

UNIVERSIDADE DO ESTADO DE SANTA CATARINA– UDESC
CENTRO DE CIÊNCIAS TECNOLÓGICAS– CCT
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA–
PPGEM

RAFAEL DE SOUZA LOPES

OTIMIZAÇÃO PARAMÉTRICA UTILIZANDO MODELO SUBSTITUTO
PARA A REDUÇÃO DE VIBRAÇÕES EM MÁQUINAS ELÉTRICAS
GIRANTES

JOINVILLE, SC
2021

RAFAEL DE SOUZA LOPES

**OTIMIZAÇÃO PARAMÉTRICA UTILIZANDO MODELO SUBSTITUTO
PARA A REDUÇÃO DE VIBRAÇÕES EM MÁQUINAS ELÉTRICAS
GIRANTES**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica do Centro de Ciências Tecnológicas da Universidade do Estado de Santa Catarina, como parte dos requisitos necessários para obtenção do título de Mestre em Engenharia Mecânica, área de concentração: Modelamento e Simulação Numérica.

Orientador: Prof. Dr. Eduardo Lenz Cardoso.

**JOINVILLE, SC
2021**

Lopes, Rafael de Souza

Otimização paramétrica utilizando modelo substituto para a redução de vibrações em máquinas elétricas girantes / Rafael de Souza Lopes. – 2021.

126 p.

Orientador: Eduardo Lenz Cardoso
Dissertação (Mestrado) – Universidade do Estado de Santa Catarina, Centro de Ciências Tecnológicas, Programa de Pós-Graduação em Engenharia Mecânica, Joinville, 2021.

1. Kriging. 2. Modelo substituto. 3. Máquina elétrica girante. 4. Vibração em duas vezes a frequência da rede. 5. Transmissibilidade de vibração. I. Lenz Cardoso, Eduardo. II. Universidade do Estado de Santa Catarina, Centro de Ciências Tecnológicas, Programa de Pós-Graduação em Engenharia Mecânica. III. Título.

RAFAEL DE SOUZA LOPES

**OTIMIZAÇÃO PARAMÉTRICA UTILIZANDO MODELO SUBSTITUTO
PARA A REDUÇÃO DE VIBRAÇÕES EM MÁQUINAS ELÉTRICAS
GIRANTES**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica do Centro de Ciências Tecnológicas da Universidade do Estado de Santa Catarina, como parte dos requisitos necessários para obtenção do título de Mestre em Engenharia Mecânica, área de concentração: Modelamento e Simulação Numérica.

BANCA EXAMINADORA

Membros:

Prof. Dr. Eduardo Lenz Cardoso
(Presidente / Orientador)
Universidade do Estado de Santa Catarina - UDESC

Prof. Dr. Ricardo de Medeiros
Universidade do Estado de Santa Catarina - UDESC

Dr. Olavo M. da Silva Junior
Fundação Stemmer para Pesquisa, Desenvolvimento e Inovação (FEESC)

Joinville, 21 de outubro de 2021.

AGRADECIMENTOS

À minha esposa Ana Paula, obrigado pelo amor incondicional durante esta jornada. Obrigado pela paciência e incentivo para a concretização deste trabalho.

Aos meus pais, meu agradecimento pelas lições de amor. Mesmo com a distância me apoiam transmitindo carinho e exemplos. Me sinto orgulhoso e privilegiado de ter vocês como meus pais.

À minha família, irmão, avós, tios(as) e primos(as), por compreenderem a ausência nos divertidos encontros de família.

Ao professor Dr. Eduardo Lenz Cardoso, pelo profissionalismo, tamanha paciência ao me ensinar e por ter me aceitado como orientando. Obrigado pelas conversas encorajadoras e conselhos enriquecedores para este trabalho.

À WEG, pela oportunidade de realizar as aulas do mestrado. Agradeço aos colegas de trabalho e aos gestores pelo convívio, gerando um clima motivador para buscar novos conhecimentos.

À FAPESC pela contribuição ao programa e suporte à pesquisa. Obrigado à UDESC e aos demais professores, por proporcionar ensino de alta qualidade, agregando intelecto e mantendo a ciência como fonte de desenvolvimento na sociedade.

RESUMO

Uma das principais preocupações no projeto de uma máquina elétrica girante (MEG) é a vibração associada a duas vezes a frequência da rede. Quando uma MEG está operando, essa vibração proveniente do estator é transmitida para a carcaça e aos mancais. Caso os valores de vibração estejam acima do solicitado, ocorrerá o desligamento imediato da MEG. Com o objetivo de reduzir esse tipo de vibração, os fabricantes de MEGs projetam componentes mecânicos para serem instalados na interface entre o estator e a carcaça. Esses componentes tem como objetivo atender os limites de vibração definidos em normas internacionais como a IEC 60034 e a NEMA MG 1, bem como as solicitações do mercado. Os procedimentos experimentais para reduzir os níveis de vibração durante a fase de projeto são caros e exigem muito tempo. Dessa forma, procedimentos de simulação e otimização numérica são empregados para evitar tais inconvenientes, reduzindo o custo durante a fase de projeto. No entanto, a otimização requer uma grande quantidade de simulações que também possuem alto custo computacional. O presente trabalho visa definir um novo mecanismo para atenuar a transmissibilidade de vibração entre o estator e a carcaça de MEG de grande porte, alterando a frequência natural do sistema e/ou reduzindo a amplitude de vibração naquela região. O projeto desse mecanismo envolve otimização paramétrica com o uso de modelos substitutos baseados na teoria de Kriging, buscando reduzir o tempo computacional. Modelos substitutos dão agilidade ao processo de desenvolvimento e otimização do produto, diminuindo o número de avaliações numéricas de alto custo computacional. Para o projeto do mecanismo, um problema de otimização é definido e por meio de modelos substitutos as funções objetivo e de restrições são aproximadas. A função objetivo é usada para ajustar a frequência natural do sistema a um determinado valor, enquanto as restrições são velocidade de vibração nos mancais, a deflexão do estator e a tensão mecânica experimentada pelo mecanismo causada pela onda de pressão girante no estator. O método de PSO (Particle Swarm Optimization) modificado para considerar as restrições é utilizado para solucionar o problema de otimização. Uma comparação entre os níveis de vibração do componente original e do otimizado é realizada para mostrar a eficácia da solução. Como resultado, o mecanismo proposto mostrou-se eficiente, reduzindo a frequência natural do sistema e atendendo às restrições. A metodologia de otimização com modelos substitutos apresentada se mostrou acurada e ágil, auxiliando na tomada de decisão durante o projeto de componentes.

Palavras-chave: Kriging. Modelo substituto. Máquina elétrica girante. Vibração em duas vezes a frequência da rede. Transmissibilidade de vibração.

ABSTRACT

One of the main concerns in the design of a rotating electric machine (REM) is the vibration associated to twice the line frequency. When a REM is operating, that vibration from stator is transmitted to the frame and to the bearings. If the vibration values are above the requested, the REM is immediately turned off. Aiming to reduce this kind of vibration, manufacturers design mechanical components to be installed in the interface between the stator and the frame. These components aims to meet vibration limits defined by international standards like IEC 60034 and NEMA MG 1, as well as market requests. Experimental procedures to reduce the vibration levels during the design phase are costly and require a large amount of time. Thus, numerical simulation and optimization procedures are used to avoid such inconveniences and to reduce the overall cost during the design phase. Nonetheless, the optimization requires a large amount of simulations which can also be very costly. This work aims to define a new mechanical component to attenuate the transmissibility of vibration between the stator and the frame in large REMs, by changing the natural frequency of the system and/or reducing vibration amplitude in that region. The design involves a parametric optimization and the use of surrogate models based on Kriging theory to decrease the overall computational time. Surrogate models give agility to engineering and optimization process, decreasing the number of evaluations of numerical simulations with high computational cost. For the mechanism design, an optimization problem is defined and through surrogate models the objective functions and constraints are approximated. The objective function is used to adjust the natural frequency of system to a given value, while the constraints are bearings vibration velocity, the stator deflection and the mechanical stress caused by the rotating pressure wave from the stator. The PSO (Particle Swarm Optimization) method with constraints handling is used to solve the optimization problem. A comparison between the vibration levels of the original and the optimized component is performed to show the efficacy of the solution. As a result, the proposed mechanism proved to be efficient, decreasing the natural frequency of the system and meeting the constraints. The methodology of optimization with surrogate models presented proved to be accurate and agile, assisting in decision making during components design.

Key-words: Kriging. Surrogate model. Rotating electric machine. Twice line frequency vibration. Transmissibility of vibration.

LISTA DE FIGURAS

Figura 1 – Circuito esquemático de uma máquina assíncrona.	23
Figura 2 – Circuito esquemático de uma máquina síncrona.	24
Figura 3 – Vista em corte de uma MEG síncrona indicando seus componentes.	25
Figura 4 – Pacote de chapas do estator com bobinas de cobre inseridas nas ranhuras	26
Figura 5 – A distância entre o diâmetro externo do rotor e o diâmetro interno do estator é conhecido como entreferro.	27
Figura 6 – Pressão magnética radial agindo sobre o dente do estator.	28
Figura 7 – Comportamento da pressão magnética e do campo magnético em função do ângulo ν	29
Figura 8 – Gráficos representando o modo de vibrar do estator em um determinado instante de tempo t , para uma MEG de 2 polos (a) e de 4 polos (b).	30
Figura 9 – Exemplo de <i>string</i> com valores binários.	35
Figura 10 – Exemplo de <i>crossover</i> de um ponto (a), <i>crossover</i> de dois pontos (b) e <i>crossover</i> uniforme (c).	36
Figura 11 – Exemplo de antes e após a etapa de mutação.	36
Figura 12 – Fluxograma do algoritmo genético (GA).	37
Figura 13 – Vetor resultante (linha contínua em azul) dos vetores da Eq. (17).	38
Figura 14 – Fluxograma do algoritmo PSO.	39
Figura 15 – Modelo de uma distribuição Gaussiana univariada conforme Eq. (27).	43
Figura 16 – Densidade de probabilidade para a Eq. (48).	46
Figura 17 – Probabilidade marginal para z_1 (à esquerda) e z_2 (à direita).	47
Figura 18 – Probabilidade condicional para $\bar{z}_2 = 3$ e $z_1 \leq 5$	48
Figura 19 – Relação de monotocidade crescente da função L com $\ln L$	49
Figura 20 – $\hat{\mu}$ e $\hat{\sigma}$ maximizam a probabilidade de representar $\mathbf{z} = \{3, 5, 9\}$, assumindo uma distribuição Gaussiana.	50
Figura 21 – Processo Gaussiano à <i>priori</i> unidimensional, onde $x \in (0, 10)$, para 5 realizações de um evento.	53
Figura 22 – Processo Gaussiano à <i>posteriori</i> para a função $Z(x) = \sin(x)$, onde $x \in (0, 10)$, para 5 realizações de um evento e 5 pontos de suporte.	54
Figura 23 – Processo Gaussiano <i>posteriori</i> para a função $Z(x) = \sin(x)$ com ruído, onde $x \in (0, 10)$, para 5 realizações de um evento e 11 pontos de suporte.	55
Figura 24 – Influência dos parâmetros l , σ_k e σ_z^2 na aproximação da função $z(x) = \sin(x)$	56
Figura 25 – Comportamento do <i>kernel</i> da Eq. (84) em relação ao parâmetro p	58
Figura 26 – Comportamento do <i>kernel</i> da Eq. (84) em relação ao parâmetro θ	59
Figura 27 – Predições da Eq. (99) em função de θ com 3 pontos de suporte.	61
Figura 28 – Exemplo de amostragem com duas variáveis (x_1 e x_2) utilizando o método de fatoração completa.	62
Figura 29 – Exemplo de plano amostral com três variáveis (x_1 , x_2 e x_3) utilizando o método do hiper-cubo latino. Projeção tridimensional (a) e suas respectivas projeções bidimensionais (b), (c) e (d).	63
Figura 30 – Comparação de amostragem com duas variáveis x_1 e x_2 , utilizando método do hiper-cubo latino puro (a) e otimizado (b).	64
Figura 31 – Processo de aquisição de novos pontos para a função da Eq. (99) utilizando como função de aquisição o EI.	67
Figura 32 – 6ª iteração para o processo EI aplicado a função da Eq. (99).	68

Figura 33 – Fluxograma do processo de otimização com modelos substitutos.	70
Figura 34 – Diferença da derivada da função de referência da Eq. (113) (com 5 pontos de suporte) e a derivada da predição de Kriging da Eq. (114).	71
Figura 35 – Problema referência de treliça com 3 barras encontrado em (ARORA, 2004).	72
Figura 36 – Gráfico de contorno da função objetivo e das restrições. O ponto preto representa o <i>design</i> ótimo, $A_1^* = 6,31$ [in ²] e $A_2^* = 3,26$ [in ²], enquanto a área cinza representa a região inviável.	73
Figura 37 – Comparativo entre os modelos substitutos da função g_2 utilizando os métodos s^2 (a) e PI (b), mostrando a disposição dos pontos de suporte (círculos branco) para do cenário A do teste 5.	75
Figura 38 – Comparativo de s^2 entre os modelos substitutos da função g_1 para o cenário B do teste 4. Considerando os métodos EI (a), s^2 (b), LCB (c) e PI (d), após 20 pontos de suporte adquiridos.	76
Figura 39 – Sistema com dois graus de liberdade representando a transmissão da vibração entre estator e carcaça.	77
Figura 40 – Espectro de velocidade de vibração no diâmetro externo no estator, na direção \bar{x} (a) e na direção \bar{y} (b).	78
Figura 41 – Espectro de velocidade de vibração no mancal dianteiro, na direção \bar{x} (a) e na direção \bar{y} (b).	79
Figura 42 – Interface entre estator e carcaça em vermelho, constituída por meio de contra placas (a) e de longarinas (b).	80
Figura 43 – Mecanismo de (HWANG <i>et al.</i> , 2010).	81
Figura 44 – Mecanismo proposto (em vermelho) instalado nas extremidades do estator.	82
Figura 45 – Principais pontos medidos durante a ODS (<i>Operating Deflection Shape</i>).	83
Figura 46 – Diagrama de cores apresentando o padrão de vibração na frequência de 114 [Hz]. Os pontos 1,2,3 e 4 representam as mesmas posições que os pontos da Fig. 45.	84
Figura 47 – Modo de vibrar encontrado pelo modelo numérico computacional.	85
Figura 48 – Corte transversal do mecanismo proposto com a localização das variáveis de projeto x_1 , x_2 e x_3	85
Figura 49 – Comportamento da função de aquisição EI para $\hat{g}_1, \hat{g}_2, \hat{g}_4$ e \hat{g}_5 . As outras duas aproximações do problema de otimização foram omitidas pois a participação do EI no número de pontos de suporte é desprezível ou nulo.	88
Figura 50 – Comportamento do modelo substituto da função objetivo \hat{y} em função de x_2 e x_3 para $x_1 = 70$ [mm].	89
Figura 51 – Comportamento dos modelos substitutos das funções de restrição \hat{g}_1 (à esquerda) e \hat{g}_2 (à direita), ambas em função de x_2 e x_3 para $x_1 = 70$ [mm].	90
Figura 52 – Comportamento do modelo substituto da função de restrição \hat{g}_3 em função de x_2 e x_3 para $x_1 = 50$ [mm].	91
Figura 53 – Comportamento do modelo substituto das funções de restrição \hat{g}_4 (à esquerda) \hat{g}_5 (à direita), ambas em função de x_2 e x_3 para $x_1 = 70$ [mm].	92
Figura 54 – Modelo da viga com a pressão magnética $\sigma_r(\nu, t)$ distribuída na aba do mecanismo proposto.	92
Figura 55 – Modelo da viga com a pressão magnética $\sigma_r(\nu, t)$ concentrada na aba do mecanismo proposto.	93
Figura 56 – Tensão mecânica no mecanismo do lado acionado (à esquerda) e do lado não-acionado (à direita).	93

Figura 57 – Análise harmônica computacional com a velocidade de vibração no mancal dianteiro na direção \bar{x}	94
Figura 58 – Análise harmônica computacional com a velocidade de vibração no mancal diateiro na direção \bar{y}	95
Figura 59 – Análise harmônica computacional com a velocidade de vibração no mancal traseiro na direção \bar{x}	95
Figura 60 – Análise harmônica computacional com a velocidade de vibração no mancal traseiro na direção \bar{y}	96

LISTA DE TABELAS

Tabela 1 – Limite de vibração da norma IEC 60034-14 para $H_{IEC} > 280$, onde H_{IEC} é a dimensão de altura de centro conforme (IEC, 2003).	31
Tabela 2 – Limite de vibração da norma NEMA MG-1 para $H_{NEMA} > 440$, onde H_{NEMA} é a dimensão de altura de centro conforme (NEMA, 2016).	31
Tabela 3 – Dados do problema de otimização - adaptado de (ARORA, 2004).	72
Tabela 4 – \mathbf{x}^* para o cenário B do teste 2.	73
Tabela 5 – Média aritmética do ótimo para cada função de aquisição dividida por $\mathbf{x}^* = \{6.31, 3.26\}$ para o cenário A.	74
Tabela 6 – Principais características da MEG do estudo de caso.	81
Tabela 7 – Equipamentos e <i>softwares</i> utilizados para análise modal experimental.	81
Tabela 8 – Tabela com o valor máximo e mínimo de cada variável de projeto.	83
Tabela 9 – Valor das funções que compõe o problema da Eq. (119) em \mathbf{x}^* e o número de pontos de suporte adquiridos pela função EI.	87
Tabela 10 – Resultados do teste 1.	122
Tabela 11 – Resultados do teste 2.	123
Tabela 12 – Resultados do teste 3.	124
Tabela 13 – Resultados do teste 4.	125
Tabela 14 – Resultados do teste 5.	126

SUMÁRIO

1	INTRODUÇÃO	19
1.1	MOTIVAÇÃO	19
1.2	OBJETIVOS	20
1.3	SEQUÊNCIA DO TRABALHO	20
2	MÁQUINA ELÉTRICA GIRANTE (MEG)	23
2.1	CONSTRUÇÃO DE UMA MEG	24
2.2	FONTES DE VIBRAÇÃO EM MEGs	26
2.3	NORMAS PARA LIMITES DE VIBRAÇÃO EM MEG	29
3	OTIMIZAÇÃO	33
3.1	ALGORITMO GENÉTICO - <i>GENETIC ALGORITHM</i> (GA)	35
3.2	OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS - <i>PARTICLE SWARM OPTIMIZATION</i> (PSO)	37
3.3	CONSIDERAÇÕES DE RESTRIÇÕES EM TÉCNICAS POPULACIONAIS	38
4	UMA VISÃO GERAL DE PROBABILIDADE	41
4.1	VARIÁVEIS ALEATÓRIAS E FUNÇÕES DE PROBABILIDADE	41
4.2	VETOR ALEATÓRIO E DISTRIBUIÇÃO MULTIVARIADA	44
4.3	VEROSSIMILHANÇA	47
5	PROCESSO ESTOCÁSTICO, PROCESSO GAUSSIANO E PREDITOR	51
5.1	PROCESSOS ESTOCÁSTICOS	51
5.2	PROCESSO GAUSSIANO	52
5.3	VALOR ESPERADO CONDICIONAL COMO PREDITOR	55
6	KRIGING, AMOSTRAGEM INICIAL E AQUISIÇÃO DE PONTOS DE SUPORTE	57
6.1	KRIGING	57
6.2	AMOSTRAGEM INICIAL	61
6.3	AQUISIÇÃO DE PONTOS DE SUPORTE	64
7	MATERIAIS E MÉTODOS	69
8	FORMULAÇÃO DO PROBLEMA E ESTRATÉGIA DE SOLUÇÃO	77
9	RESULTADOS	87
10	CONCLUSÃO	97
10.1	TRABALHOS FUTUROS	99
	REFERÊNCIAS	101
	APÊNDICE A – FUNÇÃO DENSIDADE DE PROBABILIDADE MULTIVARIADA	105
	APÊNDICE B – VALOR ESPERADO E COVARIÂNCIA CONDICIONAL	109

APÊNDICE C – DERIVADA DA FUNÇÃO DENSIDADE DE PROBABILIDADE UNIVARIADA EM RE- LAÇÃO A μ E A σ^2	111
APÊNDICE D – DERIVADA DA FUNÇÃO DENSIDADE DE PROBABILIDADE MULTIVARIADA EM RE- LAÇÃO A μ E A σ^2	113
APÊNDICE E – MATRIZ INVERSA PARTICIONADA	115
APÊNDICE F – DESENVOLVIMENTO MATEMÁTICO PARA A PREDIÇÃO DE KRIGING	117
APÊNDICE G – DESENVOLVIMENTO MATEMÁTICO PARA O EI	119
APÊNDICE H – RESULTADOS DO PROBLEMA EXEMPLO DA METODOLOGIA	121

1 INTRODUÇÃO

Equipamentos e dispositivos de alta eficiência são cada vez mais demandados pela indústria em geral. Para os fabricantes de máquinas elétricas girantes (MEGs) essas solicitações se traduzem em alto rendimento energético, baixo nível de ruído e de vibração. O alto rendimento energético é requisitado para atingir um maior retorno financeiro com menos perdas. Um baixo nível de ruído está relacionado ao conforto das pessoas que trabalham próximo ao equipamento, enquanto a redução dos níveis de vibração garante uma vida útil maior ao equipamento, evitando um comportamento danoso à MEG. Para atender aos requisitos associados a vibração, os projetos de MEGs devem seguir normas internacionais regulamentadoras tais como IEC-60034 (IEC, 2003), NEMA MG-1 (NEMA, 2016) e API-541 (API, 2014).

A vibração pode ser prejudicial tanto para a MEG quanto para os equipamentos a qual ela é aplicada. Dessa forma, frequências naturais próximas a fontes excitadoras devem ser evitadas com o intuito de reduzir a velocidade de vibração em seus componentes. O conhecimento da sensibilidade dos parâmetros modais e seus comportamentos em uma MEG é de suma importância para obter valores de vibração estipulados por norma.

A geometria e funcionamento complexos de uma MEG limitam o número de ferramentas para determinar seu comportamento modal característico. No entanto, a utilização de *software* CAE (*Computer Aided Engineering*) permite a realização de análise modal com elevada precisão, podendo antever situações críticas ainda durante a fase da concepção do produto. Portanto, a fase de projeto desse tipo de equipamento tem importante papel no ciclo de vida do produto, utilizando ferramentas para garantir o atendimento dos requisitos de normas internacionais e do mercado.

1.1 MOTIVAÇÃO

É possível encontrar diversas MEGs no cotidiano, nas mais variadas aplicações, acionando uma máquina de cortar grama ou um compressor de geladeira. Também podemos enfatizar os motores de indução de uso industrial que acionam esteiras, extrusoras, ventiladores, etc. A energia gerada para residências através de hidro geradores, diesel geradores ou turbo geradores são exemplos de MEGs de grande porte. Bem como motores de indução de média tensão que acionam compressores recíprocos nas refinarias, moinhos de bolas em mineradoras e bombas para transposição de água.

Toda esta demanda faz a indústria desse tipo de máquinas ser muito competitiva, buscando uma melhor performance nas características eletromagnéticas, térmicas e mecânicas, existindo alternativas tecnológicas para melhorar o desempenho de cada uma dessas características em uma MEG. Seja desenvolvendo um material novo para a construção dos principais componentes ou criando novas tecnologias de isoladores para reduzir a vibração. Devido ao aumento de competitividade, os próprios fabricantes definem critérios de performance mais restritos do que o estabelecido por normas internacionais. Um desses critérios é a vibração, onde se visa a redução de sua transmissibilidade, objetivando uma vida útil maior ao equipamento. Consequentemente, enfatiza-se a importância de minimizar as falhas ou paradas de produção por vibração danosa a uma MEG, como por exemplo a vibração no estator excitada em duas vezes a frequência da rede. Este tipo de vibração é originada no entreferro, passando pelo estator, indo para a carcaça e chegando aos mancais onde é realizada a medição. Dessa forma, destaca-se a inovação nos mecanismos na interface entre o estator e carcaça, os quais visam a redução da transmissibilidade de vibração. Para o desenvolvimento desses mecanismos é imprescindível a obtenção do conhecimento dos

parâmetros modais de uma MEG a fim de evitar frequências naturais que coincidam com excitações inerentes ao funcionamento. Também devido a dinâmica ágil do mercado são necessárias metodologias de desenvolvimento de componentes que permitam o lançamento do produto ou componentes de forma robusta, eficaz e rápida, proporcionando projetos de MEGs que atendam aos requisitos solicitados, com performance otimizada e vantajoso custo-benefício.

1.2 OBJETIVOS

Diante da importância das MEGs para a indústria mundial e aos requisitos de vibração solicitados, esse trabalho tem como objetivo desenvolver um mecanismo para a redução da transmissibilidade de vibração entre o estator e a carcaça. Onde a principal função do mecanismo a ser projetado é reduzir o nível de velocidade de vibração excitado em duas vezes a frequência de rede.

No desenvolvimento do mecanismo proposto é utilizado *software* comercial CAD (*Computer Aided Design*) para prototipagem virtual 3D, *software* CAE para análise em FEM (*Finite Element Method*), *software* de linguagem de programação para o desenvolvimento de modelos substitutos e de algoritmos para otimização da geometria do mecanismo. Portanto, o objetivo principal do presente trabalho é:

- Estudar a utilização de modelos substitutos para otimização paramétrica e aplicá-los no desenvolvimento de um mecanismo para a redução da transmissibilidade de vibração entre o estator e a carcaça de uma MEG.

Também pode-se elencar como objetivos secundários os seguintes pontos:

- Proposta de uma metodologia de projeto de componentes utilizando modelos substitutos;
- Validar a solução ótima através de uma simulação computacional de alto custo.

1.3 SEQUÊNCIA DO TRABALHO

No capítulo inicial dessa dissertação apresenta uma introdução geral sobre os possíveis desafios tecnológicos em uma MEG, principalmente relacionado a sua vibração. Também contém uma descrição dos objetivos deste trabalho, bem como a motivação que originou este estudo.

O Cap. 2 descreve a construção de uma MEG, relacionando seus principais componentes. Vibração em MEGs é estudada de forma mais aprofundada, enfatizando a vibração em duas vezes a frequência da rede. Além disso, descreve os limites de vibração contidos nas normas internacionais reguladoras para a construção de MEGs.

O terceiro capítulo descreve um problema de otimização e sua respectiva solução. Destacando técnicas de otimização populacionais como Algoritmos Genéticos (GA) e Otimização por Enxames de Partículas (PSO). Também é explanado quais considerações são feitas para restrições nos tipos de técnicas apresentadas.

O quarto capítulo traz uma visão global sobre incertezas, introduzindo a teoria de probabilidade a ser utilizada posteriormente. Neste capítulo são apresentadas as medidas estatísticas, uma análise de variável aleatória univariada e multivariada e o conceito de verossimilhança.

O Cap. 5 apresenta o processo estocástico juntamente com o processo Gaussiano e sua relação com medidas estatísticas. Conceitos importantes como a definição de um preditor utilizando a média condicional é demonstrada para a aplicação em modelos substitutos.

O Cap. 6 mostra a forma como modelos substitutos são construídos através da técnica de Kriging, estes modelos serão utilizados diretamente na otimização da geometria do mecanismo proposto. Neste capítulo são apresentadas formas de amostragem inicial, focando em hipercubo latino e sua forma otimizada. Também funções de aquisição para pontos de suportes são descritas e discutidas.

O sétimo capítulo descreve os materiais e os métodos a serem utilizados para atingir os objetivos. Os métodos são aplicados em um problema de teste conhecido na literatura, utilizando a teoria e técnicas apresentadas anteriormente. Neste capítulo é apresentado a metodologia para criar problemas de otimização com restrições através de modelos substitutos.

O Cap. 8 apresenta o problema de construir um mecanismo entre o estator e a carcaça para redução da transmissibilidade de vibração em uma MEG. São demonstradas medições de vibração entre estator e carcaça, apresentando os conceitos de interface atual e o respectivo modelo numérico de uma MEG. Por fim apresenta a forma do mecanismo proposto a ser implementado e o problema de otimização através de modelos substitutos.

O nono capítulo traz os resultados da aplicação da metodologia do Cap. 7 para o mecanismo proposto descrito no Cap. 8, avaliando seu impacto em uma MEG. Trazendo uma análise das funções do problema de otimização por modelos substitutos e dos fenômenos físicos decorrentes de sua implementação.

