Phi_x\left(\frac{\ln x - \mu_x}{\sigma_x}\right),$$

tomando como base a propriedade mostrada na Eq. (116) e considerando a transformação

de x , que assegura que $\mu_x = 0$ e $\sigma_x = 1$. Resolvendo a derivada, obtém-se

$$f_x(x) = \phi_x \left(\frac{\ln x - \mu_x}{\sigma_x} \right) \frac{d}{dx} \left(\frac{\ln x - \mu_x}{\sigma_x} \right),$$

que é equivalente a

$$f_x(x) = \phi_x \left(\frac{\ln x - \mu_x}{\sigma_x} \right) \frac{1}{\sigma_x x}.$$

Substituindo a função densidade de probabilidade da distribuição normal padrão, apresentada na Eq. (125), e respeitando a transformação da variável, a forma final da função densidade de probabilidade da *distribuição log-normal* é

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x x} \exp \left(-\frac{1}{2} \left(\frac{\ln x - \mu_x}{\sigma_x} \right)^2 \right).$$

De fato, os dois parâmetros da distribuição log-normal podem ser obtidos utilizando as informações da média e desvio-padrão de uma variável aleatória com distribuição normal. Tais expressões são provenientes dos conceitos apresentados na Seção B.3, que definem a média e o desvio-padrão de uma variável aleatória contínua em termos gerais. Para isto, considere uma outra variável aleatória $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. Pelas medidas de localização e dispersão computadas para Y , Haldar e Mahadevan (2000) mostram que os parâmetros da média e desvio-padrão de X , que tem distribuição log-normal, seguem a relação

$$\sigma_x = \sqrt{\ln \left(1 + \left(\frac{\sigma_Y}{\mu_Y} \right)^2 \right)} \quad (126)$$

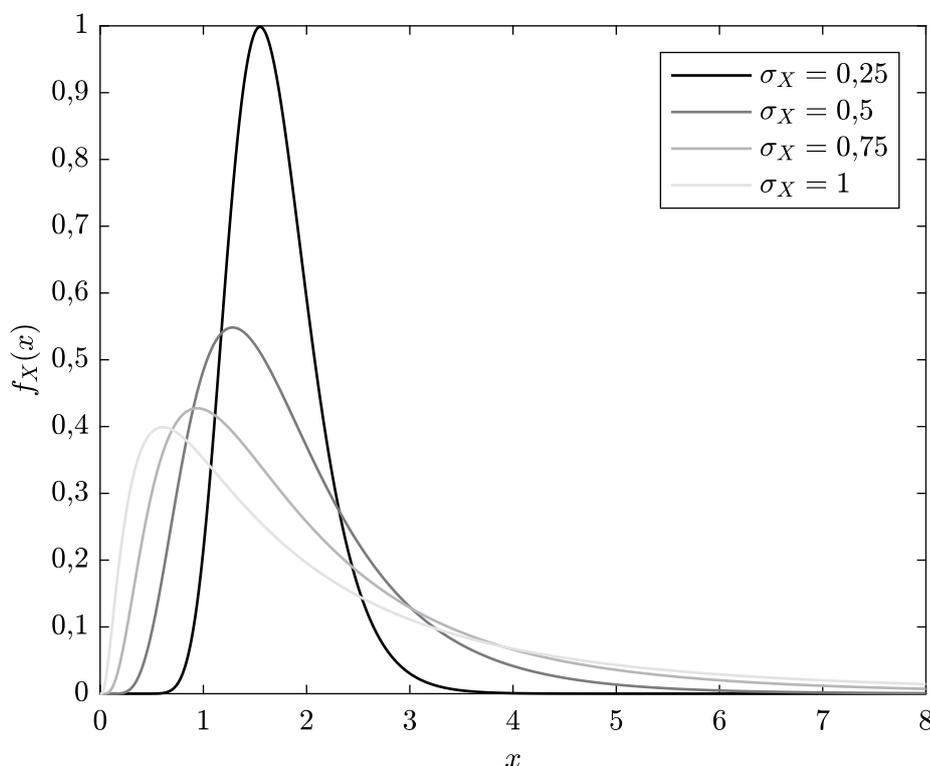
e

$$\mu_x = \ln \mu_Y - \frac{1}{2} \sigma_x^2. \quad (127)$$

Apesar de apresentar algumas similaridades com a distribuição normal, também existem diferenças relevantes entre ambas as distribuições. Inicialmente, é óbvio que a função densidade de probabilidade da distribuição log-normal só vale quando $x > 0$. Além disso, o seu comportamento não apresenta simetria em torno da média, como pode ser visto na Fig. 56, que mostra o comportamento da função densidade de probabilidade de uma variável aleatória X com distribuição log-normal, com $\mu_x = 0,5$ e diferentes valores para a variância.

Os procedimentos utilizados para o cálculo da probabilidade de ocorrência de um evento onde a variável aleatória tem distribuição normal padrão são similares aos aplicados

Figura 56 - Função densidade de probabilidade da distribuição log-normal, com $\mu_x = 0,5$ e diferentes valores para σ_x .



Fonte: O autor, 2019.

para se obter as soluções quando as variáveis têm distribuição log-normal, com exceção da necessidade de realizar a devida transformação na variável independente. De maneira similar à transformação realizada anteriormente, seja $Z = (\ln X - \mu_x) / \sigma_x$. A probabilidade de uma variável aleatória com distribuição log-normal assumir um valor entre dois limites, denotados por a e b , é calculada por

$$P[a \leq X \leq b] = \frac{1}{\sqrt{2\pi}} \int_{\zeta_1}^{\zeta_2} \exp\left(-\frac{1}{2}z^2\right) dz ,$$

tal que $\zeta_1 = (\ln a - \mu_x) / \sigma_x$ e $\zeta_2 = (\ln b - \mu_x) / \sigma_x$. Novamente, considerando $\Phi_x(x)$ como a função de distribuição acumulada da distribuição normal padrão, a probabilidade anterior pode ser expressa como $P[a \leq X \leq b] = \Phi_x(\zeta_2) - \Phi_x(\zeta_1)$ (HALDAR; MAHADEVAN, 2000).