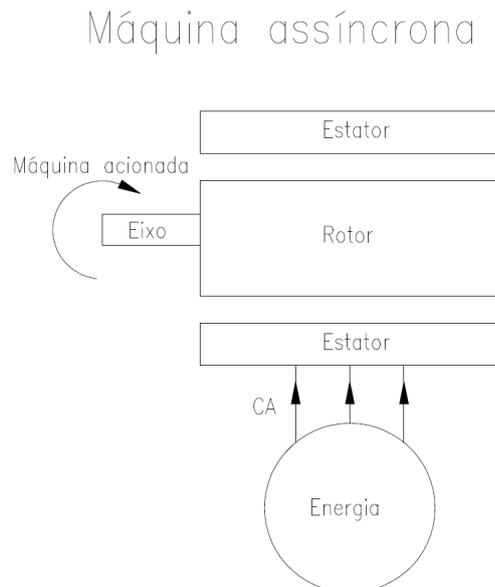
A conclusão do trabalho se encontra no décimo capítulo, além de sugestões para trabalhos futuros. Posteriormente as referências bibliográficas utilizadas no trabalho são apresentadas. Por fim, os apêndices com derivações do equacionamento pertinente ao trabalho são descritos, bem como o detalhamento de resultados de materiais e métodos.

2 MÁQUINA ELÉTRICA GIRANTE (MEG)

Segundo Chapman (2005), uma máquina elétrica girante (MEG) é um transdutor eletro-mecânico capaz de converter energia elétrica em energia mecânica ou energia mecânica em energia elétrica. Essa conversão ocorre por meio do fenômeno conhecido como magnetismo, sendo realizada através da ação de um campo magnético girante. Ainda, de acordo com Fitzgerald, Kingsley e Umans (2003), MEGs estão na categoria de equipamentos de conversão de energia contínua, nos quais se enquadram motores e geradores. Quando o equipamento converte energia elétrica em energia mecânica o mesmo é conhecido como motor, já dispositivos que convertem energia mecânica em energia elétrica são conhecidos como geradores.

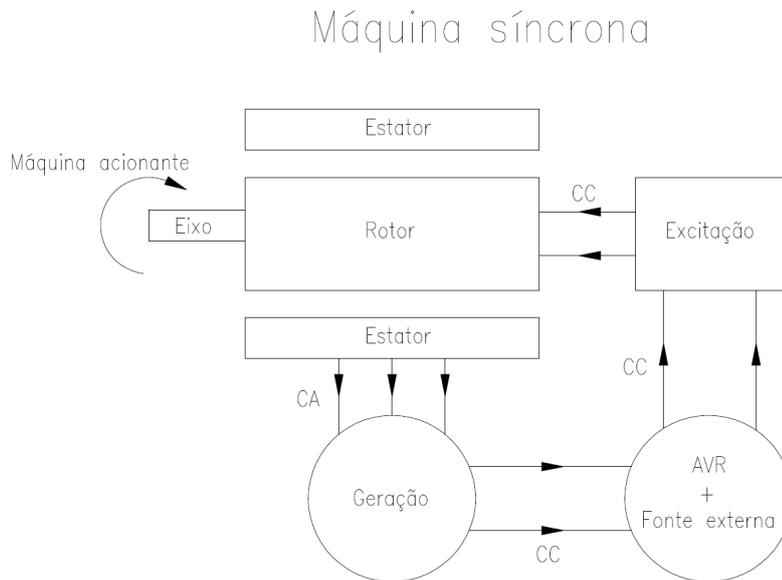
As MEGs podem ser divididas em duas classes: máquinas assíncronas e máquinas síncronas. Máquinas assíncronas são motores e geradores das quais a corrente de campo é fornecida por meio de indução magnética gerada a partir dos enrolamentos do estator, conforme mostrado na Fig 1. De acordo com Krause *et al.* (2013), é o tipo de MEG mais utilizada para converter energia elétrica em mecânica, atuando bombas, moinhos e vários outros dispositivos. Já máquinas síncronas são motores e geradores que tem sua corrente do campo magnético fornecida para o rotor por uma fonte externa independente e de corrente contínua, de acordo com a Fig 2. No caso de geradores síncronos, junto a fonte externa de energia está o AVR (*Automatic Voltage Regulator*), que é um sistema eletrônico responsável pela compensação da tensão elétrica de saída da MEG em relação ao sistema elétrico. Segundo Krause *et al.* (2013), praticamente toda a energia elétrica é advinda de MEGs síncronas, atuadas por turbinas hidráulicas ou a vapor e motores de combustão interna.

Figura 1 – Circuito esquemático de uma máquina assíncrona.



Fonte: Autor.

Figura 2 – Circuito esquemático de uma máquina síncrona.



Fonte: Autor.

2.1 CONSTRUÇÃO DE UMA MEG

De acordo com Wappler (2014), os componentes de uma MEG podem ser divididos em dois tipos: rotativos (girantes) e estacionários, possibilitando a melhor identificação de suas características funcionais e estruturais. A Fig. 3 mostra a vista em corte de uma MEG com seus principais componentes.

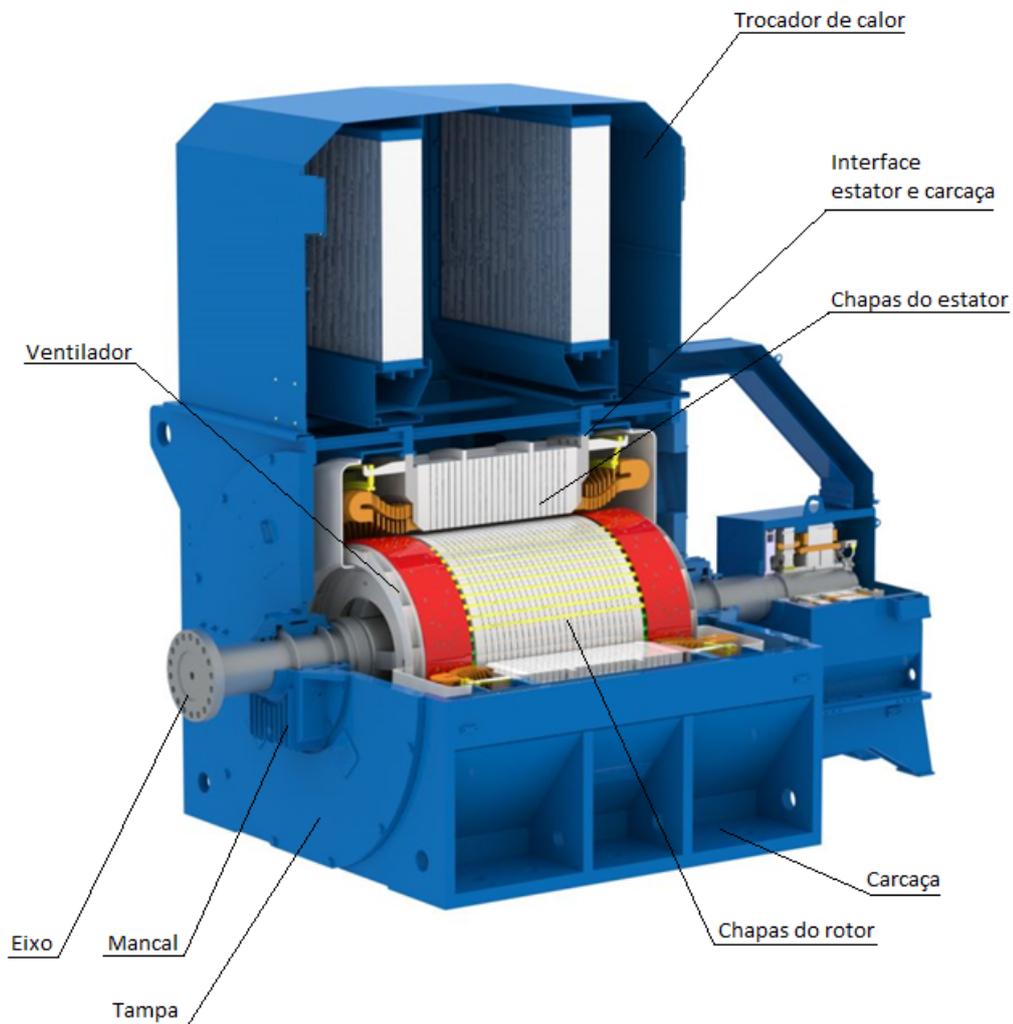
O principal componente rotativo é o rotor, apresentando movimento relativo de rotação em relação a um observador fixo. Nele está o pacote de chapas com forma cilíndrica e ranhuras no diâmetro externo para alocar bobinas de cobre. Segundo Pyrhonen, Jokinen e Hrabovcova (2013), as bobinas do rotor são responsáveis pela interação entre os campos magnéticos e o fluxo de corrente com as bobinas do estator (componente estacionário). A forma construtiva do rotor, sua corrente e tensão elétrica de alimentação definem o modo de operação (motor ou gerador) e a classe da MEG.

Outro componente rotativo importante é o eixo. Nele é fixado o pacote de chapas do rotor e onde é acoplado, através da ponta de eixo, o equipamento a ser acionado, no caso de motores ou o equipamento acionante em caso de geradores. No eixo também é fixado o ventilador do circuito interno de ventilação, responsável por realizar a circulação do ar interno da máquina e conseqüentemente realizar a troca térmica com um trocador de calor.

Os componentes estacionários são aqueles que não apresentam movimento relativo a um observador fixo. Entre eles, pode-se destacar o pacote de chapas do estator, com forma cilíndrica, o qual dispõe de ranhuras no seu diâmetro interno conforme apresentado na Fig. 4. Nas ranhuras são alocadas bobinas de cobre, as quais interagem magneticamente com as bobinas do rotor para o funcionamento de uma MEG. A carcaça é o principal componente estrutural em uma MEG, dando rigidez ao estator. Nela estão fixadas as tampas e os mancais e sua construção pode ser vertical ou horizontal. Outra função da carcaça é a fixação da MEG na base de apoio.

Outros componentes estacionários são as tampas e os mancais: a primeira faz o fechamento

Figura 3 – Vista em corte de uma MEG síncrona indicando seus componentes.



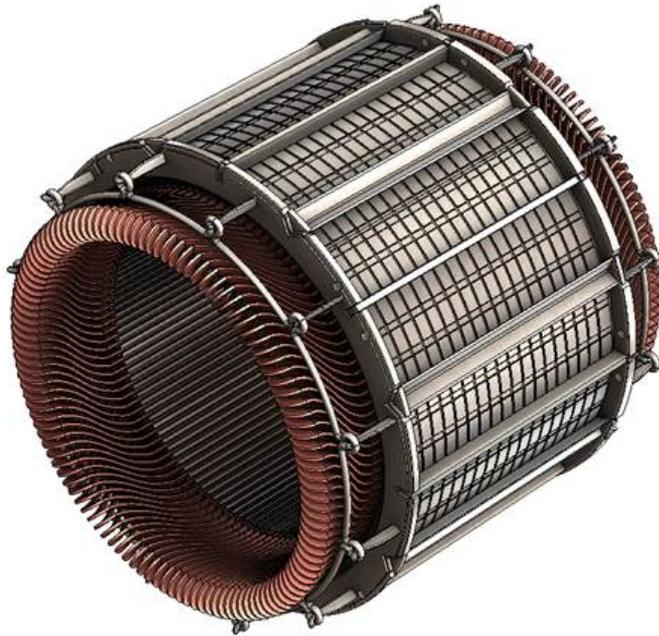
Fonte: Autor.

da carcaça e a vedação, enquanto o segundo tem como função apoiar o rotor (componente rotativo) e onde é realizada a medição de vibração em uma MEG. Os mancais tem sua posição definida em lado acionado (LA), onde está a ponta de eixo, e lado não acionado (LNA), sendo o lado oposto ao acionado.

As MEGs geram aquecimento devido suas perdas de origem mecânica (atrito com o ar através da ventilação, calor dos mancais...) e de origem eletromagnética (no ferro, Joule, efeito pelicular...). Em muitas MEGs essa característica não permite a realização da transferência de calor somente por convecção natural sem danificá-la. Para realizar a transferência de calor para outro meio, um trocador de calor é acoplado à carcaça. Os mais comuns são trocadores de calor do tipo ar-água, ar-ar e H₂-água.

A região de interface entre o principal componente rotativo, o rotor, e o principal componente estacionário, o estator, é conhecida como entreferro. É onde ocorrem as interações magnéticas entre esses componentes e como é constituído pelo ar representa uma parte não magnética do circuito eletromagnético. Isso torna o entreferro uma característica construtiva indesejada já que aumenta a resistência ao campo magnético, devido a baixa

Figura 4 – Pacote de chapas do estator com bobinas de cobre inseridas nas ranhuras



Fonte: Autor.

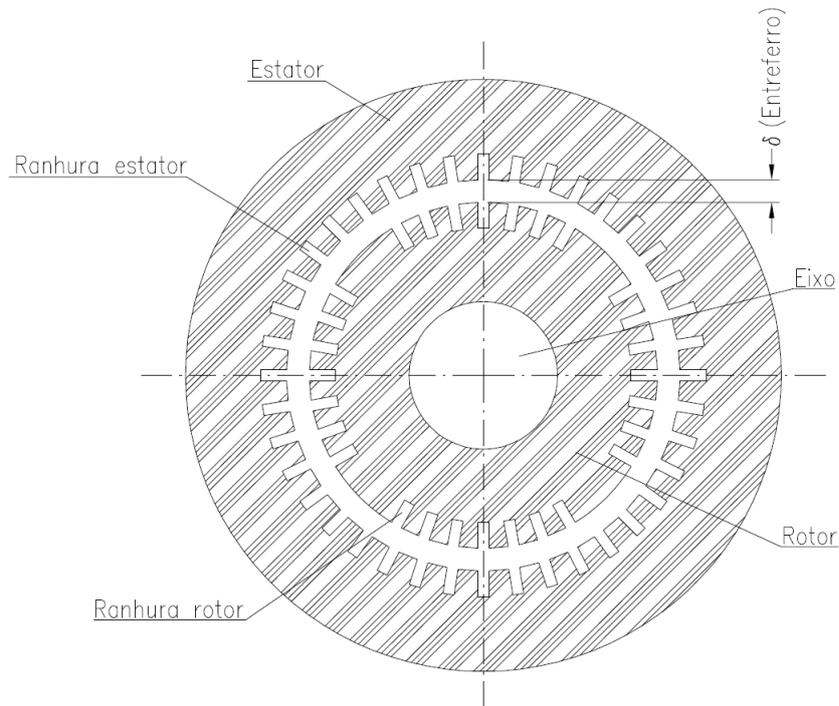
permeabilidade magnética do ar. Mas do ponto de vista mecânico é necessário, já que separa a parte girante da parte estática evitando o contato mecânico entre ambos. Portanto, seu correto dimensionamento é importante, pois tem interferência na magnetização requerida, podendo gerar perdas de origem magnética. Ainda, deve ter um comprimento radial que não permita interferência mecânica entre estator e rotor devido a deflexão radial do eixo, conforme Fig. 5.

2.2 FONTES DE VIBRAÇÃO EM MEGs

Em geral, vibrações severas induzem falhas, desgaste e fadiga aos componentes de máquinas. Esse comportamento não é diferente em MEGs, onde a vibração pode ser oriunda de diversas fontes. Ter conhecimento da origem da vibração facilita o diagnóstico de falha, quando ações corretivas para a redução dos níveis de vibração são tomadas. Para melhor identificação de sua origem, deve-se analisar o espectro de vibração, que contém características intrínsecas ao problema a ser diagnosticado, sendo que cada fonte de vibração possui um espectro característico. Finley, Hodowanec e Holter (1999) classificam duas fontes de vibração que podem ocorrer em uma MEG: as de origem mecânica e as de origem eletromagnética. Finley, Hodowanec e Holter (1999), Taylor (2005), Baccarini (2005) e Costello (1990) explanam a ocorrência dessas fontes de vibração.

Uma dessas fontes é a vibração em duas vezes a frequência da rede, a qual possui origem eletromagnética, estando relacionada ao fluxo magnético fundamental na região do entreferro. O fluxo magnético fundamental é proveniente da onda de indução magnética fundamental, a qual possui a frequência da rede elétrica como referência. Por sua vez, esse fluxo magnético induz uma onda girante de pressão, que varia com o quadrado da indução do campo magnético, a qual tem sua frequência característica em duas vezes a frequência da rede, gerando deformações dinâmicas no estator. A forma espacial e a

Figura 5 – A distância entre o diâmetro externo do rotor e o diâmetro interno do estator é conhecido como entreferro.



Fonte: Autor.

distribuição desta pressão magnética ao longo do perímetro do entreferro depende do número de polos da MEG, mas esse fenômeno ocorre independente de sua polaridade. Sabe-se que para motores de 2 polos a vibração em duas vezes a frequência da rede é mais severa, diminuindo a medida que se aumenta o número de polos. Finley, Hodowanec e Holter (1999) apresentam um comparativo para máquinas de 1000 [HP], de 2 e de 4 polos, onde o primeiro tipo apresentou uma velocidade de vibração de quatro a seis vezes maior do que o segundo. Finley, Hodowanec e Holter (1999) também afirmam que este tipo de vibração é extremamente sensível a rigidez da carcaça e da base onde está instalado o motor, tal como a planicidade dos pés da carcaça, uniformidade do entreferro e excentricidade do rotor.

A frequência angular do campo magnético f_e é definida por Chapman (2005) como

$$f_e = \frac{2f_r}{p_m}, \quad (1)$$

onde f_r em [rad/s] é a frequência angular de alimentação da rede elétrica e p_m é o número de polos da máquina. Dessa maneira, para uma MEG de 2 polos, a frequência angular do campo magnético fundamental é a mesma da rede elétrica de alimentação, enquanto para uma máquina de 4 polos a frequência do campo magnético é metade da frequência da rede elétrica de alimentação. Considerando um entreferro com geometria uniforme, Nau (2007) representa a indução magnética $b(\nu, t)$ na forma

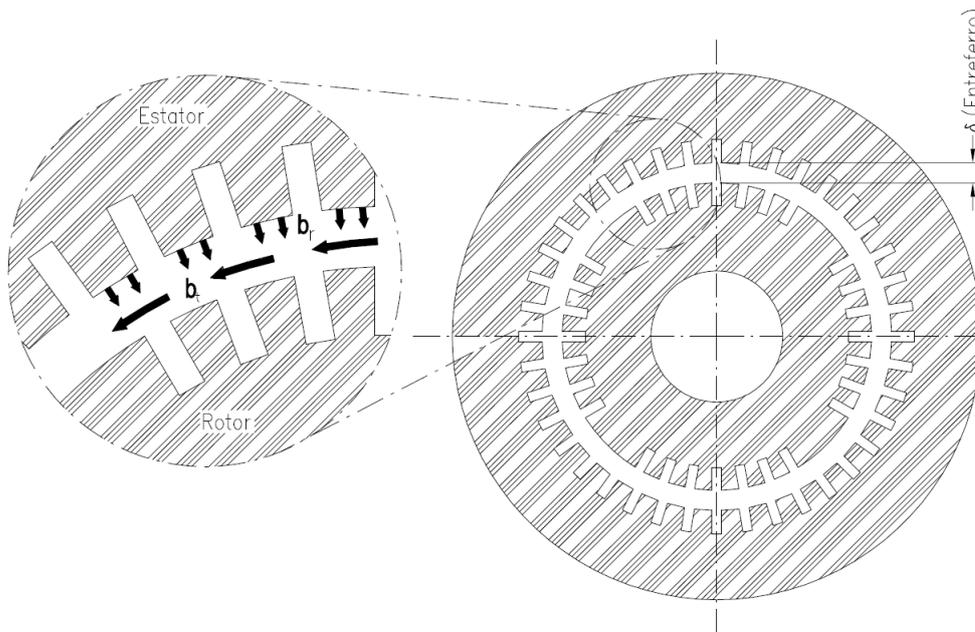
$$b(\nu, t) = \beta \cos(p_p \nu - f_r t - \phi), \quad (2)$$

onde β em $[T]$ é a amplitude da indução do campo magnético, p_p é o número de pares de polos da MEG, ν em $[rad]$ é a posição espacial angular que varia de 0 a 2π , t em $[s]$ é o tempo e ϕ em $[rad]$ é o ângulo de fase da harmônica. Sendo assim, a partir do tensor de Maxwell pode-se descrever a pressão radial σ_r em $[Pa]$ gerada pela indução do campo magnético girante que atua sobre o estator como (VALAVI; NYSVEEN; NILSSEN, 2012)

$$\sigma_r(\nu, t) = \frac{1}{2\mu_0}(b_r^2 - b_t^2), \quad (3)$$

onde μ_0 em $[H/m]$ é a permeabilidade magnética do ar, b_r e b_t são as componentes radial e tangencial, respectivamente de $b(\nu, t)$, também conhecidas como ondas de induções magnéticas radiais e tangenciais, conforme Fig. 6. A pressão magnética radial, originária de b_r , é responsável por atuar ortogonalmente ao eixo axissimétrico da MEG, causando deformação no estator e conseqüentemente vibrações indesejadas.

Figura 6 – Pressão magnética radial agindo sobre o dente do estator.



Fonte: Autor.

Tampion (1990) apresenta um modelo mínimo das componentes da indução magnética agindo no entreferro, com permeabilidade magnética infinita no rotor e no estator. Este modelo enfatiza a predominância das ondas de indução magnética radiais em relação as tangenciais na pressão magnética radial. Sutthiphornsombat (2010) afirma que na fronteira do entreferro com o estator as ondas de indução tangenciais são quase nulas devido a alta permeabilidade magnética da chapa do estator. Valavi, Nysveen e Nilssen (2012) mostram a influência da indução magnética tangencial na pressão magnética radial resultante em um gerador síncrono. Também discutem o aumento da pressão magnética radial com o aumento da corrente elétrica e afirmam que não há alteração na ordem da magnitude de σ_r mesmo desconsiderando a onda de indução tangencial b_t . Admitindo somente as ondas

de indução radial atuando na região do entreferro, a Eq. (3) pode ser reescrita como

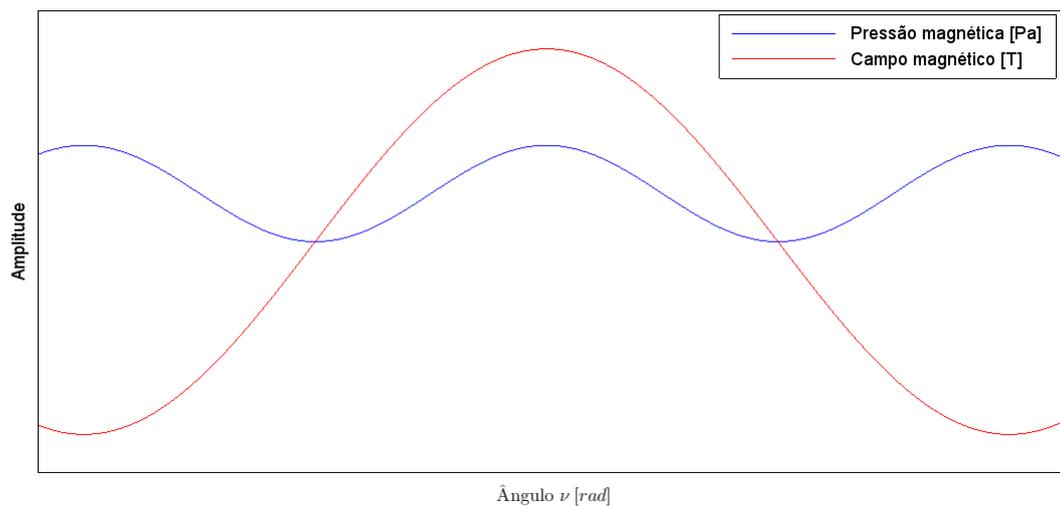
$$\sigma_r(\nu, t) = \frac{1}{2\mu_0} b_r^2. \quad (4)$$

Substituindo a Eq. (2) na Eq. (4) e utilizando a relação trigonométrica $\cos(\alpha)^2 = \frac{1}{2} [1 + \cos(2\alpha)]$, a pressão magnética radial no entreferro é expressa por

$$\sigma_r(\nu, t) = \frac{\beta^2}{4\mu_0} [1 + \cos(2p_p\nu - 2f_r t - 2\phi)]. \quad (5)$$

A Eq. (5) mostra que a onda de pressão magnética radial ocorre com frequência de duas vezes a frequência da rede elétrica de alimentação. Este fenômeno é decorrente da parcela $2f_r$, enquanto o produto $2p_p\nu$ representa a distribuição desta pressão magnética no entreferro, demonstrando assim o modo de vibrar do estator e sua respectiva deformação mecânica.

Figura 7 – Comportamento da pressão magnética e do campo magnético em função do ângulo ν .



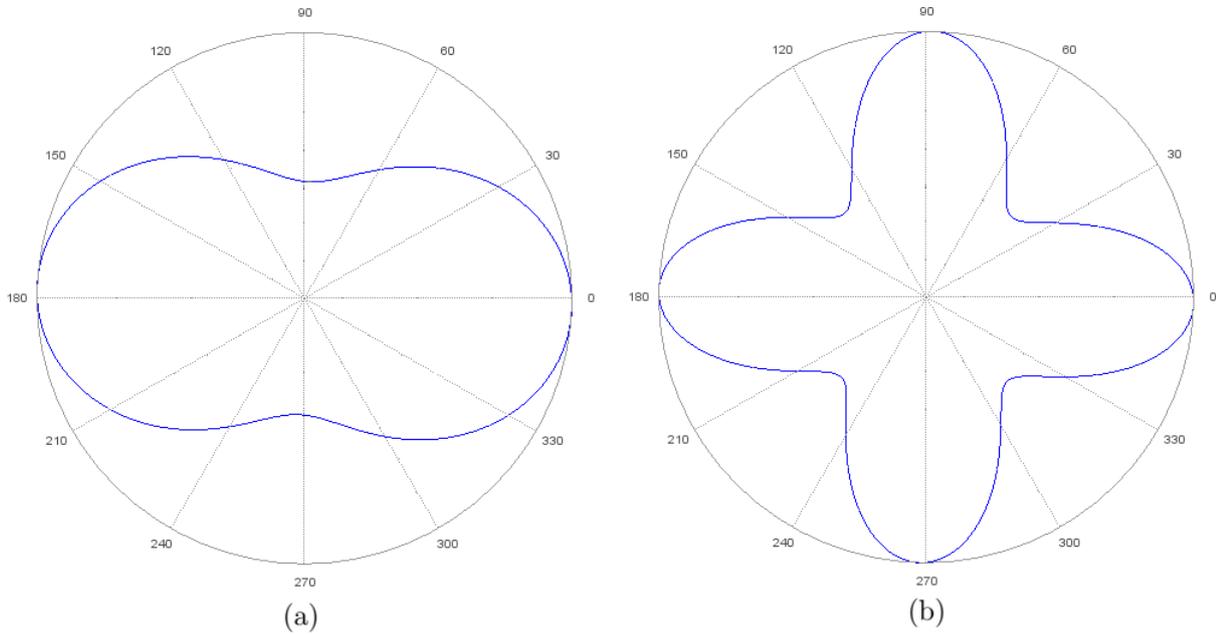
Fonte: Autor.

A Fig. 7 apresenta a variação da pressão magnética e do campo magnético para $\nu \in [0, 2\pi]$. Onde é possível visualizar a pressão magnética variando com o dobro da frequência do campo magnético, denotada pela parcela $2f_r$ da Eq. (5). A Fig. 8 mostra a onda de pressão magnética agindo no estator em direção ao rotor, representando o modo de vibrar do estator para $p_p = 1$, ou seja, em uma MEG de 2 polos. Pode-se avaliar o comportamento elíptico da pressão magnética conforme descrito na Eq. (5). Para efeito comparativo também é apresentado a forma da onda de pressão magnética para uma MEG de 4 polos ($p_p = 4$).

2.3 NORMAS PARA LIMITES DE VIBRAÇÃO EM MEG

Existem normas internacionais que limitam os níveis de vibração nos componentes das MEGs. Elas tem como objetivo evitar que projetos atinjam níveis de vibração elevados

Figura 8 – Gráficos representando o modo de vibrar do estator em um determinado instante de tempo t , para uma MEG de 2 polos (a) e de 4 polos (b).



Fonte: Autor.

a ponto de prejudicar seu funcionamento. As normas mais conhecidas e solicitadas nesse caso são a IEC-60034 (IEC, 2003), a NEMA MG-1 (NEMA, 2016) e a API-541 (API, 2014).

A norma IEC (2003) cita três parâmetros para medição de vibração: deslocamento, velocidade e aceleração. Esses parâmetros são medidos nos mancais, ou próximo deles, através de transdutores de vibração. Na prática, o parâmetro mais usual é a velocidade de vibração, devendo ser realizada a medição em banda larga sobre um espectro de frequência que varia de 10 Hz à 1000 Hz e valores em r.m.s (*root mean square*). Essa norma também determina que motores assíncronos, principalmente de 2 polos, podem registrar oscilação de vibração com uma frequência de duas vezes a frequência de rede. Nesses casos a severidade de vibração v_{vrms} em mm/s deve ser calculada com a relação

$$v_{vrms} = \sqrt{\frac{v_{vmax}^2 + v_{vmin}^2}{2}}, \quad (6)$$

onde v_{vmax} é o máximo valor eficaz (ou r.m.s) do parâmetro medido de vibração em um ciclo completo, enquanto v_{vmin} é o mínimo valor eficaz (ou r.m.s.) do parâmetro medido de vibração em um ciclo completo. Por se caracterizar somente como uma montagem rígida devido ao porte de MEGs de média e alta tensão, a vibração nos pés da carcaça não pode exceder 25% da velocidade máxima medida nos mancais. Este valor é válido para a vibração na frequência de rotação da máquina e vibração em duas vezes a frequência da rede de alimentação. Por fim, os limites da severidade de vibração nos mancais são determinados pela norma de acordo com altura do eixo em mm , grau de vibração e tipo de montagem. De acordo com a Tab. 1, o valor de velocidade de vibração usual para MEGs de média e alta tensão é de $2,3 mm/s$, quando não há exigência especial de vibração.

Tabela 1 – Limite de vibração da norma IEC 60034-14 para $H_{IEC} > 280$, onde H_{IEC} é a dimensão de altura de centro conforme (IEC, 2003).

Grau de vibração	Tamanho da carcaça Montagem	$H_{IEC} > 280$		
		Desloc. μm	Veloc. mm/s	Acel. mm/s^2
A	Suspensão livre	45	2,8	4,4
	Montagem rígida	37	2,3	3,6
B	Suspensão livre	29	1,8	2,8
	Montagem Rígida	24	1,5	2,4

Tabela 2 – Limite de vibração da norma NEMA MG-1 para $H_{NEMA} > 440$, onde H_{NEMA} é a dimensão de altura de centro conforme (NEMA, 2016).

Grau de vibração	Carcaça Montagem	$H_{NEMA} > 440$			
		Desloc. $mils\ pk-pk$	Veloc. $in/s\ pk-pk$	Acel. $g\ pk-pk$	
A	Resiliente	2,4	0,15	0,61	
	Rígida	1,9	0,12	0,49	
B	Resiliente	1,6	0,10	0,41	
	Rígida	2,4,6 polos	1,6	0,10	0,41
		>8 polos	1,3	0,08	0,33

Adicionalmente, a (IEC, 2003) traz um tópico sobre a velocidade de vibração em duas vezes frequência da rede para máquinas assíncronas. Para MEGs com altura de eixo $H > 280mm$ e Grau de Vibração A o valor de $2,3\ mm/s$ é aumentado para $2,8\ mm/s$, já que há predominância na vibração da componente em duas vezes a frequência da rede.

A norma NEMA (2016) também considera os mesmos três parâmetros de medição que a (IEC, 2003), porém as unidades de medidas estão no sistema imperial e as medições deve ser realizadas em pico-pico. Assim como na IEC (2003), a NEMA (2016) trata como dois tipos de montagens para medição da vibração: resiliente (suspensão livre) e rígida. A Tab. 2 apresenta os limites de vibração, para dados não filtrados, para cada tipo de montagem e grau de vibração.

Essa norma também possui um tópico específico referente a vibração em duas vezes a frequência da rede para motores assíncronos com dois polos, onde há dois critérios para validação:

- O valor de vibração filtrada em duas vezes a frequência da rede não deve exceder 90% do valor da Tab. 2;
- Caso seja excedido, deve-se calcular a velocidade de vibração conforme a Eq. (6), não devendo ultrapassar 80% do valor da Tab. 2.

A norma API (2014) é aplicada para MEGs designadas para o segmento de óleo e gás, e trata vibração de forma mais direta que as outras duas normas já citadas. A API (2014) estabelece que valores de velocidade de vibração não filtrados totais ou em determinada frequência não devem exceder

$$\max \left(1,8\ mm/s; \frac{2,5N_m}{1000} \right), \quad (7)$$

onde N_m em $[rpm]$ representa a máxima rotação de operação e $1,8 [mm/s]$ é medido em zero-pico (valor absoluto).

Para efeito da fixação na base, a API (2014) considera somente base rígida e para motores assíncronos dois polos a medição deve ocorrer após a estabilização da temperatura do mancal. Posteriormente a essa estabilização, a medição da vibração filtrada e da não filtrada devem ser realizadas durante 15 minutos.

3 OTIMIZAÇÃO

Otimização consiste na busca por soluções que levem a valores extremos de uma dada função objetivo sujeita a restrições (ARORA, 2004). Sistemas a serem otimizados estão relacionados a biologia, geofísica, economia e engenharias. Um problema de otimização pode ser descrito na forma

$$\begin{aligned}
 & \min_{\mathbf{x} \in S} f(\mathbf{x}) \\
 & h_i(\mathbf{x}) = 0 \quad i = 1, \dots, n_h \\
 & g_j(\mathbf{x}) \leq 0 \quad j = 1, \dots, n_g \\
 & x_i^l \leq x_i \leq x_i^u,
 \end{aligned} \tag{8}$$

onde \mathbf{x} é o vetor de variáveis de projeto, sendo a solução ótima denotada por \mathbf{x}^* , $f(\mathbf{x})$ é a função objetivo a qual queremos minimizar (ou maximizar como $-f(\mathbf{x})$), $h_i(\mathbf{x})$ e $g_j(\mathbf{x})$ são funções de restrição contidas no domínio de soluções S . As restrições podem ser definidas como:

- Restrição lateral (*Box-constraint*): $x_i^l \leq x_i \leq x_i^u$, onde \mathbf{x}^l e \mathbf{x}^u são vetores com limites inferiores e superiores respectivamente;
- Restrição de igualdade: $h_i(\mathbf{x}) = 0$ sendo $i = 1, \dots, n_h$;
- Restrição de desigualdade: $g_j(\mathbf{x}) \leq 0$ sendo $j = 1, \dots, n_g$.

Quando o problema de otimização está relacionado ao cálculo de um estrutura, o mesmo é conhecido como otimização estrutural. Podemos classificar otimização estrutural em:

- Otimização paramétrica: o vetor das variáveis de projeto \mathbf{x} está relacionado a alterações de parâmetros para alcançar a performance ótima de um projeto já definido;
- Otimização de forma: o vetor \mathbf{x} define alteração nos contornos externos e internos de uma da estrutura inicial;
- Otimização topológica: \mathbf{x} descreve a distribuição de um ou mais materiais em um domínio de projeto pré-definido.

Karush (1939) demonstra as condições necessárias para que um \mathbf{x} seja o ótimo de um problema de otimização como da Eq. (8). Posteriormente, Kuhn e Tucker (1951) apresentam de forma mais ampla o que são conhecidas como as condições de Karush-Kuhn-Tucker ou KKT.

As restrições de desigualdade tornam-se restrições de igualdade quando é adicionado uma variável de folga ϑ_j , ou *slack variable* em inglês, a $g_j(\mathbf{x})$. Portanto, restrições do tipo $g_j(\mathbf{x})$ podem ser descritas como restrições de igualdade na forma

$$g_j(\mathbf{x}) + \vartheta_j^2 = 0, \tag{9}$$

sendo a variável de folga sempre positiva ou nula. Quando ϑ_j é nula a restrição é ativa, ou seja \mathbf{x} está sob a fronteira entre a região viável e inviável. Por outro lado, quando $\vartheta > 0$ a restrição é inativa, estando na região viável. Também pode-se definir a função Lagrangiana como

$$\mathbb{L}(\mathbf{x}, \bar{\gamma}, \bar{\beta}, \vartheta) = f(\mathbf{x}) + \sum_{i=1}^{n_h} \bar{\gamma}_i h_i(\mathbf{x}) + \sum_{j=1}^{n_g} \bar{\beta}_j (g_j(\mathbf{x}) + \vartheta_j^2), \quad (10)$$

onde $\bar{\gamma}_i$ e $\bar{\beta}_j$ são multiplicadores de Kuhn-Tucker.

Assim, as condições necessárias de KKT são

$$\frac{\partial \mathbb{L}}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} + \sum_{i=1}^{n_h} \bar{\gamma}_i \frac{\partial h_i(\mathbf{x})}{\partial \mathbf{x}} + \sum_{j=1}^{n_g} \bar{\beta}_j \frac{\partial g_j(\mathbf{x})}{\partial \mathbf{x}}, \quad (11)$$

$$\frac{\partial \mathbb{L}}{\partial \bar{\gamma}} = h_i(\mathbf{x}) = 0, \quad (12)$$

$$\frac{\partial \mathbb{L}}{\partial \bar{\beta}} = g_j(\mathbf{x}) + \vartheta^2 = 0, \quad (13)$$

$$\frac{\partial \mathbb{L}}{\partial \vartheta} = 2\bar{\beta}_j \vartheta_j = 0, \quad (14)$$

$$\vartheta_j^2 \geq 0 \quad \forall j, \quad (15)$$

e

$$\bar{\beta}_j \geq 0 \quad \forall j. \quad (16)$$

Entretanto, as derivadas da Eq. (10) não são possíveis de serem calculadas quando as funções da Eq. (8) são desconhecidas. Além disso, as funções de problemas de engenharia podem ser não-convexas, ou seja, contendo de mais de um mínimo. Assim, justifica-se a utilização de técnicas de otimização que não tendem a ficarem presas em um mínimo local e não utilizem derivadas para encontrar \mathbf{x}^* . Um exemplo são técnicas de otimização populacionais, onde uma coleção de vetores de variáveis de projeto é alocada no domínio S . Diferentemente do que pode ocorrer com técnicas por gradientes, a população de indivíduos espalhados pelo domínio tem a menor tendência de ficar presa a um mínimo local. Portanto, este tipo de técnica apresenta vantagens quando utilizado para buscar o ótimo global de uma função não-convexa por exemplo. As técnicas de otimização populacionais são estocásticas, em sua essência, e baseados em fenômenos naturais como biologia evolutiva no caso dos algoritmos genéticos ou em enxame de partículas como o PSO.

3.1 ALGORITMO GENÉTICO - *GENETIC ALGORITHM* (GA)

Algoritmo genético (GA) foi desenvolvido por Holland (1975) e é uma técnica de otimização populacional baseado na biologia evolutiva, sendo como destaque a seleção natural. Para representar cada indivíduo da população o meio mais comum é por um cromossomo contendo uma *string* binária de tamanho \bar{d} com valores binários conforme mostrado na Fig. 9. Para atingir o ótimo, o GA executa etapas, as quais são: inicialização, seleção, *crossover*, mutação e geração de uma nova população.

A inicialização está relacionada a criação da população de forma aleatória, onde cada cromossomo de cada indivíduo tem sua *string* preenchida com valores binários aleatórios. Nesta etapa deve-se cuidar para que os indivíduos criados aleatoriamente não ultrapassem os limites do domínio S . A escolha do tamanho da população é importante, pois em excesso pode gerar alto custo computacional nas etapas seguintes e a falta de indivíduos pode ocasionar pouca exploração.

A segunda etapa é a seleção, após a avaliação da função objetivo dos indivíduos criados aleatoriamente os mais aptos devem ser selecionados para criar a futura geração (próxima iteração do algoritmo). A seleção deve ocorrer com o objetivo de evitar indivíduos que possuem valores de função objetivo similares, buscando uma heterogeneidade e não criando gerações as quais fiquem presas a um único valor ótimo, que pode ser local. Existem algumas métricas estatísticas para a seleção de indivíduos, Arora (2004) apresenta valores proporcionais a avaliação de função objetivo de cada indivíduo, fazendo uma analogia ao girar uma roleta. Pode-se realizar a seleção através de truncamento, onde é selecionado as melhores avaliações da função objetivo (da menor para o maior em caso de minimização) ou por meio de torneio, onde os cromossomos mais aptos são selecionados de maneira aleatória.

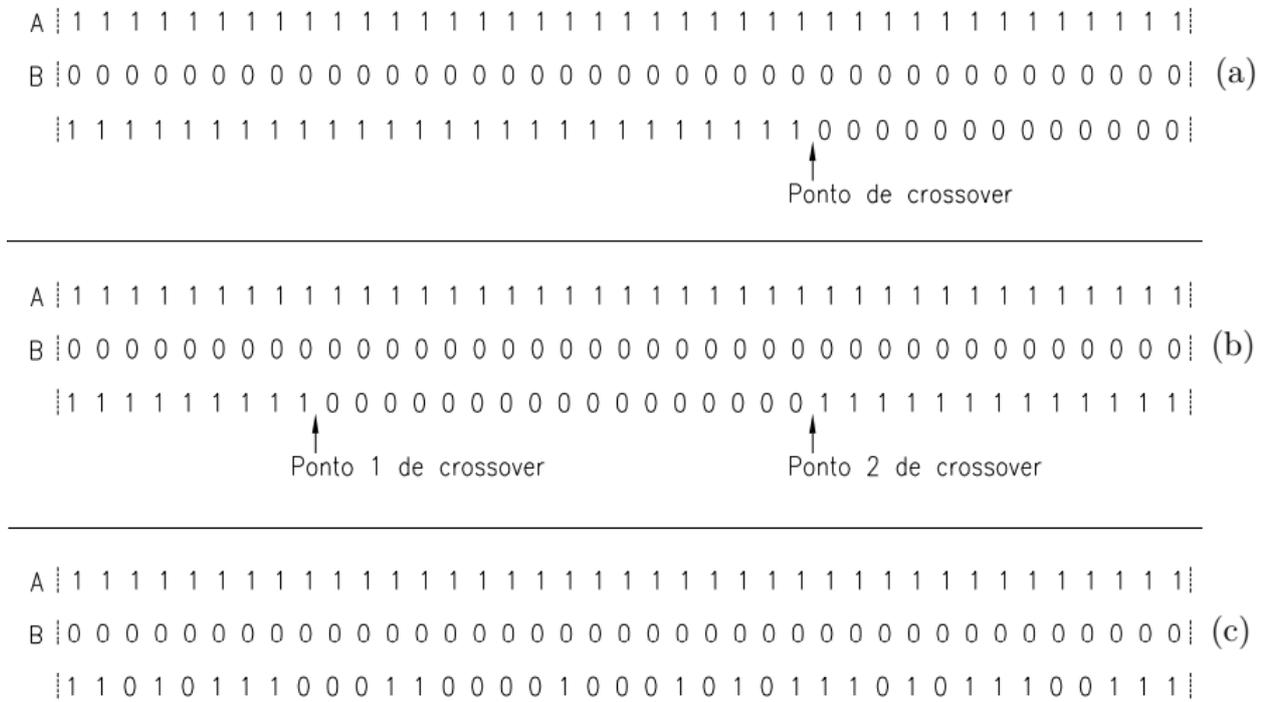
Após a seleção dos indivíduos mais aptos deve-se realizar o *crossover* para a geração da nova população. Esta etapa combina os cromossomos dos indivíduos previamente selecionados podendo ocorrer de diversas formas. Rao (2010) apresenta a mais simples, a qual em um único ponto escolhido aleatoriamente separa a *string* do novo indivíduo em duas porções, a primeira é preenchida com informações do cromossomo do selecionado "A" e a segunda porção do selecionado "B". Existe o *crossover* que utiliza dois pontos, o qual a *string* do novo indivíduo é dividida por dois pontos, tornando-se três porções, sendo essas preenchidas com informações do indivíduo "A" e "B". Por fim, o *crossover* uniforme proporciona para cada informação dos indivíduos "A" e "B" 50% de compor o cromossomo do novo indivíduo, a Fig. 10 apresenta cada uma das maneiras de realizar o *crossover*.

A mutação é a próxima etapa após o *crossover*, ela é destinada para uma melhor exploração do domínio, já que a etapa anterior pode não atingir este o objetivo. De maneira aleatória, os novos indivíduos gerados pelo *crossover* podem ter partes de suas *strings* mudadas a fim de deixar o indivíduo mais apto a função objetivo. A taxa de mutação é um fator importante, podendo levar a um mínimo local caso seja muito pequena bem como reduzir a convergência caso haja uma taxa de mutação muito elevada, a Fig. 11 mostra esquematicamente a mutação da *string* de um indivíduo.

Figura 9 – Exemplo de *string* com valores binários.

‡ 1 1 1 1 1 0 1 1 1 1 1 1 1 0 1 0 1 1 1 0 1 1 0 1 0 1 0 1 1 0 1 0 1 1 0 0 0 ‡

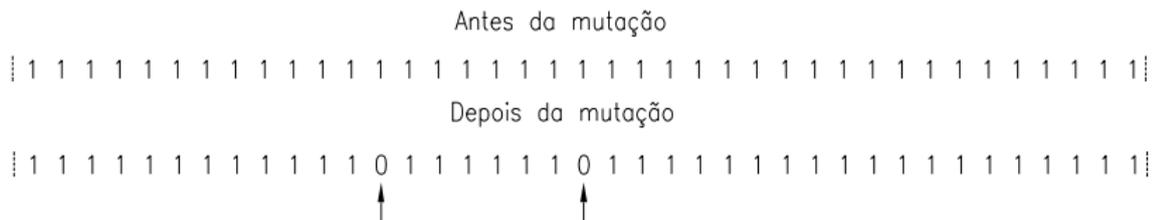
Figura 10 – Exemplo de *crossover* de um ponto (a), *crossover* de dois pontos (b) e *crossover* uniforme (c).



Fonte: Autor.

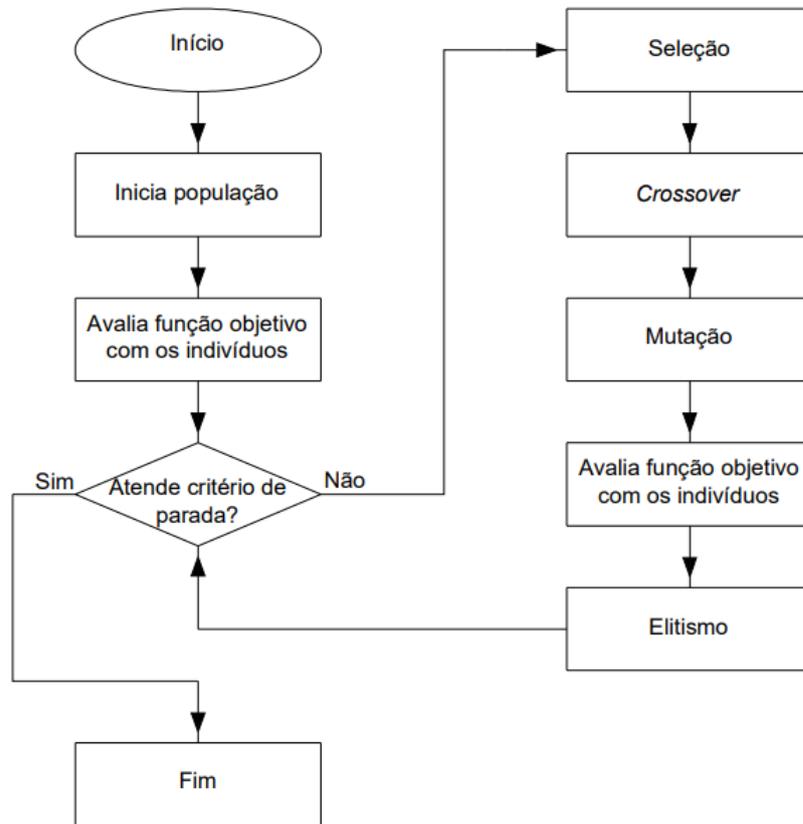
Por fim é realizado a avaliação da função objetivo com os novos indivíduos, tendo o intuito de conhecer a melhora em relação a população anterior. A concepção da nova população pode ser realizada através dos novos indivíduos criados pelas etapas citadas, adicionando os melhores da população anterior, este processo é conhecido como elitismo. Com a iteração destas etapas pode-se finalizar o algoritmo impondo alguma tolerância para a avaliação da função objetivo. A Fig. 12 mostra o fluxograma de um algoritmo genético apresentando cada etapa mencionada anteriormente.

Figura 11 – Exemplo de antes e após a etapa de mutação.



Fonte: Autor.

Figura 12 – Fluxograma do algoritmo genético (GA).



Fonte: Autor.

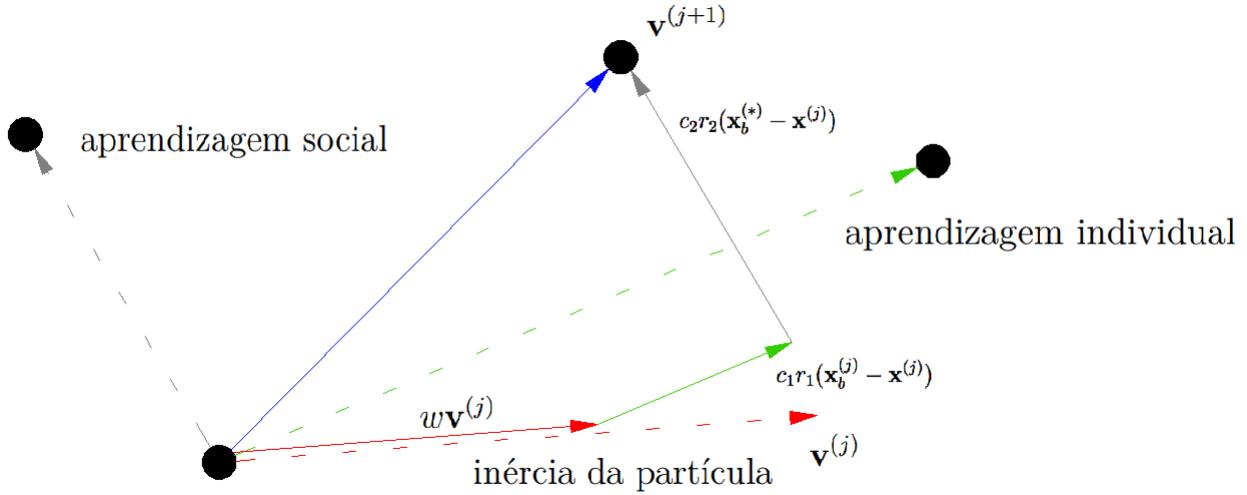
3.2 OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS - *PARTICLE SWARM OPTIMIZATION* (PSO)

A otimização por enxame de partículas (PSO), desenvolvido por Eberhart e Kennedy (1995), utiliza da física do momento linear para encontrar o ótimo global de um problema de otimização. Cada indivíduo da sua população, chamados de partículas, são independentes e possuem controle de sua velocidade e posição atualizadas e sempre carregam sua melhor solução encontrada individualmente. A melhor solução encontrada entre todas as partículas é compartilhada para toda sua população a cada iteração do algoritmo. Portanto, as partículas são aceleradas em direção a melhor solução ótima encontrada até o momento entre toda a população. Esta técnica evidencia a relação do comportamento social de um enxame, onde utiliza-se o melhor caminho encontrado até o momento para definir o ótimo global. A atualização de velocidade e posição de cada partícula realizada pelo algoritmo PSO é representada por

$$\begin{aligned} \mathbf{x}_{j+1} &= \mathbf{x}_j + \mathbf{v}_j \\ \mathbf{v}_{j+1} &= w\mathbf{v}_j + c_1r_1(\mathbf{x}_{b,j} - \mathbf{x}_j) + c_2r_2(\bar{\mathbf{x}}_b - \mathbf{x}_j), \end{aligned} \quad (17)$$

onde \mathbf{x}_{j+1} e \mathbf{v}_{j+1} são, respectivamente, a posição e a velocidade da partícula na iteração $j + 1$. c_1 representa a taxa de aprendizagem individual de cada partícula e c_2 a taxa de aprendizagem social. De acordo com Wappler (2014), esses valores dependem do problema

Figura 13 – Vetor resultante (linha contínua em azul) dos vetores da Eq. (17).



Fonte: Autor.

a ser resolvido. r_1 e r_2 são números randômicos entre 0 e a unidade. Enquanto a variável w representa a inércia da partícula e está relacionada ao seu poder de exploração, um valor alto da mesma representa uma exploração global no domínio, enquanto um valor baixo de w contempla uma exploração local. Uma estratégia para a definição de w é o seu decréscimo ao longo das iterações do algoritmo, realizando inicialmente uma busca global pelo domínio e enquanto tende ao critério de parada a busca se torna mais localizada. As variáveis $\mathbf{x}_{b,j}$ e $\bar{\mathbf{x}}_b$ são, respectivamente, a melhor solução encontrada pela partícula e a melhor solução encontrada pela população.

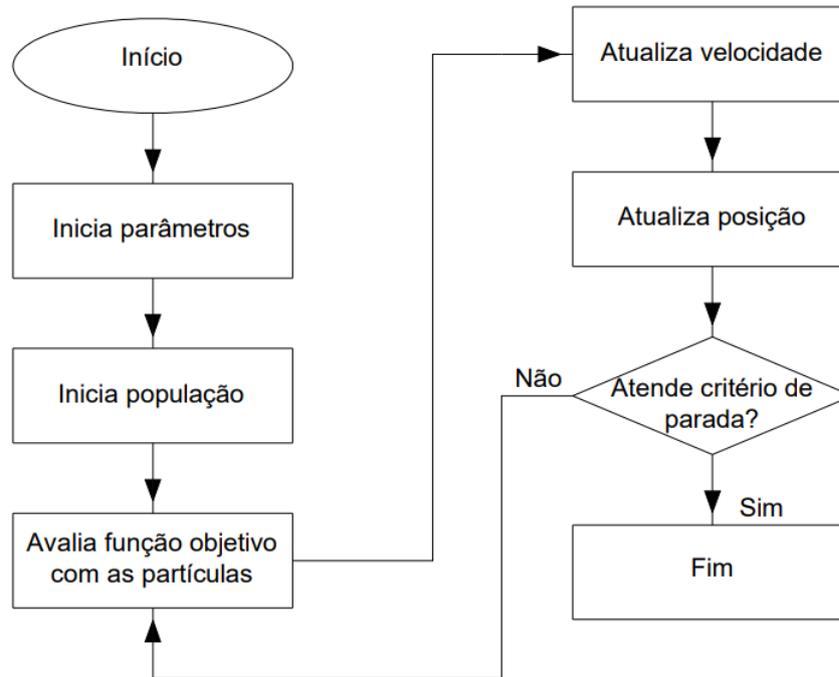
A Eq. (17) pode ser definida como a somatória de vetores gerando um vetor resultante conforme apresentado na Fig. 13, mostrando o desempenho de velocidade de cada partícula. A primeira parcela da Eq. (17) representa a influência da inércia w sobre a velocidade da partícula (vetores em vermelho na Fig. 13). Enquanto a segunda representa o incremento na velocidade devido sua taxa de aprendizagem individual (vetores em verde na Fig. 13) e a última representa o incremento devido sua taxa de sociabilidade com o restante da população (vetores em cinza na Fig. 13).

Durante toda a execução do algoritmo PSO a população é fixa, sendo iniciada aleatoriamente dentro do domínio S . Com a inicialização dos dados de velocidade e posição, ocorre uma avaliação da função objetivo a cada iteração, atualizando as partículas com novas informações de qual a melhor região do domínio a ser explorada. Diferente da técnica de algoritmo genético, no PSO existe uma interação social e colaborativa na população, quanto o algoritmo genético executa etapas para criar melhores indivíduos para a população, a Fig. 14 mostra o fluxograma do algoritmo PSO.

3.3 CONSIDERAÇÕES DE RESTRIÇÕES EM TÉCNICAS POPULACIONAIS

Conforme apresentado anteriormente, as restrições definem no domínio S uma região viável e outra inviável de soluções. Segundo Coello (2002), uma forma comum de tratar as restrições em técnicas de otimização populacionais é aplicando uma penalização no resultado da função objetivo. Por exemplo, a medida que se ranqueia as soluções da

Figura 14 – Fluxograma do algoritmo PSO.



Fonte: Autor.

função objetivo da menor para a maior em caso de minimização, as que se encontram na região inviável se tornam as últimas e conseqüentemente descartáveis para a melhor solução. Matematicamente, caso alguma restrição seja violada, definimos um novo valor para a função objetivo na forma

$$\bar{f}(\mathbf{x}) = f(\mathbf{x}) + \mathcal{C}(\mathbf{x}). \quad (18)$$

Onde $\bar{f}(\mathbf{x})$ é o novo valor da função objetivo f e $\mathcal{C}(\mathbf{x})$ é a constante de penalização. Esta técnica de penalização torna o problema com restrição em irrestrito, porém a utilização de penalização depende de estratégias ao ser aplicada. O valor de penalização muito alto pode trazer uma variável de projeto da região inviável para a viável rapidamente de forma a não explorar a região de fronteira. Por outro lado, um valor de penalização muito baixo haverá uma exploração demasiada da região inviável, gerando alto custo computacional e não explorando a região de fronteira.

É possível definir a constante \mathcal{C} como uma função de penalização, Coello (2002) e Mezura-Montes e Coello (2011) trazem várias técnicas para estes casos, como estática, dinâmica, adaptativa e entre outras. Nessas técnicas, a nova função de penalização $\mathcal{C}(\mathbf{x})$ é muitas vezes dependente de parâmetros definidos inicialmente pelo usuário e do número de restrições. Tornando um problema com muitas restrições demasiadamente custoso do ponto de vista computacional.

Deb (2000) desenvolveu uma técnica onde a função objetivo em uma região inviável depende não somente da quantidade de restrições violadas mas também da população de soluções da iteração j . Essa técnica estabelece regras de torneio para avaliação de pares de soluções da função objetivo, as quais são:

- Qualquer solução viável é preferida em relação a uma solução inviável;

- Entre duas soluções viáveis, a que possui a melhor solução da função objetivo é preferida;
- Entre duas soluções inviáveis, a que possui o menor valor da violação da restrição é preferida.

Realizando uma análise para essas três avaliações, a primeira compara somente valores da função objetivo, enquanto a segunda é uma análise se a solução é ou não viável, por fim, a terceira compara valores de violação de restrição. A ideia é não computar a função objetivo de um vetor de variável de projeto que se encontre na região inviável do domínio S . Portanto, ocorre um comparativo somente em termos de violação das restrições, trazendo menor custo computacional. Assim, essa técnica faz o algoritmo descartar as soluções do espaço inviável. Deb (2000) equaciona na forma

$$\begin{cases} \bar{f}(\mathbf{x}) = f(\mathbf{x}) & \text{se } g_i \leq 0 \\ \bar{f}(\mathbf{x}) = f_w(\mathbf{x}) + \sum_{j=1}^n \max(0, g_j) & \text{caso contrário,} \end{cases} \quad (19)$$

onde f_w é a pior solução da função objetivo na região viável e a segunda parcela representa o somatório das restrições violadas. Dessa forma, ocorre a convergência da população da região inviável para a região viável. Portanto, elimina-se a necessidade de computar a função objetivo da população na região inviável, enquanto as soluções viáveis mantêm seus valores originais.

4 UMA VISÃO GERAL DE PROBABILIDADE

Incertezas do cotidiano podem ter sua ocorrência quantificadas através do uso da teoria de probabilidades. A probabilidade $P(A)$ é a definição numérica do evento aleatório A ocorrer, quantificando o grau de possibilidade da sua realização. Para cada evento existe um espaço amostral S que reúne todos os seus possíveis resultados. A probabilidade conjunta é quando dois eventos A e B ocorrem de forma independente, ou seja, onde um evento não influencia no resultado do outro, representada por

$$P(A, B) = P(A)P(B). \quad (20)$$

A probabilidade condicional é a probabilidade de um evento A ocorrer, dado que um evento B é verdadeiro, sendo denotada por $P(A|B)$ na forma

$$P(A|B) = \frac{P(A, B)}{P(B)} = \frac{P(A)P(B)}{P(B)}, \quad (21)$$

onde $P(B) > 0$.

Os eventos A e B podem se referir a variáveis aleatórias unidimensionais ou vetores aleatórios multidimensionais. A realização desses eventos podem ser distribuídas conforme uma função de probabilidade, as quais são parametrizadas por medidas estatísticas. Por sua vez, esses parâmetros podem ser ajustados para o aumento da probabilidade de um evento ocorrer, essa técnica é conhecida como verossimilhança.

4.1 VARIÁVEIS ALEATÓRIAS E FUNÇÕES DE PROBABILIDADE

Uma variável aleatória Z é uma medida que depende de um experimento ou fatores aleatórios (do termo em latim *alea*, que significa sorte), assim para um evento no espaço amostral S existe uma função Z que mapeia um número real, ou seja, $Z : S \rightarrow \mathbb{R}$. Usualmente, uma variável aleatória é representada por uma letra maiúscula, no caso Z , e suas realizações são denotadas pela correspondente minúscula z . Quando a imagem de Z é finita ou infinita contável a variável aleatória é conhecida por ser discreta. Por outro lado, quando a imagem de Z é infinita incontável é chamada de variável aleatória contínua. Um exemplo de representação da probabilidade de uma variável aleatória discreta é $P(Z = z)$, determinando qual a probabilidade da variável aleatória Z assumir o valor de z . Já para variáveis aleatórias contínuas, um exemplo de representação de probabilidade de um evento é $P(Z \leq z)$.

Através de várias realizações de uma variável aleatória discreta pode-se observar o comportamento do seu valor médio. Com isso podemos computar seu valor esperado como

$$\mu = E[Z] = \sum_{i=1}^n z_i P(Z_i), \quad (22)$$

onde n é o número de realizações possíveis. O valor esperado também é conhecido como média e representa o primeiro momento central. Sendo a uma constante e \tilde{Z} outra variável aleatória independente assim como Z , o valor esperado apresenta as seguintes propriedades (BOROVKOV, 2013):

- $E[a] = a$;
- $E[Z] + E[\check{Z}] = E[Z + \check{Z}]$;
- $E[a + Z] = a + E[Z]$;
- $E[aZ] = aE[Z]$.

Outra medida estatística importante é a variância σ^2 ou Var , a qual pode ser representada por

$$\sigma_Z^2 = Var[Z] = \frac{1}{n} \sum_{i=1}^n (z_i - E[Z])^2. \quad (23)$$

A variância indica o quão a variável Z está dispersa em relação ao valor esperado $E[Z]$ e representa o segundo momento central. Como propriedades da variância pode-se citar (BOROVKOV, 2013):

- $Var[a] = 0$;
- $Var[Z] + Var[\check{Z}] = Var[Z + \check{Z}]$;
- $Var[a + Z] = Var[Z]$;
- $Var[aZ] = a^2 Var[Z]$.

De posse da variância σ_Z^2 é possível calcular o desvio padrão σ_Z como

$$\sigma_Z = \sqrt{\sigma_Z^2}, \quad (24)$$

o qual representa o desvio absoluto entre a variável aleatória Z e o valor esperado $E[Z]$.

Diferente das variáveis aleatórias discretas, não se pode associar uma variável aleatória contínua a somente uma distribuição de probabilidade, visto que variáveis aleatórias contínuas assumem infinitos valores reais dentro de um intervalo infinitesimal $[z, z + dz]$. A probabilidade deste tipo de variável assumir o valor de uma variável aleatória discreta é nula. Portanto para estes casos, define-se a probabilidade marginal de uma variável aleatória contínua Z em um intervalo $[a, b]$ na forma

$$P(a \leq Z \leq b) = \int_a^b f_Z(z) dz. \quad (25)$$

Também é possível determinar a probabilidade acumulada de uma variável aleatória contínua. A qual representa a probabilidade de Z ser menor ou igual a z , ou seja, $F_Z(z) = P(Z \leq z)$, logo

$$F_Z(z) = P(Z \leq z) = \int_{-\infty}^z f_Z(z) dz. \quad (26)$$

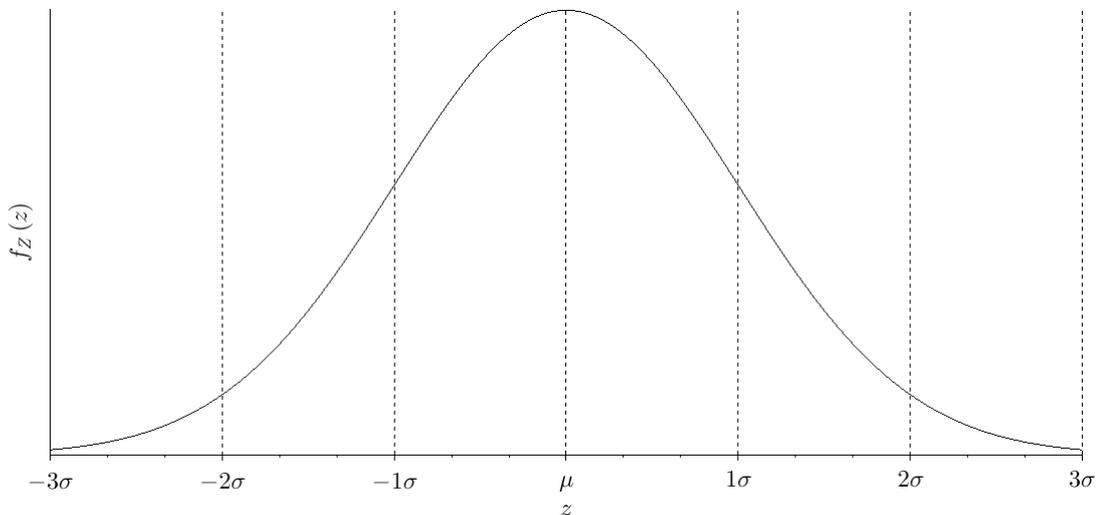
A função f_Z contida nas Eqs. (25) e (26) é conhecida como função densidade de probabilidade de Z . f_Z pode assumir valores maiores que 1 e tem como propriedade sua integral de

$-\infty$ a ∞ ser unitária. Existem diversas formas de apresentar f_Z , uma delas é conhecida como distribuição Gaussiana (BOROVKOV, 2013), representada por

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{(z - \mu)^2}{\sigma^2} \right) \right]. \quad (27)$$

A Fig. 15 apresenta a forma da Eq. (27) de representar a associação de uma distribuição Gaussiana a uma variável aleatória $Z \sim \mathcal{N}(\mu, \sigma^2)$. Os termos entre parênteses são conhecidos como os parâmetros que definem uma distribuição Gaussiana. Também podemos representar f_Z em sua forma parametrizada como $f_Z(z|\boldsymbol{\kappa})$, onde $\boldsymbol{\kappa}$ representa os parâmetros de f_Z , neste caso não se resumindo somente a μ e σ^2 .

Figura 15 – Modelo de uma distribuição Gaussiana univariada conforme Eq. (27).



Fonte: Autor.

Na Fig. 15 o valor esperado representa o centro da distribuição Gaussiana de probabilidade. Com posse da Eq. (25) podemos generalizar a expressão do valor esperado $E[Z]$ na forma

$$\mu = E[Z] = \int_{-\infty}^{\infty} z f_Z(z) dz. \quad (28)$$

Da mesma forma que o valor esperado $E[Z]$, a variância $Var[Z]$ pode ser generalizada como

$$\sigma_Z^2 = Var[Z] = E[(Z - E[Z])^2] = E[Z^2] - E[Z]^2. \quad (29)$$

É possível normalizar a variável aleatória Z para escalonar e transladar a distribuição $Z \sim \mathcal{N}(E[Z], Var[Z])$ para $\check{Z} \sim \mathcal{N}(aE[Z] + b, a^2Var[Z])$. Esta normalização se dá pela relação $\check{Z} = aZ + b$, tal que

$$E[\check{Z}] = aE[Z] + b \quad (30)$$

e

$$\text{Var}[\check{Z}] = a^2 \text{Var}[Z], \quad (31)$$

onde

$$\check{Z} = \frac{Z + \mu}{\sigma}, \quad (32)$$

sendo $\sigma > 0$. Percebe-se nas Eqs. (30) e (31) se $a = \sigma = 1$ e $b = \mu = 0$, podemos transladar e escalonar a distribuição Gaussiana $Z \sim \mathcal{N}(E[Z], \text{Var}[Z])$ para $\check{Z} \sim \mathcal{N}(0, 1)$. Dessa maneira, é possível associar a variável \check{Z} a uma função densidade de probabilidade Gaussiana, obtendo a função densidade de probabilidade normal, com as mesmas propriedades da Eq. (27), na forma

$$f_{\check{Z}}(\check{z}) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\check{z})^2\right]. \quad (33)$$

4.2 VETOR ALEATÓRIO E DISTRIBUIÇÃO MULTIVARIADA

Um vetor aleatório \mathbf{Z} com N -dimensões é um conjunto com N variáveis aleatórias na forma $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_N\}$. Anderson (1962) computa a probabilidade marginal de uma variável aleatória Z_i na forma

$$f_{Z_i}(z_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{Z}} dz_1 \cdots dz_{i-1} dz_{i+1} \cdots dz_N. \quad (34)$$

Ou a probabilidade conjunta entre duas variáveis Z_i e Z_j de \mathbf{Z} como

$$f_{Z_i Z_j}(z_i, z_j) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{Z}} dz_1 \cdots dz_{i-1} dz_{i+1} \cdots dz_{j-1} dz_{j+1} \cdots dz_N. \quad (35)$$

Dessa forma, pode-se calcular a probabilidade condicional para um dado \bar{z}_2 , assim $f_{Z_1|Z_2=\bar{z}_2}(z_1)$ é obtida com

$$f_{Z_1|Z_2=\bar{z}_2}(z_1) = \frac{f_{Z_1 Z_2}(z_1, \bar{z}_2)}{f_{Z_2}(\bar{z}_2)}. \quad (36)$$

Análogo a Eq. (26), podemos associar o vetor aleatório \mathbf{Z} a uma função de distribuição acumulativa multivariada $F_{\mathbf{Z}}$, tal que

$$F_{\mathbf{Z}}(\mathbf{Z}) = \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_N} f_{\mathbf{Z}}(\mathbf{z}) dz_1 \cdots dz_N. \quad (37)$$

$f_{\mathbf{Z}}$ nas Eqs. (34-37) representa a função de densidade de probabilidade multivariada. O vetor aleatório \mathbf{Z} de dimensão N possui uma distribuição multivariada se qualquer combinação linear seja representada por uma distribuição univariada. Ou seja, se a relação

$$\boldsymbol{\lambda}^T \mathbf{Z} \sim \mathcal{N}(\boldsymbol{\lambda}^T E[\mathbf{Z}], \boldsymbol{\lambda}^T \mathbf{C} \boldsymbol{\lambda}), \quad \forall \boldsymbol{\lambda}^T \in \mathbb{R}^N \quad (38)$$

for observada, pode-se afirmar que \mathbf{Z} possui uma distribuição multivariada na forma $\mathbf{Z} \sim \mathcal{N}(E[\mathbf{Z}], \mathbf{C})$. Onde $\boldsymbol{\lambda}^T$ é um vetor qualquer de dimensão N , $E[\mathbf{Z}]$ é o vetor com o valor esperado de cada variável de dimensão N e \mathbf{C} é a matriz de variâncias-covariâncias de dimensão $N \times N$. Portanto, bem como na sua forma univariada, pode-se definir uma distribuição multivariada através dos parâmetros $E[\mathbf{Z}]$ (ANDERSON, 1962), onde

$$E[\mathbf{Z}]_i = E[Z_i] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} z_i f_{\mathbf{Z}}(z_i) dz_i \cdots dz_N \quad (39)$$

e \mathbf{C} , tal que

$$C_{ij} = Cov[Z_i, Z_j], \quad (40)$$

onde

$$Cov[Z_i, Z_j] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (z_i - E[Z_i])(z_j - E[Z_j]) f_{Z_i} f_{Z_j}(z_i, z_j) dz_i dz_j. \quad (41)$$

Também análogo a forma univariada, podemos associar a distribuição multivariada $f_{\mathbf{Z}}$ a uma função de densidade de probabilidade multivariada Gaussiana como

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{C}|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (\mathbf{z} - E[\mathbf{Z}])^T \mathbf{C}^{-1} (\mathbf{z} - E[\mathbf{Z}]) \right], \quad (42)$$

onde $|\mathbf{C}|$ representa o determinante da matriz de variâncias-covariâncias. No Apêndice A é apresentado o desenvolvimento matemático da transformação da função de densidade de probabilidade univariada definida na Eq. (27) para multivariada conforme a Eq. (42).

Por fim, pode-se definir a distribuição condicional multivariada como $\mathbf{Z}_1 | \mathbf{Z}_2 \sim \mathcal{N}(E[\mathbf{Z}_1 | \mathbf{Z}_2], \mathbf{C}[\mathbf{Z}_1 | \mathbf{Z}_2])$. Partindo de um vetor aleatório composto por \mathbf{Z}_1 e \mathbf{Z}_2 na forma $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)^T$ o qual cada um tem seu valor esperado $E[\mathbf{Z}] = \{E[\mathbf{Z}_1], E[\mathbf{Z}_2]\}$ e definindo a matriz de variâncias-covariâncias particionada como

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}, \quad (43)$$

pode-se definir o valor esperado condicional na forma

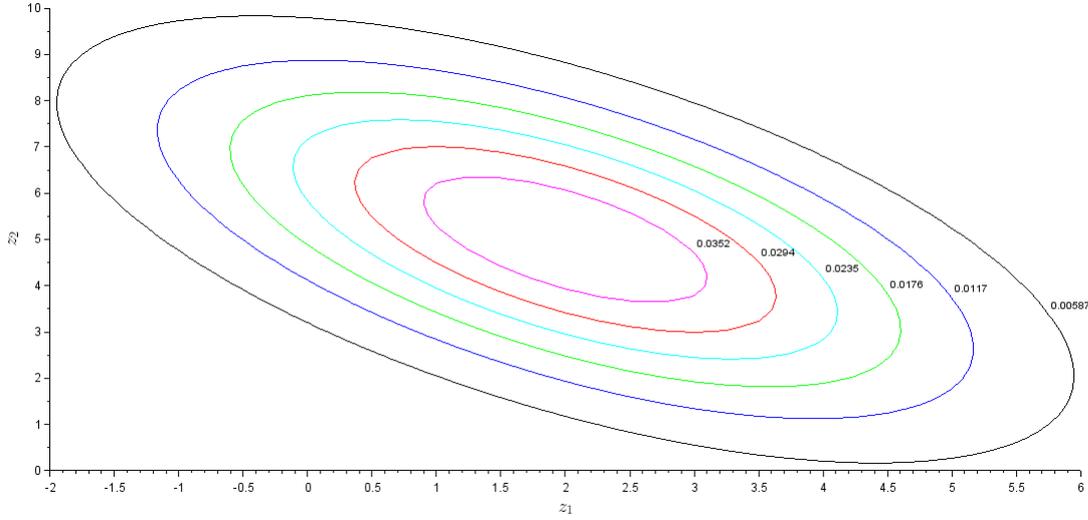
$$E[\mathbf{Z}_1 | \mathbf{Z}_2] = E[\mathbf{Z}_1] + \mathbf{C}_{12} \mathbf{C}_{22}^{-1} (\mathbf{Z}_2 - E[\mathbf{Z}_2]), \quad (44)$$

bem como a matriz de variâncias-covariâncias condicional, onde

$$\mathbf{C}[\mathbf{Z}_1 | \mathbf{Z}_2] = \mathbf{C}_{11} - \mathbf{C}_{12} \mathbf{C}_{22}^{-1} \mathbf{C}_{12}^T. \quad (45)$$

No Apêndice B é apresentado o desenvolvimento matemático para se obter as Eq. (44) e (45).

Figura 16 – Densidade de probabilidade para a Eq. (48).



Fonte: Autor.

Para exemplificar, consideramos uma distribuição Gaussiana bidimensional $N = 2$ com o vetor de valores esperados

$$E[\mathbf{Z}] = (2, 5)^T \quad (46)$$

e matriz de variâncias-covariâncias

$$\mathbf{C} = \begin{bmatrix} 4 & -3 \\ -3 & 6 \end{bmatrix} \quad (47)$$

onde

$$f_{\mathbf{Z}} = \frac{1}{(2\pi)|\mathbf{C}|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (z_1 - 2, z_2 - 5)^T \begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 0.266 \end{bmatrix} (z_1 - 2, z_2 - 5) \right]. \quad (48)$$

Graficando a Eq. (48) por meio de isolinhas conforme mostrado na Fig. 16, nota-se as maiores densidades de probabilidade na região próxima ao valor esperado $E[\mathbf{Z}] = (2, 5)^T$. Também é possível representar a probabilidade marginal para cada variável z_1 e z_2 , conforme Fig. 17.

Dado $\bar{z}_2 = 3$, a probabilidade condicional representada pela Eq. (36) é apresentada pela Fig. 18 sendo um corte transversal da distribuição da Fig. 16 em $z_2 = 3$ e escalonada por $f_{Z_2}(\bar{z}_2)$. Considerando $z_1 \leq 5$ a probabilidade acumulada é aproximadamente 0.897 representada pela área azul. Para essa condição, realizando a avaliação do valor esperado através da Eq. (39), onde $f_{\mathbf{Z}}(z_i) = f_{Z_1|Z_2=3}$, chegaremos em um valor esperado de $z_1 = 3$. De fato, pois

$$P(z_1 \leq 3 | \bar{z}_2 = 3) = \int_{-\infty}^3 f_{Z_1|z_2=3}(z_1) dz_1 = 0.5, \quad (49)$$

sendo este valor visualmente perceptível no centro do gráfico da Fig. 18.

Aplicando o conceito de densidade condicional Gaussiana multivariada no exemplo proposto, pode-se obter os mesmos resultados. Utilizando a Eq. (44) temos

$$E[Z_1|z_2 = 3] = 2 + (-3)\frac{1}{6}(3 - 5) = 3 \quad (50)$$

e a Eq. (45), onde

$$C[Z_1|z_2 = 3] = 4 - (-3)\frac{1}{6}(-3) = \frac{5}{2} \quad (51)$$

e aplicando na Eq. (42) na forma

$$f_{z_1 \leq 3 | \bar{z}_2 = 3}(z_1) = \frac{1}{\sqrt{2\pi\frac{5}{2}}} \exp \left[-\frac{1}{2} \left(\frac{z_1 - 3}{\sqrt{\frac{5}{2}}} \right)^2 \right], \quad (52)$$

onde percebe-se os mesmos resultados obtidos anteriormente.

4.3 VEROSSIMILHANÇA

A estimativa de uma função desconhecida tem como objetivo buscar a melhor distribuição que demonstre os valores realizados. Uma forma é assumir a representação das realizações por uma função densidade de probabilidade dependente dos parâmetros $\boldsymbol{\kappa}$, ou seja, $f_Z(z|\boldsymbol{\kappa})$. Esses parâmetros podem ser ajustados para aumentar a probabilidade de conotar as realizações da função desconhecida. Assim, utiliza-se da verossimilhança para definição dos parâmetros $\boldsymbol{\kappa}$ maximizando a probabilidade de representar as realizações \mathbf{z} .

Considerando uma amostra independente Z , com dimensão n e realizações $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$, Cressie (2015) define a função de verossimilhança L como

$$L(\boldsymbol{\kappa}|\mathbf{z}) = \prod_{i=1}^n f_Z(z_i|\boldsymbol{\kappa}). \quad (53)$$

Figura 17 – Probabilidade marginal para z_1 (à esquerda) e z_2 (à direita).

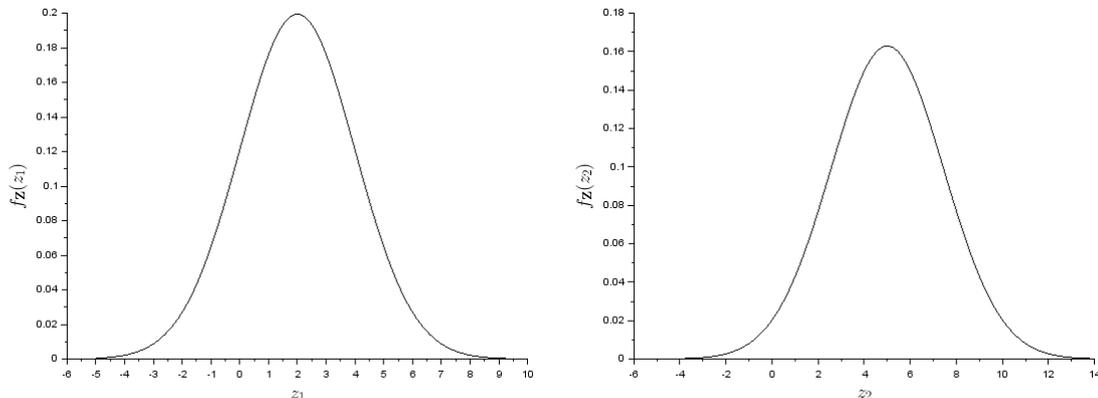
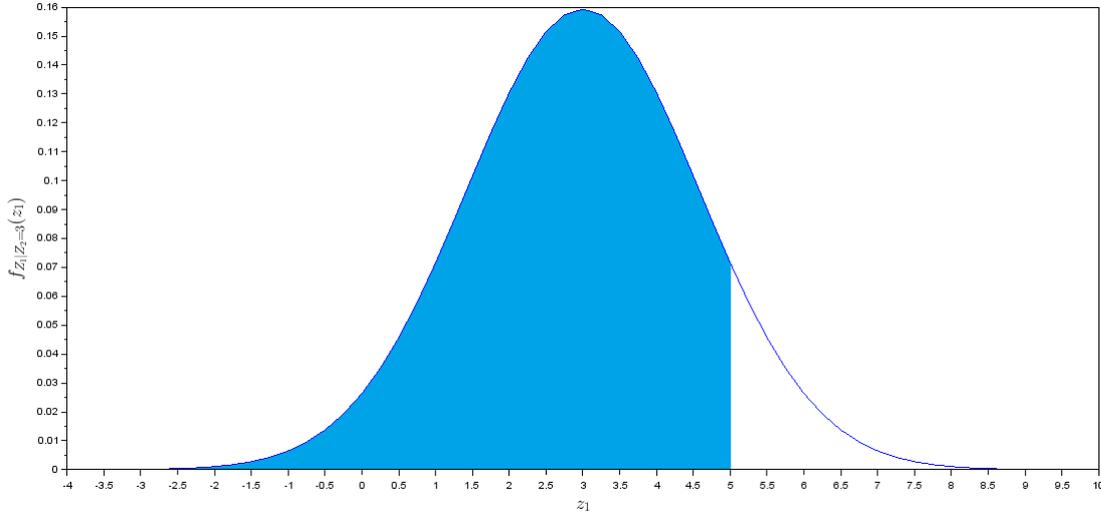


Figura 18 – Probabilidade condicional para $\bar{z}_2 = 3$ e $z_1 \leq 5$.

Fonte: Autor.

Nota-se que a verossimilhança é uma função de $\boldsymbol{\kappa}$, demonstrando a probabilidade de obter z_1, z_2, \dots, z_n pela função de densidade de probabilidade assumida que contém os parâmetros $\boldsymbol{\kappa}$. O produtório está relacionado a definição de probabilidade conjunta, com realizações independentes, conforme a Eq. (20).

Conforme o número de realizações de \mathbf{z} aumentam pode-se melhorar a estimativa da função desconhecida refinada pela verossimilhança. De forma que obtendo todas as realizações há total confiança nos parâmetros $\boldsymbol{\kappa}$ encontrados. Definindo $\hat{\boldsymbol{\kappa}}$ como os parâmetros para a estimativa da máxima verossimilhança (FISHER, 1912), então temos

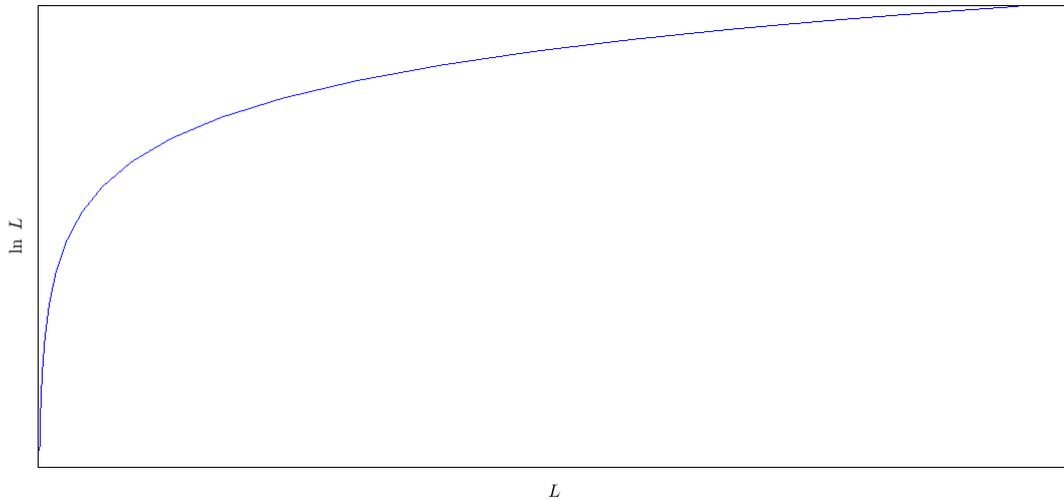
$$L(\hat{\boldsymbol{\kappa}}|\mathbf{z}) \geq L(\boldsymbol{\kappa}|\mathbf{z}). \quad (54)$$

Para uma melhor tratabilidade matemática é comum aplicar o logaritmo natural (\ln) na função verossimilhança. Enders (2010) explica que em geral a verossimilhança possui um valor entre 0 e 1, portanto seu produtório se torna um valor pequeno. Ao aplicar \ln , trabalha-se com números razoáveis e conseqüentemente o produtório de L na Eq. (53) se torna um somatório. Ainda esta consideração mantém a função de verossimilhança monotônica crescente, conforme mostrado na Fig. 19. Em vista disso, o problema é a maximização global de $\ln L$, com condição necessária

$$\frac{\partial \ln L}{\partial \boldsymbol{\kappa}} = \mathbf{0}. \quad (55)$$

Deve-se considerar que o valor máximo da função da Eq. (53) não é o mesmo quando aplicado seu logaritmo natural, porém $\hat{\boldsymbol{\kappa}}$ traz o valor máximo tanto da Eq. (53) quanto do seu logaritmo natural. Este processo também é conhecido como *argmax*, onde se deseja encontrar os argumentos (parâmetros) para obter o máximo da função. A consideração de $\ln L$ torna o valor da função negativo, consolidando a necessidade de utilizar o negativo do logaritmo natural para obter um número positivo, ou seja, aplica-se $-\ln L$. Portanto, o problema torna-se a minimização global da função $-\ln L$ em relação aos parâmetros $\boldsymbol{\kappa}$ ou *argmin*.

Figura 19 – Relação de monotocidade crescente da função L com $\ln L$.



Fonte: Autor.

Considerando $\kappa = \{\mu, \sigma^2\}$ e $Z \sim \mathcal{N}(\mu, \sigma^2)$, a Eq. (27) pode ser utilizada como função verossimilhança L , tal que

$$L(\mu, \sigma^2 | z_1, z_2, \dots, z_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{(z_i - \mu)^2}{\sigma^2} \right) \right], \quad (56)$$

onde seu logaritmo natural é

$$\ln L(\mu, \sigma^2 | z_1, z_2, \dots, z_n) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2. \quad (57)$$

Com o intuito de definir $\hat{\kappa}$, inicialmente pode-se derivar a Eq. (57) em relação a $\kappa_1 = \mu$, conforme condição da Eq. (55), resultando em

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu) = 0. \quad (58)$$

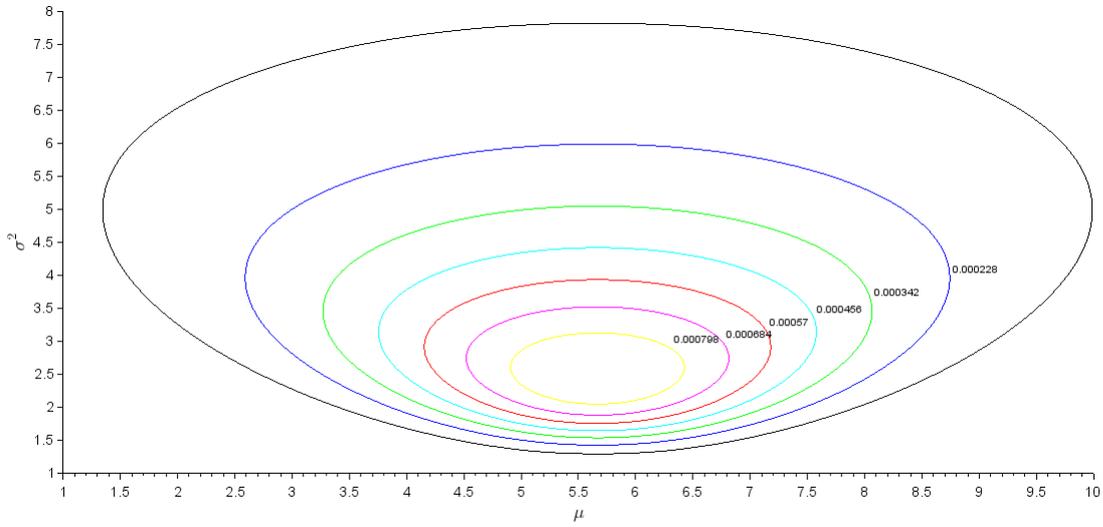
Essa equação é solucionada para

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n z_i. \quad (59)$$

Realizando o mesmo procedimento para $\kappa_2 = \sigma^2$ obtemos

$$\frac{\partial \ln L}{\partial \sigma^2} = -n + \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 = 0, \quad (60)$$

Figura 20 – $\hat{\mu}$ e $\hat{\sigma}$ maximizam a probabilidade de representar $\mathbf{z} = \{3, 5, 9\}$, assumindo uma distribuição Gaussiana.



Fonte: Autor.

resolvendo para

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \mu)^2, \quad (61)$$

definindo $\hat{\boldsymbol{\kappa}} = \{\hat{\mu}, \hat{\sigma}^2\}$. No Apêndice C é demonstrado o desenvolvimento matemático para encontrar $\hat{\mu}$ e $\hat{\sigma}^2$ a partir da Eq. (57).

Para exemplificar a utilização da máxima verossimilhança, considere o vetor $\mathbf{z} = \{3, 5, 9\}$, onde $\hat{\mu} = 5.66$ e $\hat{\sigma}^2 = 2.49$, os quais são os parâmetros que maximizam a probabilidade de representar \mathbf{z} , conforme mostrado na Fig. 20.

5 PROCESSO ESTOCÁSTICO, PROCESSO GAUSSIANO E PREDITOR

Eventos não incertos possuem características determinísticas, reproduzindo sempre o mesmo resultado. Por outro lado, um processo estocástico ocorre quando há incerteza na realização de um evento e o mesmo se desenvolve de forma aleatória e imprevisível ao longo do tempo (CRESSIE, 2015). Quando a entrada do processo estocástico é associada a uma variável aleatória com distribuição Gaussiana, pode-se particularizar o processo estocástico em Processo Gaussiano (GP). É possível aplicar GPs para definir preditores, permitindo inferir funções desconhecidas. Um exemplo é o preditor utilizando a média condicional.

5.1 PROCESSOS ESTOCÁSTICOS

Um processo estocástico $Z = \{Z(t), t \in T\}$ é uma coleção de variáveis aleatórias, sendo que para cada $t \in T$, $Z(t)$ é uma variável aleatória contida em um mesmo espaço de probabilidade S . Onde t é o tempo ou espaço e T é o conjunto de indexadores do processo estocástico. De forma prática, pode-se afirmar que é um fenômeno que varia incertamente à medida que observações vão ocorrendo ao longo do tempo. É comum definir processos estocásticos em função do tempo, porém podem ser definidos para qualquer indexador. Por exemplo, $Z_0, Z_1, Z_2 \dots$ é uma coleção de variáveis aleatórias observadas discretamente. Dessa forma, assim como variáveis aleatórias, processos estocásticos são classificados como discretos ou contínuos. O primeiro é representado quando T é um conjunto finito ou contável, enquanto o segundo T é um intervalo de números reais. Portanto, é possível associar um vetor aleatório \mathbf{Z} a um processo estocástico, mantendo as análises realizadas anteriormente para condição multivariada. Inclusive é possível aplicar medidas estatísticas como o valor esperado e a variância em um processo estocástico.

Considerando o valor esperado em um dado t_i na forma $E[\mathbf{Z}(t_i)]$, é possível computar a autocovariância em um momento t_j como

$$\begin{aligned} Cov[\mathbf{Z}(t_i), \mathbf{Z}(t_j)] &= E[(\mathbf{Z}(t_i) - E[\mathbf{Z}(t_i)])(\mathbf{Z}(t_j) - E[\mathbf{Z}(t_j)])^T] \\ Cov[\mathbf{Z}(t_i), \mathbf{Z}(t_j)] &= E[\mathbf{Z}(t_i)\mathbf{Z}(t_j)^T] - E[\mathbf{Z}(t_i)]E[\mathbf{Z}(t_j)]^T. \end{aligned} \quad (62)$$

Também é possível computar sua autocorrelação na forma

$$Cor[\mathbf{Z}(t_i), \mathbf{Z}(t_j)] = \frac{Cov[\mathbf{Z}(t_i), \mathbf{Z}(t_i)]}{\sqrt{\sigma_{\mathbf{Z}(t_i)}^2 \sigma_{\mathbf{Z}(t_j)}^2}}. \quad (63)$$

Quando um processo estocástico possui sua distribuição de probabilidade inalterada com o avanço do indexador t , o mesmo é chamado de processo estacionário. Ou seja, as características do processo não se alteram com o indexador t , portanto a translação do processo em t é irrelevante. Dado um intervalo do indexador \bar{t} , pode-se afirmar que um processo estocástico é estritamente estacionário quando

$$f(\mathbf{Z}(t_i + \bar{t})) = f(\mathbf{Z}(t_i)). \quad (64)$$

Portanto, um processo estacionário se desenvolve em torno do valor esperado E ao longo do indexador t . Onde tanto $E[\mathbf{Z}(t_i)]$ quanto e sua autocovariância $Cov[\mathbf{Z}(t_i), \mathbf{Z}(t_j)]$ são constantes em t .

Um processo estocástico é dito fracamente estacionário se seu valor esperado é constante, sua autocovariância é finita e dependente exclusivamente da distância. Essas propriedades podem ser definidas como:

- $E[\mathbf{Z}(t_i)] = cte$;
- $Cov[Z(t)^2] < \infty$;
- $Cov[Z(t), Z(t + \bar{t})]$ depende de $\bar{t} \forall t$.

Percebe-se que essas propriedades estão relacionadas para os momentos de primeira e de segunda ordem. Um exemplo de processos fracamente estacionários são os processos Gaussianos, os quais são definidos pelo valor esperado e a variância.

Em contrapartida do processo estocástico, ocorre o processo determinístico, onde é possível prever seu estado futuro sem aleatoriedade. Um processo determinístico sempre vai produzir o mesmo resultado para as mesmas entradas, um exemplo são as leis da física definidas através de equações diferenciais.

5.2 PROCESSO GAUSSIANO

Processo Gaussiano (GP) é um caso particular de processo estocástico, de forma que uma família finita de variáveis aleatórias $Z(\mathbf{x})$ atribuída a cada posição $\mathbf{x} \in \mathbb{R}^n$ do domínio S possuem uma distribuição Gaussiana multivariada. Assim, uma variável aleatória $Z(\mathbf{x})$ possui uma distribuição Gaussiana multivariada em um conjunto discreto $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, conhecidos como pontos de suportes, representada por

$$\mathbf{z}|\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K}), \quad (65)$$

onde $\mathbf{z} = \{Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_N)\}$, $\boldsymbol{\mu} = \{E[Z(\mathbf{x}_1)], \dots, E[Z(\mathbf{x}_N)]\}$ e \mathbf{K} é a matriz de variâncias-covariâncias nos pontos de suportes \mathbf{X} .

Em um processo Gaussiano a matriz de variâncias-covariâncias é assumida como \mathbf{K} , na qual estão contidas as funções de similaridade, conhecidas como *kernels* (WILLIAMS; RASMUSSEN, 2006). Essas funções especificam a relação estatística entre as amostras \mathbf{X} e dependem exclusivamente da distância entre as próprias amostras na forma $\|\mathbf{x}_i - \mathbf{x}_j\|$, ou seja, dependem da posição relativa. A matriz \mathbf{K} é simétrica e positiva definida, enquanto o *kernel* pode ter várias formas para mapear $\mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}$. Williams e Rasmussen (2006) trazem alguns exemplos de *kernels*, um deles é o exponencial quadrático (ou RBF - *radial basis function* em inglês) na forma

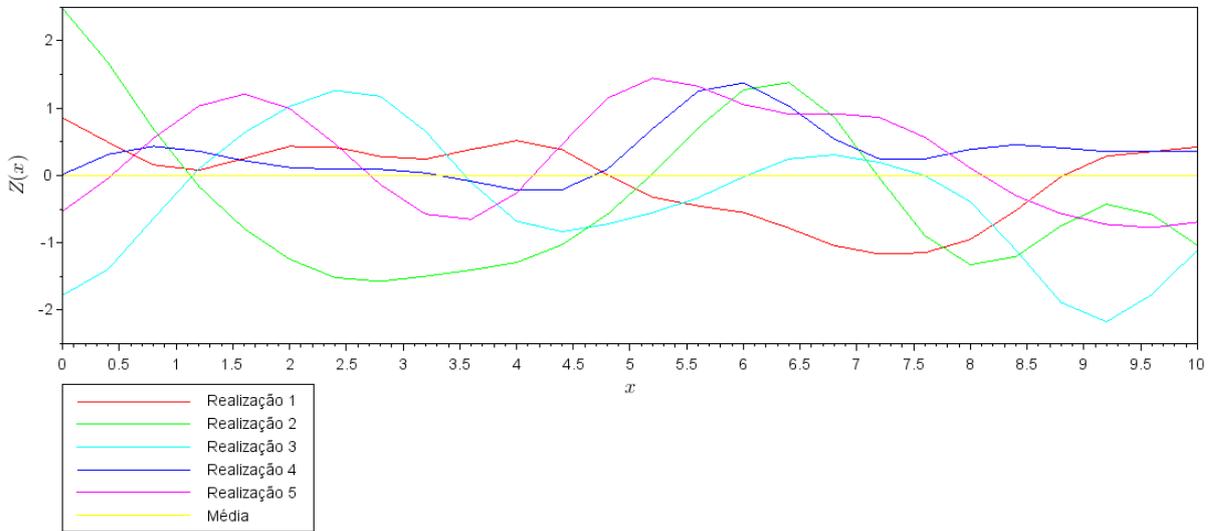
$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_k \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2l^2}\right) \quad (66)$$

onde l determina a característica de abrangência e escala do *kernel*.

Um processo Gaussiano pode ter seu comportamento definido exclusivamente pelo momento de segunda ordem, ou seja, somente pela matriz \mathbf{K} . Dessa forma, é assumido que a média $\boldsymbol{\mu}$ é nula, onde o GP torna-se

$$\mathbf{z}|\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu} = \mathbf{0}, \mathbf{K}). \quad (67)$$

Figura 21 – Processo Gaussiano à *priori* unidimensional, onde $x \in (0, 10)$, para 5 realizações de um evento.



Fonte: Autor.

A Eq. (67) apresenta uma quantidade incerta para representar os dados, sendo definida somente pela matriz \mathbf{K} , este tipo de Processo Gaussiano é conhecido como à *priori* (MACKAY, 2003). A Fig. 21 mostra um exemplo de 5 realizações de um evento utilizando a média nula e o *kernel* exponencial quadrático para formar a matriz \mathbf{K} em um processo Gaussiano à *priori*.

À medida em que \mathbf{z} realizações sem ruído de um evento em \mathbf{X} pontos de suporte são obtidas é possível construir um processo Gaussiano à *posteriori*. Nesse caso pode-se prever novos valores de \mathbf{z}_* em N_* novos pontos \mathbf{X}_* na forma de probabilidade condicional. Considerando uma distribuição Gaussiana multivariada aumentada na forma

$$\begin{pmatrix} \mathbf{z} \\ \mathbf{z}_* \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu}_{**} \end{pmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^T & \mathbf{K}_{**} \end{bmatrix} \right) \quad (68)$$

tal que

$$\mathbf{z}_* | \mathbf{X}_*; \mathbf{X}, \mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*), \quad (69)$$

sendo

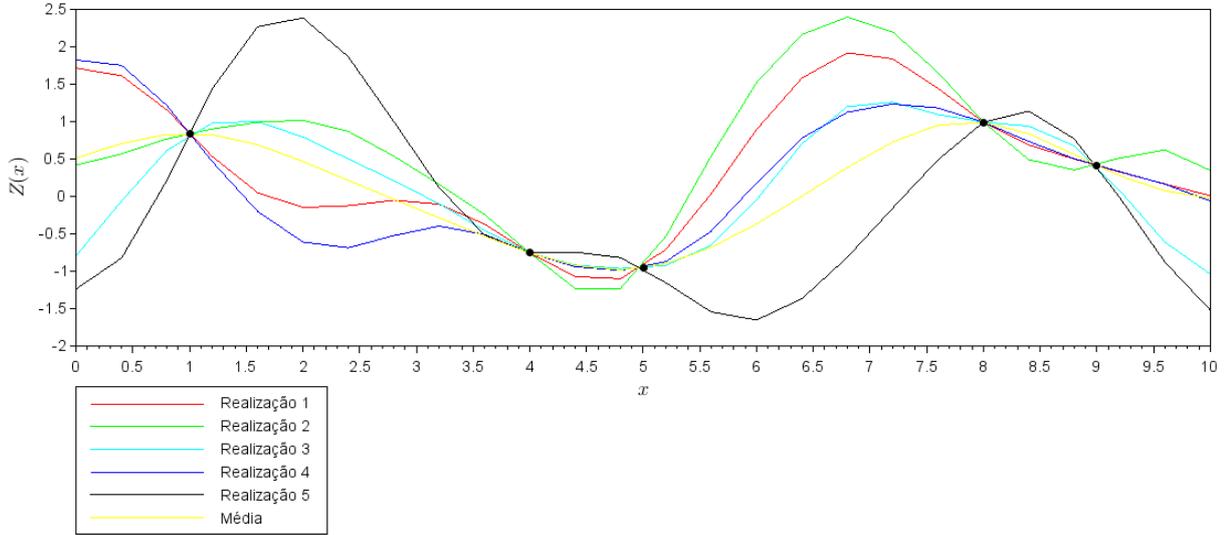
$$\boldsymbol{\mu}_* = \boldsymbol{\mu}_{**} + \mathbf{K}_*^T \mathbf{K}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \quad (70)$$

e

$$\boldsymbol{\Sigma}_* = \mathbf{K}_{**} - \mathbf{K}_*^T \mathbf{K}^{-1} \mathbf{K}_*, \quad (71)$$

onde \mathbf{K}_* possui a dimensão $N \times N_*$, enquanto \mathbf{K}_{**} tem dimensão $N_* \times N_*$. Observa-se a mesma natureza entre as Eqs. (70) e (71) com as Eqs. (44) e (45) respectivamente, pois são

Figura 22 – Processo Gaussiano à *posteriori* para a função $Z(x) = \sin(x)$, onde $x \in (0, 10)$, para 5 realizações de um evento e 5 pontos de suporte.



Fonte: Autor.

tratadas como distribuição Gaussiana multivariada condicionada a \mathbf{X}_*, \mathbf{X} e \mathbf{z} . Na Fig. 22 é apresentado um exemplo de predição à *posteriori* com 5 realizações e 5 pontos de suporte.

Quando as observações \mathbf{z} possuem ruído, o mesmo pode ser adicionado ao modelo. Esse ruído é considerado como uma distribuição Gaussiana com valor esperado nulo e variância $\sigma_z^2 \mathbf{I}$, ou seja, $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma_z^2 \mathbf{I})$. Logo, é adicionado de forma independente em cada observação e na matriz de variâncias-covariâncias \mathbf{K} na forma

$$\mathbf{z}_\epsilon = \mathbf{z} + \epsilon \quad (72)$$

e

$$\mathbf{K}_\epsilon = \mathbf{K} + \sigma_z^2 \mathbf{I}. \quad (73)$$

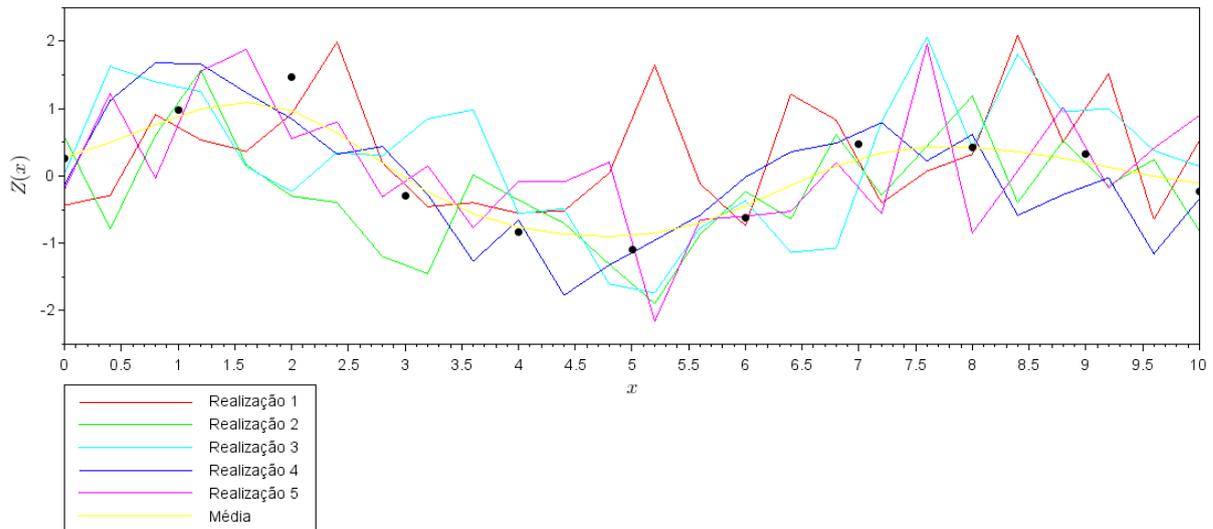
Finalmente, pode-se definir a distribuição Gaussiana com ruídos na forma

$$\mathbf{z}_* | \mathbf{X}_*, \mathbf{X}, \mathbf{z}_\epsilon = \mathcal{N} \sim (\mathbf{z}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_* + \sigma_z^2 \mathbf{I}), \quad (74)$$

onde é definida pelas Eqs. (70) e (71), porém substituindo a matriz \mathbf{K} pela Eq. (73). Percebe-se que a adição de um ruído não altera o modelo e o mesmo não é agregado a predição, mas modifica as observações e a covariância. A Fig. 23 é apresentado um exemplo de predição à *posteriori* com 5 realizações e 11 pontos de suporte.

Uma avaliação a ser realizada é a sensibilidade dos parâmetros σ_k e l no ajuste do modelo, assim como a variação do ruído σ_z^2 . Tomando como base o *kernel* exponencial quadrático, na Fig. 24 percebe-se para l menores a aproximação tende as observações, mostrando um modelo mais oscilante. Por outro lado para l maiores a aproximação se

Figura 23 – Processo Gaussiano *posteriori* para a função $Z(x) = \sin(x)$ com ruído, onde $x \in (0, 10)$, para 5 realizações de um evento e 11 pontos de suporte.



Fonte: Autor.

torna mais suave se distanciando das observações. Por sua vez, σ_k controla a translação vertical da aproximação, por ser uma constante que multiplica o restante da função *kernel*. Por fim, σ_z^2 menores trazem a aproximação mais próximas das observações, já σ_z^2 elevados evitam a sobreposição entre as observações e a aproximação.

5.3 VALOR ESPERADO CONDICIONAL COMO PREDITOR

Nas Figs. 22 e 23, é perceptível que o valor esperado (linha amarela) da Eq. (70) realiza uma aproximação adequada para o modelo. Para verificar se o valor esperado condicional é o melhor preditor deve-se determinar algumas propriedades dos momentos condicionais (ANDERSON, 1962). Considerando Z_1 e Z_2 variáveis aleatórias define-se a lei da expectativa total como

$$E[E[Z_1|Z_2]] = E[Z_1] \quad (75)$$

e a variância condicional na forma

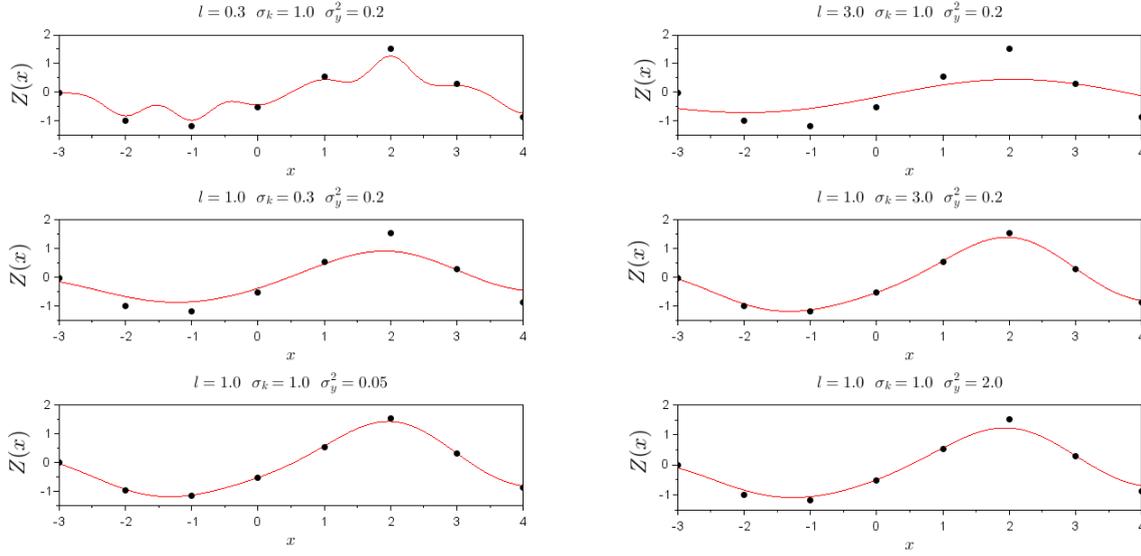
$$Var[Z_1|Z_2] = E[(Z_1 - E[Z_1|Z_2])^2|Z_2] = E[Z_1^2|Z_2] - (E[Z_1])^2. \quad (76)$$

Pode-se realizar combinações entre a variância e o valor esperado, podendo iniciar com a variância do valor esperado como

$$Var[E[Z_1|Z_2]] = E[(E[Z_1|Z_2])^2] - (E[E[Z_1|Z_2]])^2 = E[(E[Z_1|Z_2])^2] - (E[Z_1])^2, \quad (77)$$

também havendo o valor esperado da variância, definido por

Figura 24 – Influência dos parâmetros l , σ_k e σ_z^2 na aproximação da função $z(x) = \sin(x)$.



Fonte: Autor.

$$E[Var[Z_1|Z_2]] = E[(Z_1 - E[Z_1|Z_2])^2] = E[E[Z_1^2|Z_2]] - E[(E[Z_1|Z_2])^2]$$

$$E[Var[Z_1|Z_2]] = E[Z_1^2] - E[(E[Z_1|Z_2])^2]. \quad (78)$$

Observando que o termo $E[(E[Z_1|Z_2])^2]$ é comum nas Eqs. (77) e (78), é possível construir a relação da variância de Z_1 como

$$Var[E[Z_1|Z_2]] + E[Var[Z_1|Z_2]] = E[Z_1^2] - (E[Z_1])^2 = Var[Z_1]. \quad (79)$$

O critério para definir o melhor preditor é encontrar a função $\bar{h}(Z_2)$ a qual minimiza

$$E[(Z_1 - \bar{h}(Z_2))^2], \quad (80)$$

podendo ser reescrita como

$$E[(Z_1 - \bar{h}(Z_2))^2] = E[(Z_1 - E[Z_1|Z_2] + E[Z_1|Z_2] - \bar{h}(Z_2))^2]$$

$$E[(Z_1 - \bar{h}(Z_2))^2] = E[(Z_1 - E[Z_1|Z_2])^2 + 2(Z_1 - E[Z_1|Z_2])(E[Z_1|Z_2] - \bar{h}(Z_2)) - (\bar{h}(Z_2))^2]$$

$$E[(Z_1 - \bar{h}(Z_2))^2] = E[E[(Z_1 - E[Z_1|Z_2])^2|Z_2]] + E[2(E[Z_1|Z_2] - \bar{h}(Z_2))E[(Z_1 - E[Z_1|Z_2])|Z_2]] + E[(E[Z_1|Z_2] - \bar{h}(Z_2))^2]. \quad (81)$$

Onde o segundo termo é nulo, pois sua segunda parte torna-se $(E[Z_1|Z_2] - E[Z_1|Z_2])$, assim a Eq. (81) é resumida a

$$E[(Z_1 - \bar{h}(Z_2))^2] = E[Var[Z_1|Z_2]] + E[(E[Z_1|Z_2] - \bar{h}(Z_2))^2]. \quad (82)$$

Logo, se $\bar{h}(Z_2) = E[Z_1|Z_2]$ o segundo termo torna-se zero, minimizando a Eq. (80). O valor esperado condicional $E[Z_1|Z_2]$ é apresentado na Eq. (44) na forma multivariada, bem como na Eq. (70).

6 KRIGING, AMOSTRAGEM INICIAL E AQUISIÇÃO DE PONTOS DE SUPORTE

Podemos utilizar os conceitos dos Caps. 4 e 5 para definir modelos substitutos de funções desconhecidas, uma das formas é utilizando Kriging ordinário. Através de um *kernel* conhecido, o método de Kriging correlaciona um conjunto de dados previamente definido, criando o modelo substituto. Os dados definidos previamente podem ser gerados através de métodos de amostragem como estratificação completa ou hipercubo latino (HCL). Além disso, o modelo substituto pode receber novos dados que podem ser adquiridos de forma estratégica através de funções de aquisições, um exemplo é por meio do Expected Improvement (EI).

6.1 KRIGING

O método de Kriging foi desenvolvido pelo matemático e geólogo Georges Matheron (MATHERON, 1963), com base no estudo realizado pelo engenheiro de minas sul-africano Danie Krieger (KRIGE, 1951). O objetivo do estudo de Krieger era compreender a distribuição de ouro em uma mina através de poucas amostras, neste caso realizando o menor número de perfurações. Atualmente a *krigagem* (termo também designado por Matheron) possui várias aplicações em projetos de engenharia, principalmente relacionados ao projeto de otimização de sistemas mecânicos (SIMPSON *et al.*, 2001).

Considerando que não há ruídos nas realizações \mathbf{y} nos pontos de suporte \mathbf{X} , o preditor

$$\hat{y}(\mathbf{x}) = \mu + Z(\mathbf{X}), \quad (83)$$

é assumido para descrever uma função desconhecida, onde $Z(\mathbf{X})$ é um processo Gaussiano com média nula e variância σ^2 . A primeira parcela da Eq. (83) (μ) representa a tendência global, enquanto $Z(\mathbf{X})$ denota o ajuste local para a predição. Definindo o *kernel*

$$Cor(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\sum_{k=1}^N \theta_k |x_{i,k} - x_{j,k}|^{p_k}\right), \quad (84)$$

o qual correlaciona os valores de $\mathbf{y}(\mathbf{X})$ através das amostras \mathbf{X} , onde N é a dimensão do espaço amostral S , θ e p são parâmetros que definem o tamanho da correlação e a suavidade da predição respectivamente. A matriz de variâncias-covariâncias é definida na forma

$$\mathbf{C} = \sigma^2 \mathbf{\Psi}, \quad (85)$$

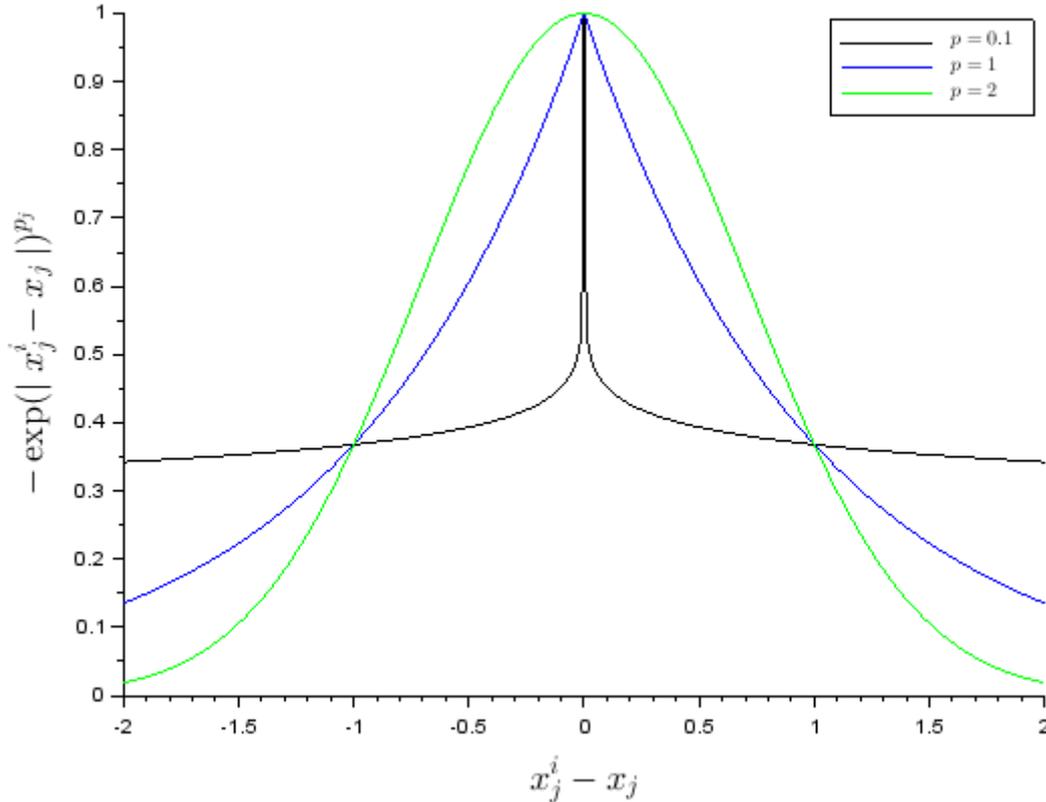
onde $\mathbf{\Psi}$ é a matriz de correlação com o *kernel* da Eq. (84), denotada por

$$\mathbf{\Psi} = \begin{bmatrix} Cor(\mathbf{x}_i, \mathbf{x}_j) & \cdots & Cor(\mathbf{x}_i, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ Cor(\mathbf{x}_N, \mathbf{x}_j) & \cdots & Cor(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}. \quad (86)$$

Inicialmente pode-se analisar o comportamento do *kernel* da Eq. (84) em relação aos parâmetros θ e p . Na Fig. 25 é possível observar que $\exp(-|x_{i,k} - x_{j,k}|^{p_k}) \rightarrow 1$ à

medida que a diferença entre duas amostras tende a zero ($x_{i,k} - x_{j,k} \rightarrow 0$), denotando que $y(x_{i,k})$ e $y(x_{j,k})$ possuem alta correlação. Por outro lado, quando $x_{i,k} - x_{j,k} \rightarrow \infty$ o valor de $\exp(-|x_{i,k} - x_{j,k}|^{p_k}) \rightarrow 0$, apontando que as amostras $y(x_{i,k})$ e $y(x_{j,k})$ não possuem correlação. Na Fig. 25 observa-se um gradiente de correlação contínuo para $p = 2$. Esse valor é o mais utilizado na literatura, trazendo suavidade para o *kernel* da Eq. 84.

Figura 25 – Comportamento do *kernel* da Eq. (84) em relação ao parâmetro p



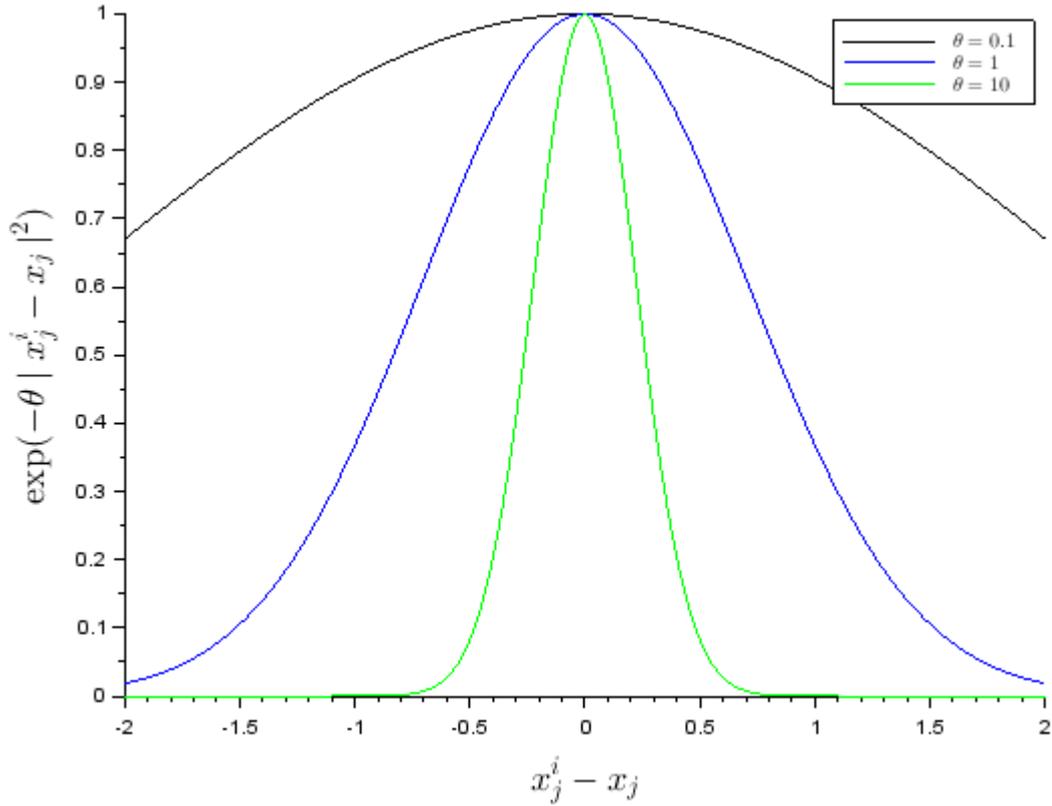
Fonte: Autor.

A Fig. 26 mostra como o parâmetro θ afeta o *kernel* da Eq. (84). Para valores baixos de θ haverá influência das amostras distantes na predição, enquanto para valores altos de θ observa-se que os pontos próximos da predição terão maior correlação.

Devido a $Z(\mathbf{X})$ ser um processo Gaussiano pode-se associar \hat{y} a uma distribuição Gaussiana multivariada na forma $\hat{y}(\mathbf{x}) \sim \mathcal{N}(\mathbf{1}\mu, \mathbf{C})$. Percebe-se que a matriz de variâncias-covariâncias \mathbf{C} é composta pela variância σ^2 e por Ψ , que por sua vez é dependente dos parâmetros θ e p . Os valores dos parâmetros que maximizam a probabilidade de representar \mathbf{y} são definidos através da utilização da técnica da verossimilhança. Assim, \hat{y} é associado a uma distribuição Gaussiana multivariada através de quatro parâmetros μ , σ , θ e p . Os dois primeiros parâmetros podem ser determinados conforme explicado no Capítulo 4. Por outro lado θ é não-convexo e definido através de Algoritmo Genético, e como definido previamente $p = 2$.

É possível definir a função verossimilhança conforme Eq. (42), porém reescrita em

Figura 26 – Comportamento do *kernel* da Eq. (84) em relação ao parâmetro θ



Fonte: Autor.

função dos parâmetros explícitos e implícitos citados anteriormente como

$$\mathcal{L}(\mu, \sigma^2, \theta, p|\mathbf{X}, \mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}} |\Psi|^{\frac{1}{2}}} \exp \left[-\frac{(\mathbf{y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2} \right]. \quad (87)$$

Aplicando o logaritmo natural temos

$$\ln(\mathcal{L}(\mu, \sigma^2, \theta, p|\mathbf{X}, \mathbf{y})) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{1}{2} \ln |\Psi| - \frac{(\mathbf{y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\mu)}{2\sigma^2}. \quad (88)$$

Portanto, utilizando o conceito da máxima verossimilhança o parâmetro μ é definido como

$$\begin{aligned} \frac{\partial \ln(\mathcal{L}(\mu, \sigma^2, \theta, p|\mathbf{X}, \mathbf{y}))}{\partial \mu} &= 0 \\ \hat{\mu} &= \frac{\mathbf{1}^T \Psi^{-1} \mathbf{y}}{\mathbf{1}^T \Psi^{-1} \mathbf{1}}. \end{aligned} \quad (89)$$

De forma independente segue-se o mesmo conceito para σ^2 , logo

$$\begin{aligned} \frac{\partial \ln(\mathcal{L}(\mu, \sigma^2, \theta, p|\mathbf{X}, \mathbf{y}))}{\partial \sigma^2} &= 0 \\ \hat{\sigma}^2 &= \frac{(\mathbf{y} - \mathbf{1}\hat{\mu})^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{N}. \end{aligned} \quad (90)$$

A derivação das Eqs. (89) e (90) a partir da Eq. (88) é demonstrada no Apêndice D. Finalmente, substituindo as Eqs. (89) e (90) na Eq. (88) e removendo os termos constantes chega-se a função ln-verossimilhança concentrada, tal que

$$\ln(\mathcal{L}(\theta, p|\mathbf{X}, \mathbf{y})) = -\frac{N}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln |\Psi|, \quad (91)$$

tornando uma função dependente de θ e p .

Com os parâmetros $\boldsymbol{\kappa}$ definidos, pode-se prever novos valores \hat{y} . Segundo Jones (2001), pode-se realizar isto adicionando x^+ amostras e sua respectiva predição \hat{y} na matriz de correlação Ψ e no vetor $\mathbf{y}(\mathbf{X})$ respectivamente. Pode-se iniciar essa análise determinando uma nova matriz de correlação aumentada

$$\tilde{\Psi} = \begin{bmatrix} \Psi & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{bmatrix} \quad (92)$$

e um novo vetor de realizações aumentado

$$\tilde{\mathbf{y}} = \{\mathbf{y}, \hat{y}\}^T, \quad (93)$$

tornando o problema dependente somente de \hat{y} . Onde o termo unitário na diagonal de $\tilde{\Psi}$ na Eq. (92) indica que \hat{y} tem uma correlação unitária entre si. Enquanto

$$\boldsymbol{\psi} = \exp \left(- \sum_{k=1}^N \theta_k |x_{i,k} - x_k^+|^{p_k} \right) \quad (94)$$

é o *kernel* que correlaciona as amostras realizadas \mathbf{x} com a amostra da predição x^+ .

Substituindo as Eqs. (92) e (93) na Eq. (88) obtemos

$$\ln(\mathcal{L}(\hat{y}|\mathbf{X}, \mathbf{y}, \hat{\mu}, \hat{\sigma}^2, \theta, p)) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln |\tilde{\Psi}| - \frac{(\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})^T \tilde{\Psi}^{-1} (\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})}{2\hat{\sigma}^2}, \quad (95)$$

tendo somente o último termo dependente de \hat{y} , tornando

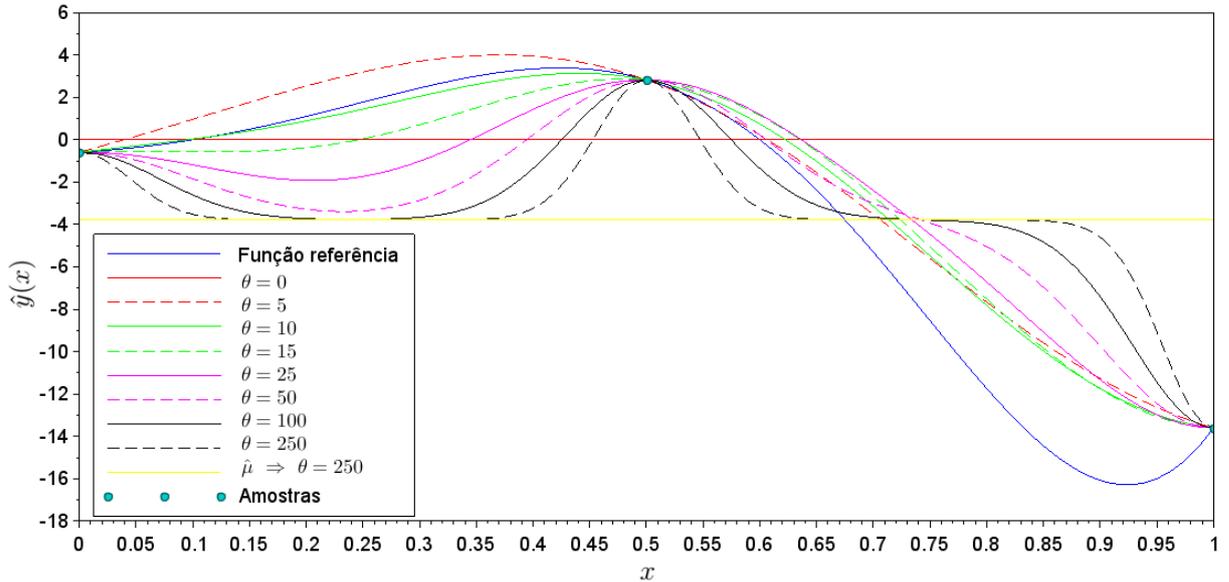
$$\ln(\mathcal{L}(\hat{y}|\mathbf{X}, \mathbf{y}, \hat{\mu}, \hat{\sigma}^2, \theta, p)) = -\frac{(\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})^T \tilde{\Psi}^{-1} (\tilde{\mathbf{y}} - \mathbf{1}\hat{\mu})}{2\hat{\sigma}^2}. \quad (96)$$

Resolvendo $\tilde{\Psi}^{-1}$ pelo método da inversa particionada conforme apresentado no Apêndice E, derivando a Eq. (96) em relação a \hat{y} e igualando a zero chega-se no preditor \hat{y} , logo

$$\frac{\partial \ln(\mathcal{L}(\hat{y}|\mathbf{X}, \mathbf{y}, \hat{\mu}, \hat{\sigma}^2, \theta, p))}{\partial \hat{y}} = 0$$

$$\hat{y}(\mathbf{x}) = \hat{\mu} + \boldsymbol{\psi}^T \Psi^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}). \quad (97)$$

Figura 27 – Predições da Eq. (99) em função de θ com 3 pontos de suporte.



Fonte: Autor.

Percebe-se que a Eq. (97) remete a Eq. (83), com a média estimada $\hat{\mu}$ e o vetor ψ com seus respectivos pesos $\Psi^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu})$. O Apêndice F apresenta o desenvolvimento matemático para a predição da Eq. (97) partindo da Eq. (91). Outra característica é a similaridade com as Eqs. (44) e (70), sendo o valor esperado condicional definido como melhor preditor no Capítulo 5. Também nota-se que é possível interpolar exatamente as realizações \mathbf{y} na forma

$$\hat{y}(\mathbf{x}) = \hat{\mu} + y_i - \hat{\mu}. \quad (98)$$

Por mais que o parâmetro θ é definido utilizando Algoritmo Genético, pode-se analisar o comportamento da predição da Eq. (97) em função deste parâmetro. A Fig. 27 apresenta predições para a função unidimensional

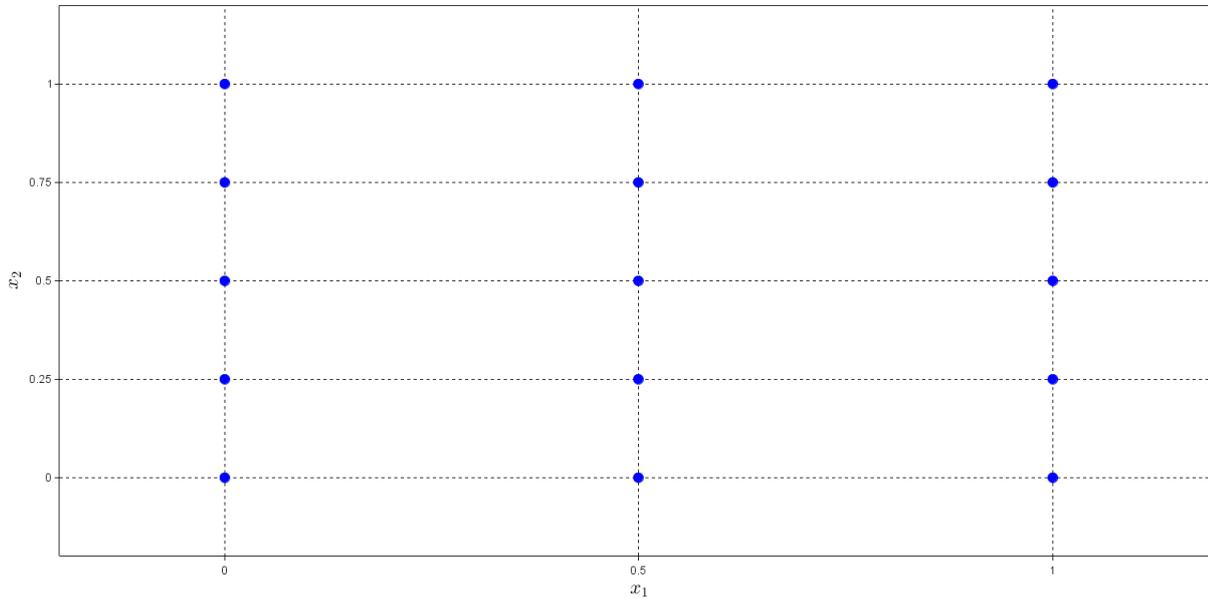
$$y(x) = \sin \left[2\pi \left(x - \frac{1}{10} \right) \right] \exp(x\pi) \quad (99)$$

utilizando a Eq. (97). É possível observar para θ nulo a predição tende a zero, uma vez que a média estimada $\hat{\mu}$ também é nula e a matriz Ψ se torna singular, pois é unitária. Para este caso, também é possível observar que valores maiores que zero até 15 apresentam uma boa predição para a função da Eq. (99), mesmo com um erro associado. Quando θ possui um valor alto, a predição tende a degenerar para a média estimada $\hat{\mu}$, esse fenômeno pode ser observado para $\theta \geq 100$.

6.2 AMOSTRAGEM INICIAL

O método de Kriging ordinário está condicionado aos valores das amostras \mathbf{X} e suas respectivas realizações \mathbf{y} . Inicialmente métodos de amostragem podem ser utilizados para

Figura 28 – Exemplo de amostragem com duas variáveis (x_1 e x_2) utilizando o método de fatoração completa.



Fonte: Autor.

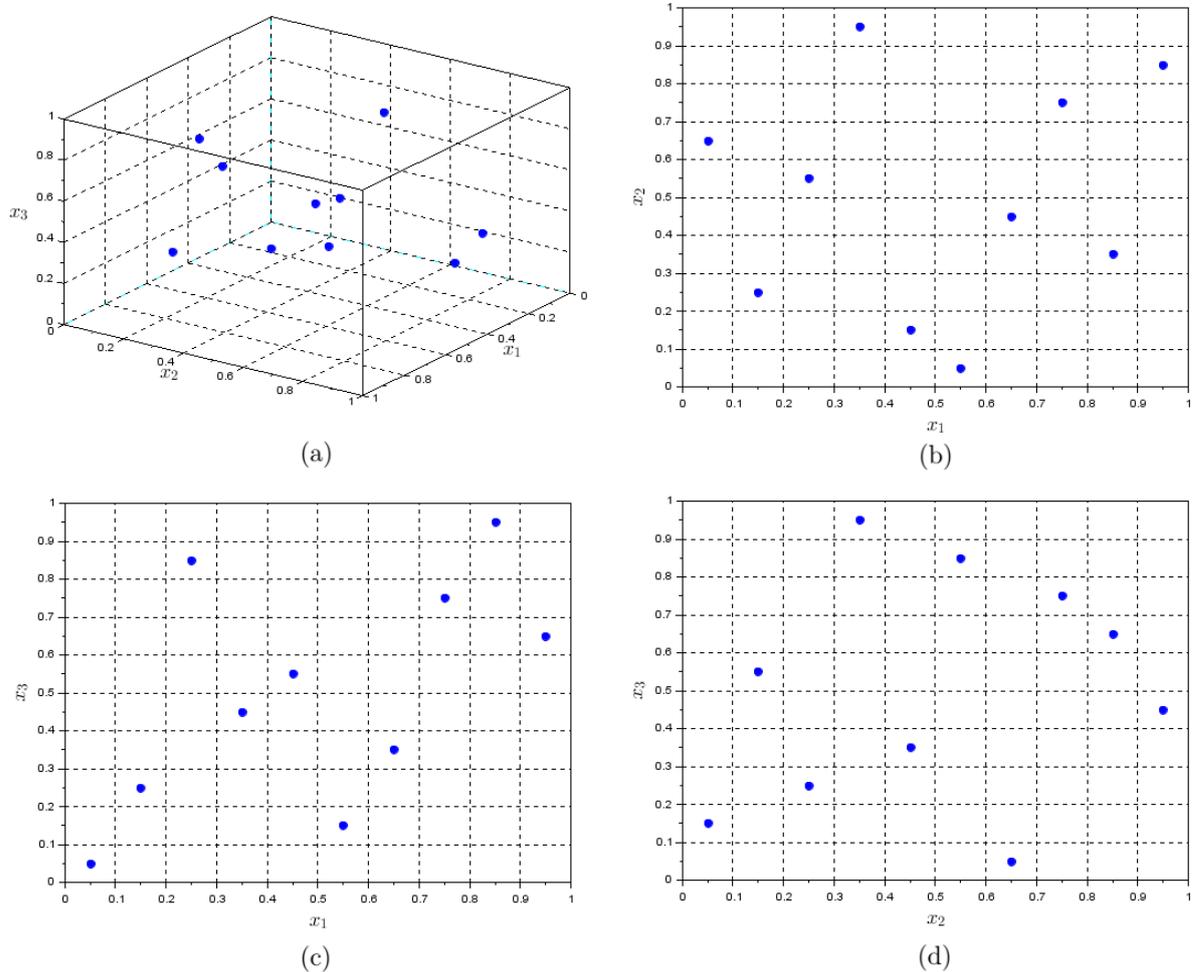
definir as amostras \mathbf{X} e obter as realizações \mathbf{y} , exemplos destes métodos são a fatoração completa (FORRESTER; SOBESTER; KEANE, 2008) e o hipercubo latino (MCKAY; BECKMAN; CONOVER, 1979).

O primeiro consiste em dividir o domínio S em um *grid* retangular de amostras. Dessa forma, a quantidade de n amostras é determinada pelo produto dos componentes do vetor \mathbf{b} , onde $n = b_1 \times b_2 \times \dots \times b_k$, sendo k a dimensão das amostras. A Fig. 28 mostra um exemplo do método de fatoração completa, onde as n amostras são divididas uniformemente. Um problema deste método é quando poucas divisões em uma variável qualquer não tornam a amostragem *space-filling*.

Por sua vez, o hipercubo latino é uma extensão melhorada da amostragem por fatoração completa. Neste método também divide-se o domínio para cada variável em um *grid* retangular e uniforme. Neste caso os espaços gerados pelo *grid* são conhecidos como estratos, onde uma única amostra é alocada. Esta alocação ocorre de maneira controlada, de forma que seja possível "sair" do domínio em qualquer direção paralela a qualquer eixo sem encontrar outro estrato ocupado. Mesmo assim o método do hipercubo latino puro não garante um plano amostral *space-filling*, pois a seleção do estrato onde a amostra é armazenada ainda ocorre controladamente mas de forma aleatória, conforme mostrado na Fig. 29.

Esse problema pode ser contornado utilizando a métrica *maxmin* definida por Johnson, Moore e Ylvisaker (1990) e posteriormente modificada por Morris e Mitchell (1995). Considere um vetor \mathbf{d} contendo a lista da maior para a menor distância igual para todos os pares possíveis da amostragem inicial \mathbf{X} . Também considera o vetor \mathbf{j} , o qual aloca a quantidade de número de pares de pontos em \mathbf{X} separados pelas distâncias contidas em \mathbf{d} , percebe-se que \mathbf{j} possui o mesmo tamanho de \mathbf{d} . As quais essas as distâncias contidas

Figura 29 – Exemplo de plano amostral com três variáveis (x_1 , x_2 e x_3) utilizando o método do hipercubo latino. Projeção tridimensional (a) e suas respectivas projeções bidimensionais (b), (c) e (d).



Fonte: Autor.

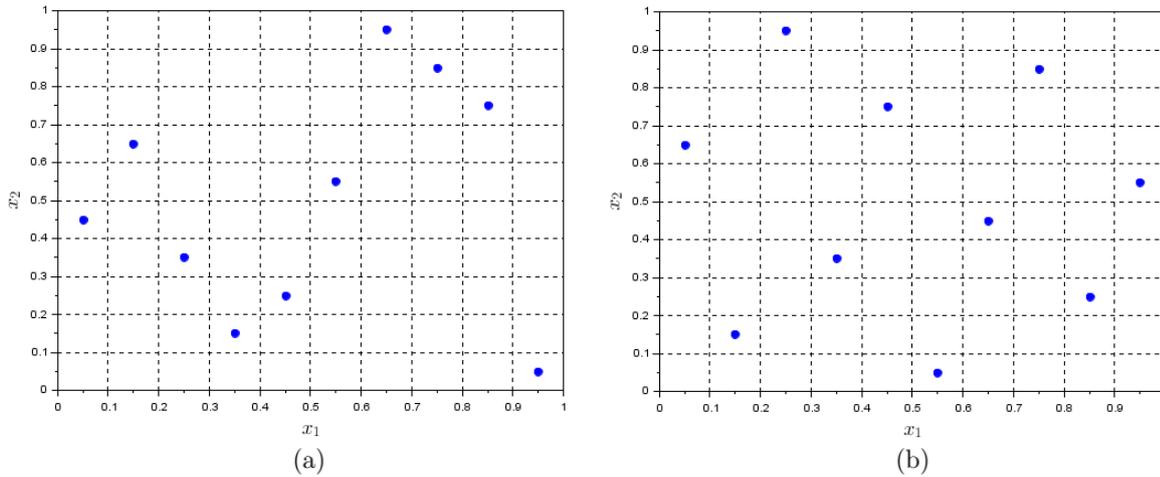
em \mathbf{d} são representadas por

$$d_m(\mathbf{x}_i, \mathbf{x}_j) = \left(\sum_{r=1}^k |\mathbf{x}_i - \mathbf{x}_j|^{\bar{p}} \right)^{\frac{1}{\bar{p}}}, \quad (100)$$

onde \mathbf{x}_i e \mathbf{x}_j são amostras distintas das amostras \mathbf{X} e $\bar{p} = 1$ representa a norma Manhattan (distância retangular), enquanto $\bar{p} = 2$ representa a norma Euclidiana. Forrester, Sobester e Keane (2008) citam que na literatura existem diferenças entre selecionar uma norma ou outra para avaliação da amostragem, porém a norma Manhattan é mais vantajosa do ponto de vista computacional.

A métrica *maxmin* afirma que as amostras \mathbf{X} são mais uniforme quando maximiza d_1 , entre os planos o qual é verdade, minimiza j_1 , entre os planos o qual é verdade, maximiza d_2 , entre os planos o qual é verdade, minimiza j_2 e assim por diante. Morris e Mitchell

Figura 30 – Comparação de amostragem com duas variáveis x_1 e x_2 , utilizando método do hipercubo latino puro (a) e otimizado (b).



Fonte: Autor.

(1995) ranqueiam o melhor \mathbf{X} de acordo com a função

$$\bar{\gamma}_q = \left(\sum_{j=1}^m \frac{\mathbf{j}_j}{\mathbf{d}_j^q} \right)^{\frac{1}{q}}, \quad (101)$$

definindo que o menor valor de $\bar{\gamma}_q$ possui o melhor \mathbf{X} com propriedades de um plano *space-filling*. Morris e Mitchell (1995) ainda recomendam minimizar a Eq. (101) para $q = 1, 2, 4, 10, 20, 50$ e 100 , já Forrester, Sobester e Keane (2008) mostram uma investigação sobre quais são os melhores valores de q para a Eq. (101).

Entretanto, para obter um \mathbf{X} otimizado é necessário aplicar um algoritmo de otimização na Eq. (101). Morris e Mitchell (1995) utilizam um algoritmo com técnica probabilística de otimização conhecido como *simulated annealing*. Enquanto Forrester, Sobester e Keane (2008) consideram para otimização uma alternativa de mutação utilizando a técnica de operação evolutiva (*Evolutionary Operation* (EVOP) em inglês).

A Fig. 30 mostra um comparativo entre duas amostragens \mathbf{X} em um domínio $S = [0, 1]^k$ com duas dimensões (x_1 e x_2) e $n = 10$ amostras. A primeira utilizando o método do hipercubo latino randômico e a segunda utilizando o mesmo método aplicando um técnica de otimização. É possível notar que a amostragem (a) não completa tanto o domínio quanto (b), caracterizando esta última como uma amostragem *space-filling*.

6.3 AQUISIÇÃO DE PONTOS DE SUPORTE

Com uma amostragem inicial determinada através de um hipercubo latino otimizado e com o modelo de Kriging construído, a ideia é definir novos pontos de suporte que minimizem o erro global da estimativa. Uma maneira de realizar a aquisição de novos pontos de suporte é computando o erro quadrático médio (*Mean Squared Error* (MSE) em

inglês), o qual Sacks *et al.* (1989) definem como

$$s^2(\mathbf{x}) = \hat{\sigma}^2 \left[1 - \psi^T \Psi^{-1} \psi + \frac{(1 - \psi^T \Psi^{-1} \mathbf{1})^2}{\mathbf{1}^T \Psi^{-1} \mathbf{1}} \right], \quad (102)$$

onde o terceiro termo está relacionado à incerteza sob a média estimada $\hat{\mu}$ e geralmente é bem pequeno e omitido (FORRESTER; SOBESTER; KEANE, 2008). Para realizações sem ruídos, observa-se que para a n -ésima resposta o peso $\psi^T \Psi^{-1} \psi$ é o valor da unidade, configurando $s^2 = 0$ para os pontos de suportes definidos previamente. De acordo com Jones (2001), é mais conveniente trabalhar com a raiz quadrada do MSE ou erro padrão para medição de incerteza na predição. O MSE pode ser utilizado como estratégia para adicionar novos pontos de suporte, porém é equivalente a preencher as regiões distantes dos mesmos. Dessa forma, não fica preso a um mínimo local da função a ser predita, ao mesmo tempo que pode ser substituída por um método de amostragem inicial com maior número de pontos de suporte.

Outra forma é dar um peso para a influência do erro s , controlando sua importância na exploração. Este método é conhecido como *Lower Confidence Bound* em inglês ou somente LCB (COX; JOHN, 1992), definido por

$$LCB(\mathbf{x}) = \hat{y}(\mathbf{x}) - \mathcal{V}s, \quad (103)$$

onde \mathcal{V} é a constante que controla o balanço da exploração. Quando $\mathcal{V} \rightarrow 0$, $LCB(\mathbf{x}) \rightarrow \hat{y}(\mathbf{x})$ tendendo a um refino local. Por outro lado, quando $\mathcal{V} \rightarrow \infty$, s é maximizado e torna uma exploração global, basicamente o que ocorre quando a Eq. (102) é a métrica para definir novos pontos de suporte. A escolha da constante \mathcal{V} é uma das dificuldade deste método, podendo tender a somente uma exploração global ou somente a um refino local.

Em busca do mínimo da função a ser predita, pode-se computar a probabilidade de obter uma realização $y(\mathbf{x})$ menor que o mínimo de $\mathbf{y}(\mathbf{X})$ até o momento, definida como y_{min} . É possível definir a melhora $I(\mathbf{x})$ para $y(\mathbf{x}) \sim \mathcal{N}(\hat{y}(\mathbf{x}), s^2(\mathbf{x}))$ como

$$I(\mathbf{x}) = \max(y_{min} - y(\mathbf{x}), 0), \quad (104)$$

determinando que há melhora quando $y(\mathbf{x}) < y_{min}$. De outra forma, o método da probabilidade de melhora (KUSHNER, 1964), *Probability of Improvement* em inglês ou somente PI, busca maximizar a probabilidade de um evento que tem $I(\mathbf{x}) > 0$. Utilizando as Eqs. (26) e (27) é possível calcular a probabilidade de $I(\mathbf{x})$ como

$$\begin{aligned} PI(\mathbf{x}) &= P(I(\mathbf{x}) > 0) \\ PI(\mathbf{x}) &= P(y(\mathbf{x}) < y_{min}) \\ PI(\mathbf{x}) &= \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^0 \exp \left[-\frac{(I - y(\mathbf{x}))^2}{2s^2} \right] dI, \end{aligned} \quad (105)$$

utilizando a função erro (ANDREWS, 1998), a Eq. (105) resulta em

$$PI(\mathbf{x}) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{y_{min} - \hat{y}(\mathbf{x})}{s\sqrt{2}} \right) \right]. \quad (106)$$

Como s é nulo em \mathbf{x} não existe a possibilidade de gerar um novo ponto de suporte nos pontos já amostrados. Jones (2001) explica que a Eq. (106) é extremamente sensível a escolha do alvo y_{min} , onde fará uma busca exaustiva próximo do alvo, buscando pontos de suporte infinitesimalmente menor que y_{min} ao invés de pontos com maior ganho e incerteza. Isso é contornado adicionando a variável $\zeta \geq 0$ para um melhor balanço entre a busca local e global, tornando a Eq. (106) em

$$PI(\mathbf{x}) = \frac{1}{2} \left[1 + erf \left(\frac{y_{min} - \hat{y}(\mathbf{x}) - \zeta}{s\sqrt{2}} \right) \right]. \quad (107)$$

Uma forma de definir ζ é iniciando com um valor alto e ir reduzindo ao longo da busca por pontos de suporte.

O PI define pontos de suporte somente pela probabilidade da melhoria, não pela sua magnitude. Uma forma é através da maximização da melhoria esperada, ou EI (*Expected Improvement* em inglês). Segundo Mockus, Tiesis e Zilinskas (1978) a ideia é maximizar a melhoria esperada em relação a y_{min} . O novo ponto de suporte \mathbf{x}^+ para $y(\mathbf{x}) \sim \mathcal{N}(\hat{y}(\mathbf{x}), s^2)$ é definido como

$$\mathbf{x}^+ = argmax E[\max(y_{min} - y(\mathbf{x}), 0)]. \quad (108)$$

Utilizando a Eq. (27) como verossimilhança e determinando o valor esperado de $I(\mathbf{x})$ aplicando a Eq. (28), o $E[I(\mathbf{x})]$ é

$$E[I(\mathbf{x})] = \int_{-\infty}^{\infty} I \frac{1}{\sqrt{2\pi s^2}} \exp \left[-\frac{1}{2} \left(\frac{(y_{min} - I - \hat{y})^2}{s^2} \right) \right] dI, \quad (109)$$

no qual sua solução é

$$E[I(\mathbf{x})] = (y_{min} - \hat{y}) \mathcal{H} \left(\frac{y_{min} - \hat{y}}{s} \right) + s \mathcal{I} \left(\frac{y_{min} - \hat{y}}{s} \right). \quad (110)$$

Onde $\mathcal{H}(\cdot)$ representa a função de distribuição cumulativa de probabilidade e $\mathcal{I}(\cdot)$ é a função de densidade de probabilidade para uma distribuição Gaussiana. O Apêndice G apresenta o desenvolvimento matemático para obter a Eq. (110) partindo da Eq. (109). Observando a Eq. (110), $E[I(\mathbf{x})] = 0$ nos pontos de suporte, pois s também é nulo. O primeiro termo avalia a probabilidade de melhoria de uma predição \hat{y} ser menor que y_{min} , similar ao método PI. O segundo termo representa a busca em incertezas onde a incerteza s é alta, atuando longe dos pontos de suporte. Dessa forma, podemos reescrever a Eq. (110) como

$$E[I(\mathbf{x})] = \begin{cases} (y_{min} - \hat{y}) \mathcal{H} \left(\frac{y_{min} - \hat{y}}{s} \right) + s \mathcal{I} \left(\frac{y_{min} - \hat{y}}{s} \right) & \text{if } s \geq 0 \\ 0 & \text{if } s = 0 \end{cases} \quad (111)$$

As Figs. 31 e 32 representam a evolução do processo EI para a função da Eq. (99), onde a função a ser predita é definida com 6 iterações.

Figura 31 – Processo de aquisição de novos pontos para a função da Eq. (99) utilizando como função de aquisição o EI.

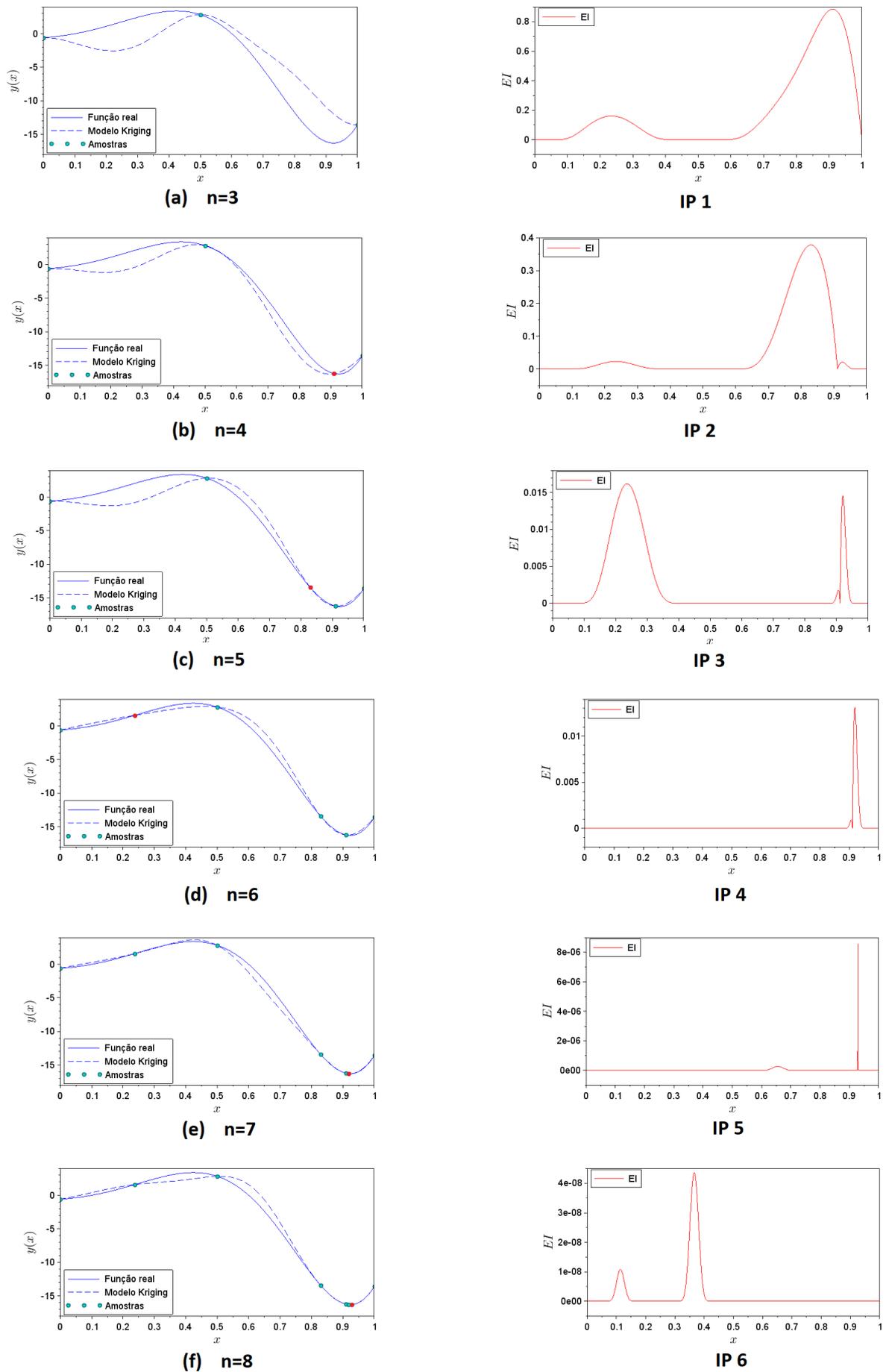
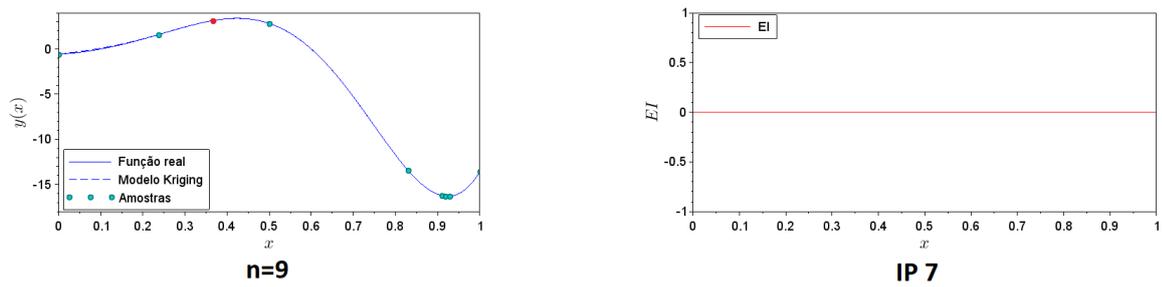


Figura 32 – 6^a iteração para o processo EI aplicado a função da Eq. (99).

Fonte: Autor.

7 MATERIAIS E MÉTODOS

Um exemplo de aplicação do método de Kriging é a geração de modelos substitutos para funções caixas pretas ou *black-box function* em inglês. Este tipo de função relaciona uma entrada \mathbf{x} (*input* em inglês) com uma saída $y(\mathbf{x})$ (*output* em inglês). Funções caixas pretas são conhecidas por não estarem disponível para algum tipo de tratamento matemático, como o cálculo de sua derivada. Problemas resolvidos em *softwares* de elementos finitos comerciais ou uma análise modal feita em laboratório podem ser classificadas como este tipo de função. Ambos os casos demandam um longo tempo de processamento, no caso do *softwares* de elementos finitos, ou preparação e execução, no caso da análise modal em laboratório.

Para evitar um grande número de análises custosas, é possível utilizar modelos substitutos criados pelo método de Kriging para uma predição de rápida e fácil exploração. Desse modo, permite-se a construção de um problema de otimização conforme a Eq. (8), porém com a função objetivo e as restrições definidas por modelos substitutos. Baseado na Eq. (8) pode-se definir um problema de otimização na forma

$$\begin{aligned} \min_{\mathbf{x} \in S} \quad & \hat{y}(\mathbf{x}) \\ \hat{h}_i(\mathbf{x}) = 0 \quad & i = 1, \dots, n_h \\ \hat{g}_j(\mathbf{x}) \leq 0 \quad & j = 1, \dots, n_g \\ x_i^l \leq x_i \leq x_i^u, \end{aligned} \tag{112}$$

onde $\hat{y}(\mathbf{x})$ representa a função objetivo em forma de modelo substituto (sucendo $y(\mathbf{x})$), $\hat{h}_i(\mathbf{x})$ e $\hat{g}_j(\mathbf{x})$ representam as restrições de igualdade e desigualdade respectivamente em forma de modelo substituto (sucendo $h_i(\mathbf{x})$ e $g_i(\mathbf{x})$). Com a Eq. (112) evita-se a criação de uma rede densa de respostas $\mathbf{y}(\mathbf{x})$ do modelo de alto custo computacional. A criação desses modelos substitutos ocorrem de forma independente conforme fluxograma da Fig. 33.

Tendo em mãos os modelos substitutos de cada função que compõe o problema de otimização é possível definir técnicas de otimização para definir o ótimo \mathbf{x}^* de $\hat{y}(\mathbf{x})$. É possível aplicar técnicas de otimização baseadas em gradientes, pois a função de predição do método de Kriging na Eq. (97) é conhecida. Porém, a utilização dessas técnicas pode acarretar em diversos riscos. Um exemplo é a função a ser predita ter vários mínimos locais ou até mesmo não ser contínua, portanto a seleção de um ponto inicial \mathbf{x}_0 pode fazer com que o processo fique preso em um mínimo local. Outro risco é a garantia que a derivada do modelo substituto terá o mesmo comportamento da derivada da função a ser predita. Por exemplo, tomando a derivada da função da Eq. (99) como

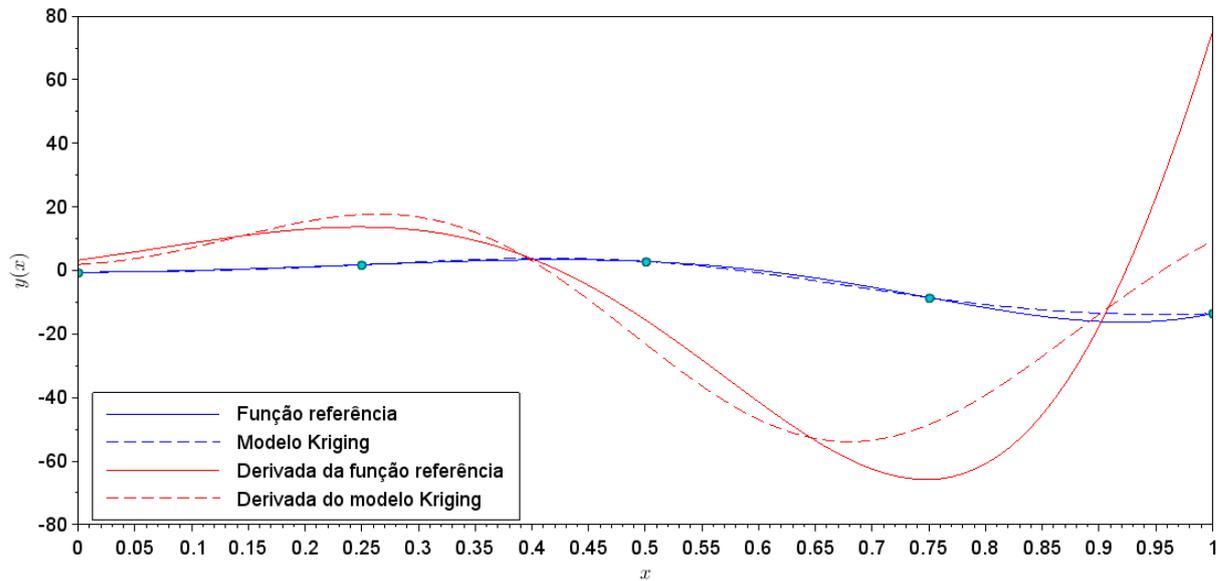
$$\frac{\partial y(x)}{\partial x} = \pi \exp(\pi x) \left\{ 2 \cos \left[2\pi \left(x - \frac{1}{10} \right) \right] + \sin \left[2\pi \left(x - \frac{1}{10} \right) \right] \right\}. \tag{113}$$

e da Eq. (97) na forma

$$\frac{\partial \hat{y}(x)}{\partial x} = [\theta p |x_j^{(i)} - x|^{(p-2)} (x_j^{(i)} - x) \exp(-\theta |x_j^{(i)} - x|^p)] \Psi^{-1}(\mathbf{y} - \mathbf{1}\hat{\mu}). \tag{114}$$

Pode-se verificar pela Fig. 34 que há divergência entre a aproximação da derivada do modelo substituto e da função de referência, mesmo nos pontos de suporte. Para evitar

Figura 34 – Diferença da derivada da função de referência da Eq. (113) (com 5 pontos de suporte) e a derivada da predição de Kriging da Eq. (114).



Fonte: Autor.

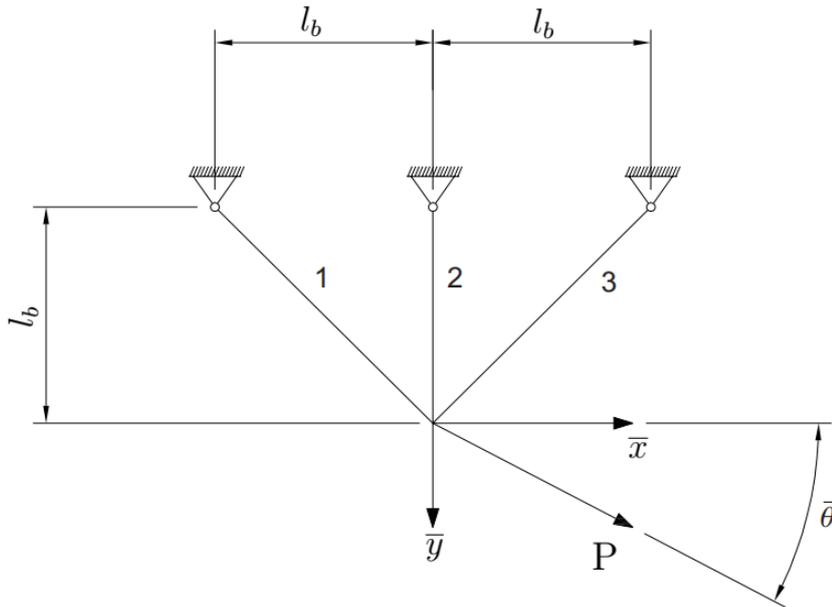
e

$$g_3 = \sigma_3 = \frac{1}{\sqrt{2}} \left[-\frac{P_{\bar{x}}}{A_1} + \frac{P_{\bar{y}}}{(A_1 + \sqrt{2}A_2)} \right] \leq \sigma_{a3}, \quad (118)$$

onde $P_{\bar{x}}$ e $P_{\bar{y}}$ representam as componentes de P nos respectivos eixos \bar{x} e \bar{y} . O resultado das restrições descritas acima devem ser menor ou igual σ_a , a qual representa a tensão admissível do material selecionado. Além das restrições de tensões, restrições laterais para as variáveis de projeto são determinadas como $1 \leq A_1, A_2 \leq 10$ [in^2]. Tomando as considerações da Tab. 3, a solução ótima é $A_1^* = 6,31$ [in^2] e $A_2^* = 3,26$ [in^2], enquanto o valor da função objetivo é de 21,11 [lb]. A Fig. 36 mostra o gráfico de contorno da função objetivo contida na Eq. (115) e das restrições das Eqs. (116) a (118), bem como a solução ótima \mathbf{x}^* do sistema. Os valores das restrições em \mathbf{x}^* são 5000 [psi], 3663,22 [psi] e -1338 [psi] respectivamente, onde a restrição g_1 está ativa. O mínimo da função objetivo é $\mathbf{x}_{Vol}^* = \{1, 1\}$, para $g_1(\mathbf{x})$ o mínimo é $\mathbf{x}_{\sigma_1}^* = \{10, 10\}$, para $g_2(\mathbf{x})$ é $\mathbf{x}_{\sigma_2}^* = \{10, 10\}$ e por fim $g_3(\mathbf{x})$ tem o mínimo de $\mathbf{x}_{\sigma_3}^* = \{1, 10\}$.

Um dos passos do fluxograma da Fig. 33 é a aquisição de novos pontos de suporte por meio de uma função de aquisição. Através do problema selecionado pode-se realizar uma análise da melhor função de aquisição para a metodologia proposta. Portanto, os cenários A,B e C são definidos para avaliar as funções de aquisição s^2 , LCB, PI e EI apresentadas no Cap. 6. A diferença entre os cenários é a quantidade de pontos iniciais do hipercubo latino. No cenário A o hipercubo latino otimizado tem $n_i = 15$ pontos de suportes iniciais, no cenário B possui $n_i = 10$ pontos de suporte e no cenário C tem $n_i = 5$. O problema é executado 5 vezes para cada cenário, onde o hipercubo latino é limitado as restrições laterais.

Figura 35 – Problema referência de treliça com 3 barras encontrado em (ARORA, 2004).



Fonte: Autor.

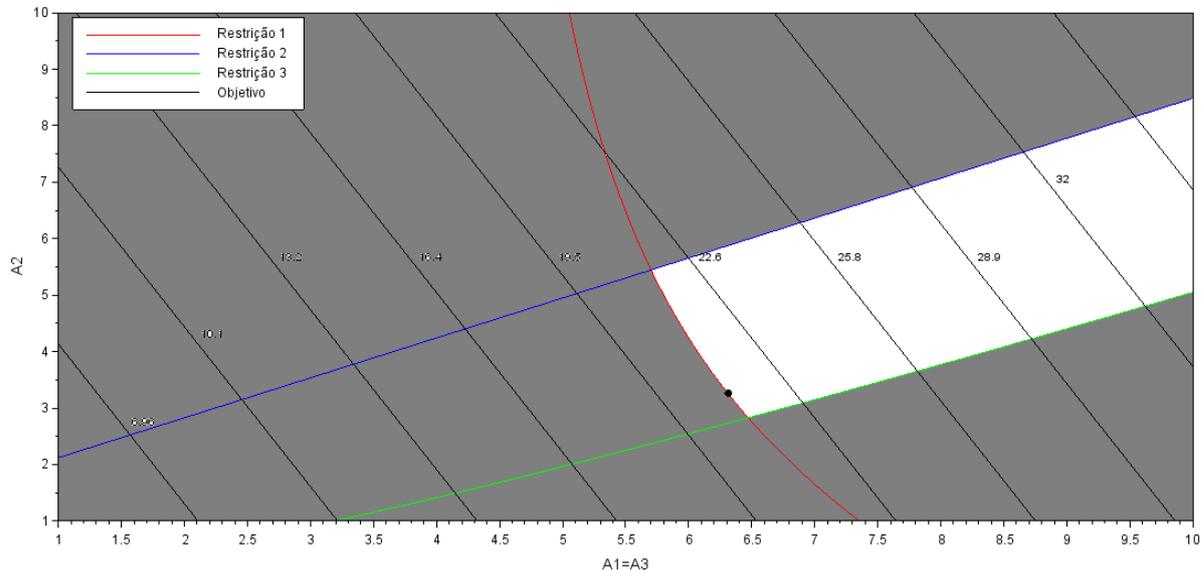
A partir da criação dos modelos de Kriging, as funções de aquisição são aplicadas uma de cada vez para avaliar o respectivo comportamento. Os valores dos parâmetros κ são $0.001 \leq \theta \leq 100$ e $p = 2$, onde θ é definido através de um Algoritmo Genético com uma população de 150 indivíduos (BAUDIN; COUVERT; STEER, 2010). Em relação a parada do algoritmo, o orçamento destinado para cada função de aquisição é de 20 iterações. Adicionalmente para os métodos EI, PI e s^2 é definido uma tolerância de parada quando seus valores chegam a 1×10^{-2} . Para o método LCB foi considerado a constante $\mathcal{V} = 1$. Após a obtenção dos modelos substitutos do problema de otimização pelo método de Kriging, a técnica de otimização PSO é utilizada para encontrar o ótimo \mathbf{x}^* . Para tal, é definida uma população de 50 partículas, com $c_1 = c_2 = 1,5$ e w decrescendo ao longo das iterações e as restrições são tratadas conforme (DEB, 2000), explicado no Cap. 3.

O Apêndice H apresenta os resultados para as 5 execuções dos cenários A, B e C. Em uma análise inicial a função objetivo possui um comportamento linear, onde todas as funções de aquisições adicionaram poucos pontos de suporte para o modelo substituto. As funções de restrições possuem um comportamento não-linear, de forma geral é adquirido um maior número de pontos de suporte para essas funções. O cenário C demonstrou o pior desempenho em todos os 5 testes, justificado pelo pequeno número de pontos de

Tabela 3 – Dados do problema de otimização - adaptado de (ARORA, 2004).

Estrutura	Dados
Tensão admissível (barras 1 e 3) - σ_{a1} e σ_{a3} [psi]	5000
Tensão admissível (barra 2) - σ_{a2} [psi]	20000
Altura - l_b [in]	10
Carga - P [lb]	40000
Peso específico - ρg [lb/in ³]	0,10

Figura 36 – Gráfico de contorno da função objetivo e das restrições. O ponto preto representa o *design* ótimo, $A_1^* = 6,31$ [in²] e $A_2^* = 3,26$ [in²], enquanto a área cinza representa a região inviável.



Fonte: Autor.

Tabela 4 – \mathbf{x}^* para o cenário B do teste 2.

Função de aquisição	x_1	x_2
s^2	6.33	3.14
LCB	6.29	3.26
PI	6.34	3.15
EI	6.32	3.1
HCL	6.28	3.25

suporte para LCB, PI e EI. Mesmo quando utilizado s^2 que adiciona 20 pontos de suporte ao modelo substituto apresentou um desempenho aquém do esperado. Ressalta-se que as funções de aquisição LCB, PI e EI exploram a região do ótimo de cada função do problema com poucas iterações, mesmo nos cenários B e C. Entretanto, não são efetivos para determinar o mínimo do problema de otimização, pois dependem dos pontos iniciais definidos pelo hipercubo latino. Essa observação é apresentada na Tab. 4, onde o cenário B do teste 2 é executado somente com o hipercubo latino, encontrando \mathbf{x}^* próximo dos valores encontrados quando utilizado funções de aquisição.

A aquisição de novos pontos de suporte por s^2 se mostrou custosa, pois o orçamento de 20 iterações foi atingido em todos os cenários e execuções. No cenário A, este método traz problemas ao deixar o domínio denso de pontos de suporte, tornando a matriz Ψ mal condicionada e dificultando a inversão da mesma pelo método de Cholesky. A Fig. 37 apresenta o modelo de Kriging para g_2 com a disposição dos pontos de suporte para o cenário A do teste 5, comparando as funções de aquisição s^2 e PI. O s^2 adiciona novos pontos de suporte nas bordas do domínio onde seu valor é maior devido a redução da correlação pelo parâmetro θ . Esse comportamento não auxilia na busca pelo ótimo global que está longe dessas regiões. Portanto, a adição de novos pontos de suporte por s^2

Tabela 5 – Média aritmética do ótimo para cada função de aquisição dividida por $\mathbf{x}^* = \{6.31, 3.26\}$ para o cenário A.

Função de aquisição	m_{x_1} [%]	m_{x_2} [%]
s^2	-0.6	38
LCB	1.11	4.6
PI	2.66	-11.9
EI	1.31	-4.1

apresenta menor acurácia em relação aos outros métodos mesmo com maior número de pontos de suportes.

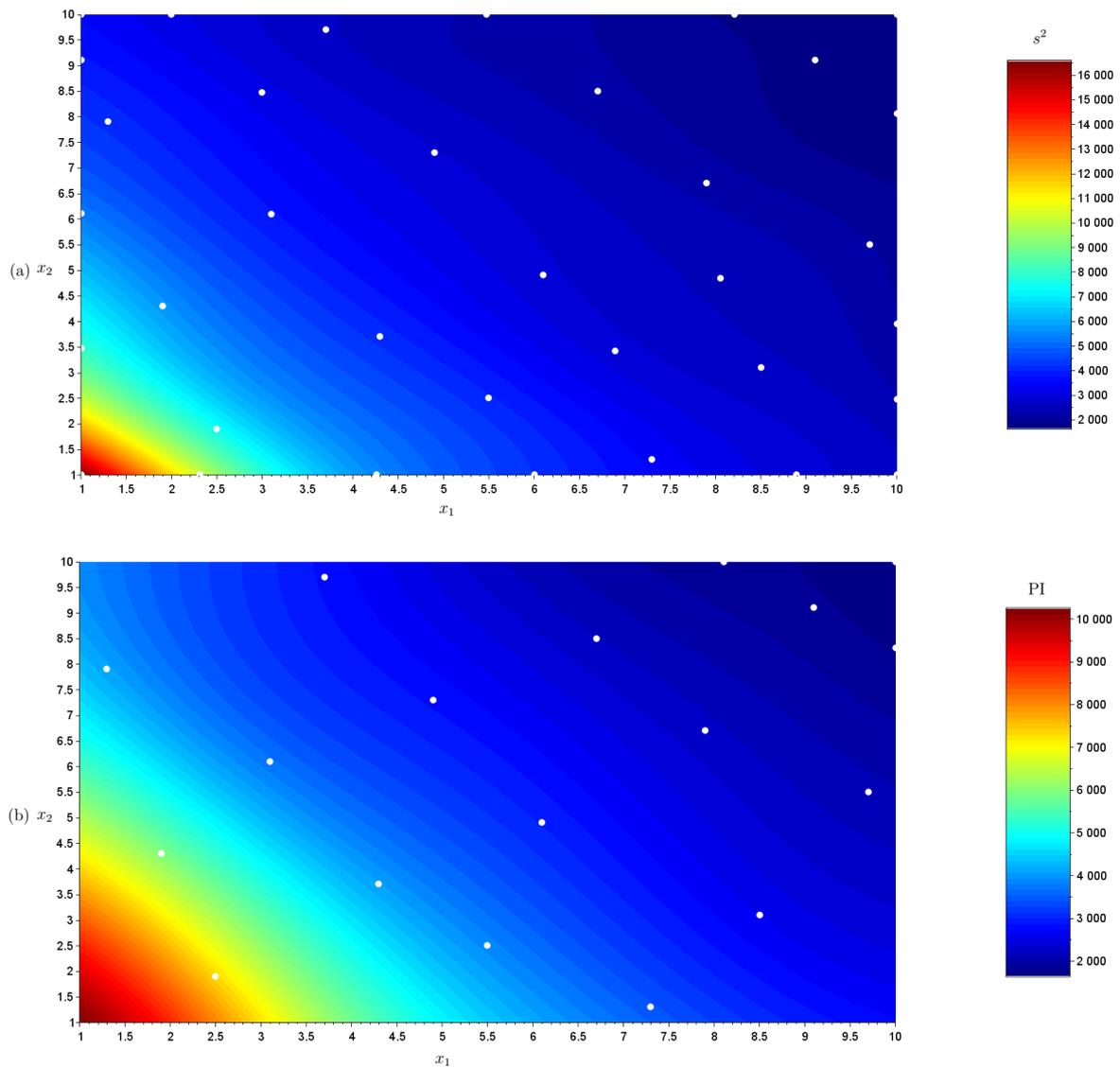
De forma geral, o LCB se caracteriza por ter mais pontos de suporte que o PI devido a constante $\mathcal{V} = 1$, realizando uma busca mais local. Mas ambos se mostram eficiente em definir o mínimo de cada função com a aquisição de poucos pontos de suporte, principalmente em relação ao s^2 . Ambos possuem acurácia similar ao encontrar o mínimo global sendo dependentes da disposição de pontos definidos inicialmente pelo hipercubo latino.

Percebe-se que o EI apresenta uma variância estimada s^2 menor ao final de cada execução em relação aos métodos PI e LCB. Esse comportamento é gerado principalmente pela influência da segunda parcela da Eq. (110), a qual leva em consideração o erro na forma de s para um balanço de busca de novos pontos de suporte. A Fig. 38 mostra um comparativo da variância s^2 para a função g_1 , considerando o cenário B do teste 4.

Para calcular a acurácia de cada função de aquisição, após a execução do PSO a média aritmética de todas as execuções é computada e dividida por \mathbf{x}^* do problema, esses valores são apresentados na Tab. 5. Levando em consideração somente o cenário A, o EI apresentou a melhor aproximação em relação ao ótimo do problema. O LCB também possui uma aproximação aceitável, porém a definição da constante \mathcal{V} pode não ser adequada a outros problemas. O PI mostrou uma grande divergência em relação a variável de projeto x_2^* , assim como o s^2 . Novamente apresentando que a redução da variância dos modelos substitutos do problema de otimização não é garantia de uma melhor acurácia em relação ao ponto ótimo.

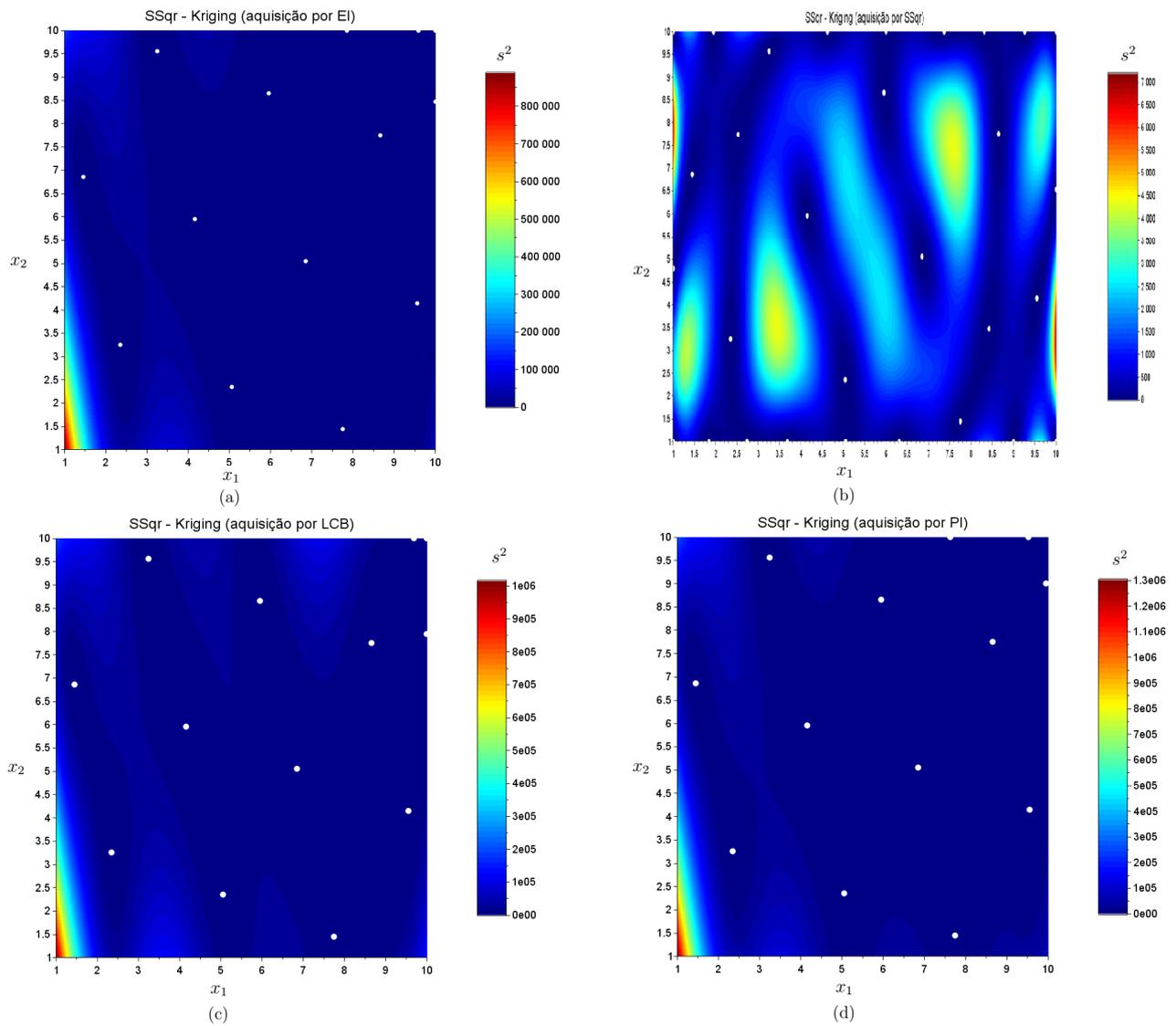
O método EI será utilizado como função de aquisição de novos pontos de suporte para a metodologia do fluxograma da Fig. 33 após os testes realizados. Dessa forma, evita-se a necessidade da consideração de algum fator de balanço entre a busca local e global como a constante \mathcal{V} para o LCB. Também será considerado um hipercubo latino com $n_i = 15$ pontos de suporte como amostragem inicial, pois observou-se uma melhor relação de custo computacional por preenchimento do domínio.

Figura 37 – Comparativo entre os modelos substitutos da função g_2 utilizando os métodos s^2 (a) e PI (b), mostrando a disposição dos pontos de suporte (círculos branco) para do cenário A do teste 5.



Fonte: Autor.

Figura 38 – Comparativo de s^2 entre os modelos substitutos da função g_1 para o cenário B do teste 4. Considerando os métodos EI (a), s^2 (b), LCB (c) e PI (d), após 20 pontos de suporte adquiridos.

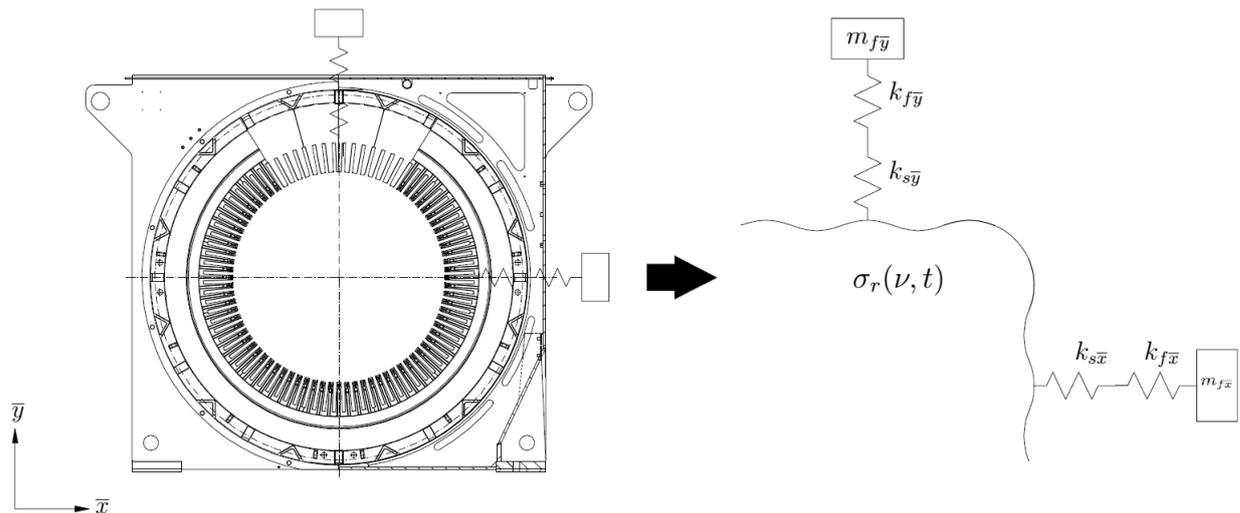


Fonte: Autor.

8 FORMULAÇÃO DO PROBLEMA E ESTRATÉGIA DE SOLUÇÃO

A vibração em duas vezes a frequência da rede apresentada no Cap. 2 é gerada no estator e transmitida para a carcaça, que por sua vez transmite para as tampas e posteriormente aos mancais. Nos mancais estão posicionados os transdutores de vibração e caso identifiquem uma velocidade de vibração maior que o valor definido por norma, é realizado o desligamento da MEG. Analisando a transmissão da vibração com a estrutura de uma MEG, a mesma está relacionada com as características modais, principalmente com a rigidez radial do conjunto estator e carcaça. Essa transmissão de vibração também está relacionada com a amplitude da força magnética líquida, proveniente da pressão magnética radial apresentada na Eq. (5). Dessa maneira, observando axialmente ao eixo \bar{z} (saindo da folha), podemos construir um sistema com dois graus de liberdade para representar essa transmissão, conforme ilustrado na Fig 39.

Figura 39 – Sistema com dois graus de liberdade representando a transmissão da vibração entre estator e carcaça.

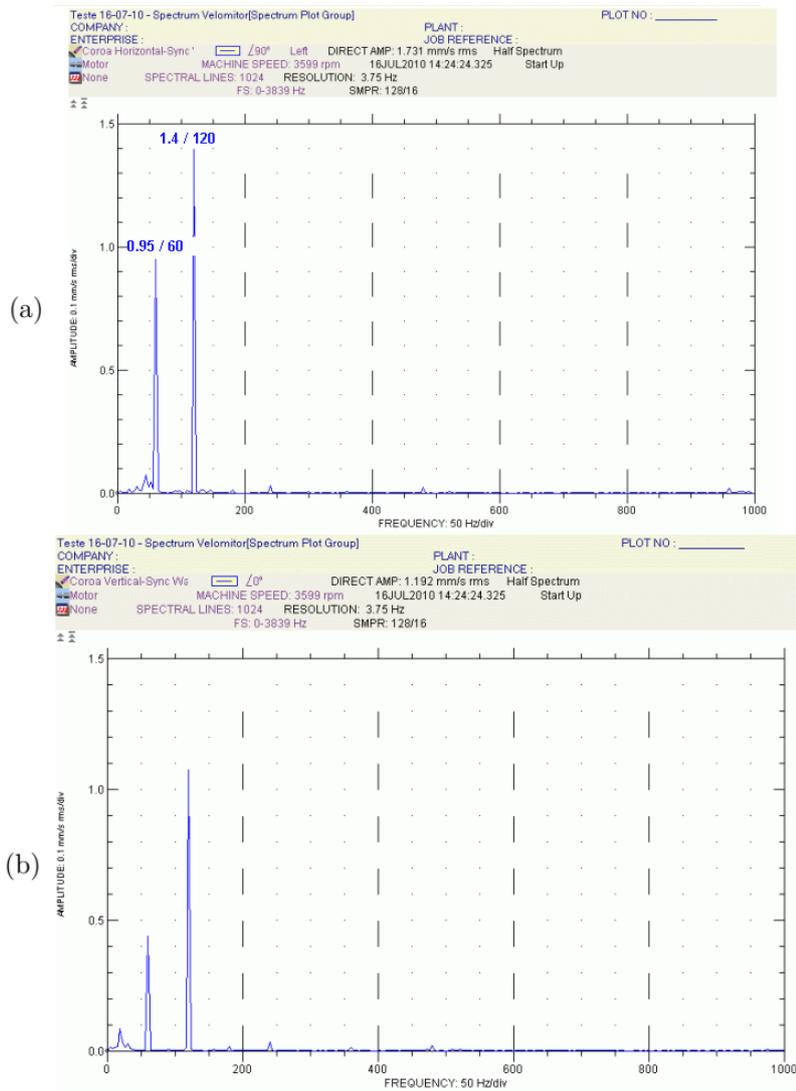


Fonte: Autor.

Onde $\sigma_r(\nu, t)$ representa a pressão magnética radial no entreferro conforme Eq. (5), $k_{s\bar{x}}$, $k_{f\bar{x}}$ e $m_{f\bar{x}}$ representam respectivamente a rigidez do estator, a rigidez da carcaça e a massa do sistema na direção \bar{x} e $k_{s\bar{y}}$, $k_{f\bar{y}}$ e $m_{f\bar{y}}$ representam respectivamente a rigidez do estator, a rigidez da carcaça e a massa do sistema na direção \bar{y} .

A medição da vibração de forma experimental auxilia no entendimento desse fenômeno. Para uma melhor compreensão, uma medição comparativa utilizando transdutores é realizada no diâmetro externo do estator e nos mancais de uma máquina 2 polos. As informações podem ser divididas em quatro espectros na banda de duas vezes a frequência da rede. A Fig. 40 ilustra os espectros de velocidades no diâmetro externo do estator na direção \bar{x} e \bar{y} , enquanto a Fig. 41 mostra os espectros de vibração no mancal dianteiro também para as mesmas direções. Analisando os espectros é possível confirmar a menor rigidez do sistema na direção horizontal \bar{x} , onde houve a transmissão da vibração de aproximadamente 72%, enquanto na direção vertical \bar{y} foi de 23% da vibração originada no estator. Salienta-se que nesse caso a vibração nos mancais não ultrapassa os valores

Figura 40 – Espectro de velocidade de vibração no diâmetro externo no estator, na direção \bar{x} (a) e na direção \bar{y} (b).



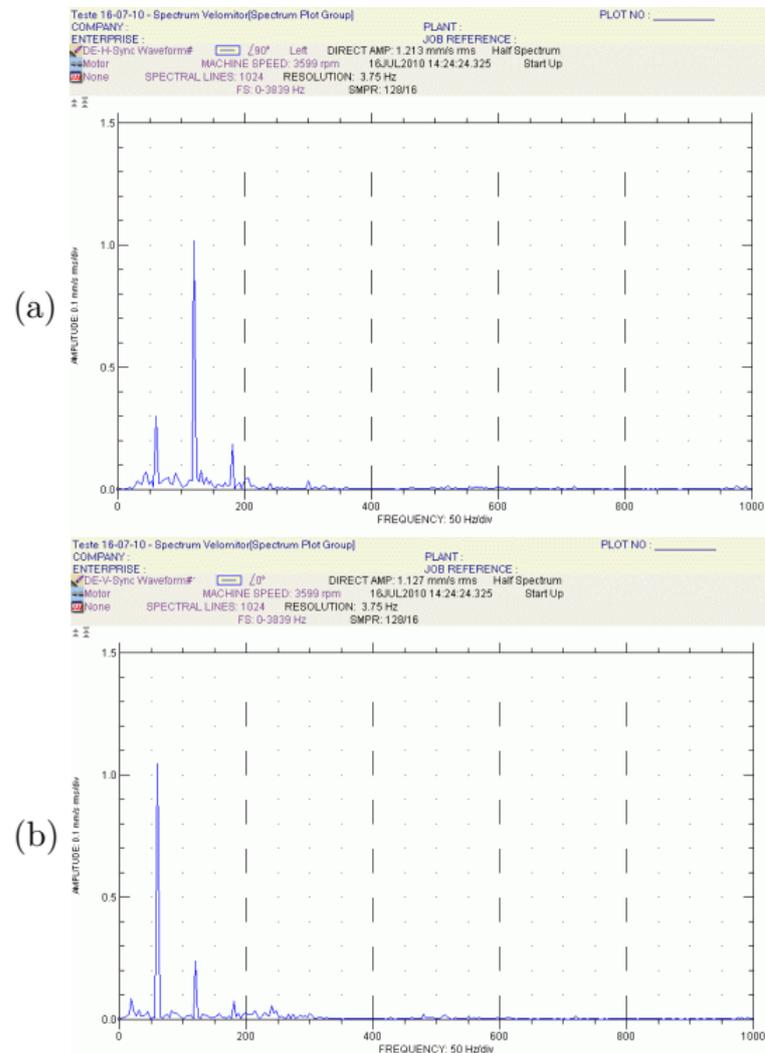
Fonte: Autor.

recomendados por norma. Isso pode ser observado comparando o espectro da Fig. 41 com as normas apresentadas no Cap. 2.

O componente de interface entre estator e carcaça tem um papel importante na transmissão da vibração para o mancal, pois modifica a rigidez entre estator e carcaça. Muitas vezes essas interfaces não permitem a atenuação da vibração proveniente do estator, sendo construídas com o intuito de facilitar a fabricação e a refrigeração da MEG. Devido a isso, a interface entre estator e carcaça é realizada de duas maneiras: por longarinas ou por contra placas, conforme ilustra a Fig 42.

Na interface com longarinas, as quais estão distribuídas circularmente na carcaça, o pacote de chapas do estator é fixado por interferência mecânica. Devido a esta interferência o diâmetro externo das chapas do estator é maior que o diâmetro usinado nas longarinas. A inserção do estator na carcaça ocorre por meio de prensagem, posteriormente é pinado radialmente evitando movimentação axial. Já na interface com contra placas, a carcaça

Figura 41 – Espectro de velocidade de vibração no mancal dianteiro, na direção \bar{x} (a) e na direção \bar{y} (b).

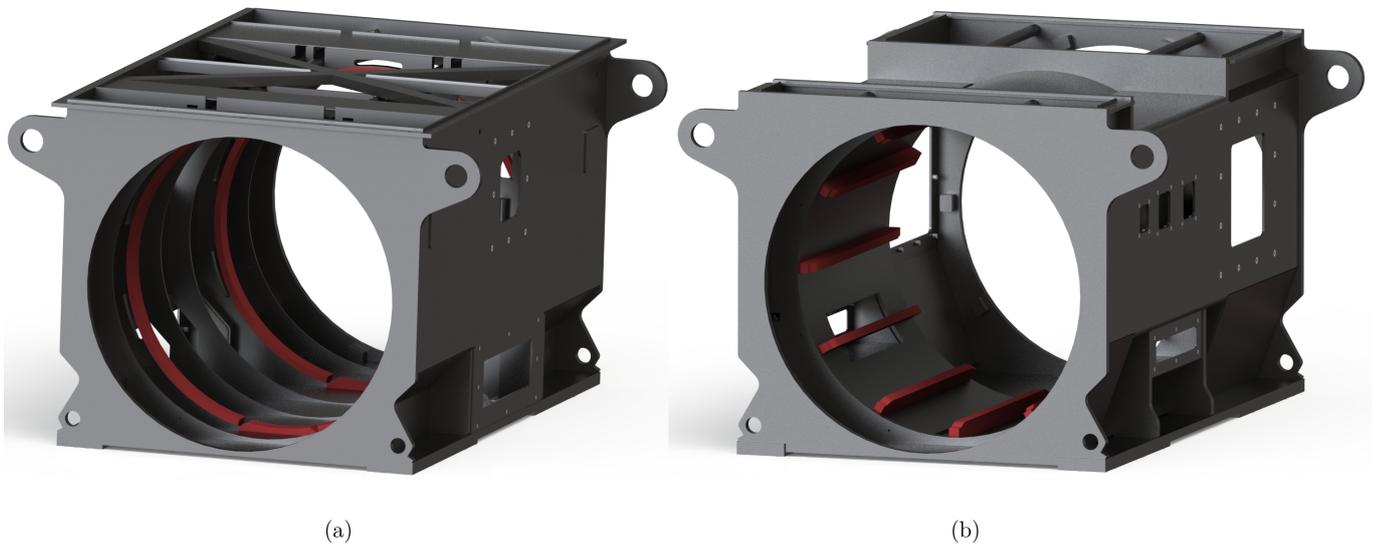


Fonte: Autor.

possui placas no sentido radial da máquina onde o estator é apoiado. A contra placa dianteira possui um assento usinado para definição da posição axial do pacote de chapas do estator e após este posicionamento é realizado a soldagem na placa dianteira e traseira.

Penniman e Taylor (1941) foram um dos primeiros a estudar alterações na interface estator e carcaça, criando uma estrutura mais flexível para aplicação em máquinas síncronas 2 polos. Laffoon e Rose (1940) experimentam suportes flexíveis para este fim e ao longo do tempo foram desenvolvidas várias patentes referente a esta interface. Fonçatti (2018) aplica um sistema com 6 graus de liberdade encontrado em (HARRIS C. M.; PIERSOL, 2002) com o intuito de analisar numericamente a relação entre estator e carcaça. Onde é comparado um modelo reduzido em elementos finitos com um modelo analítico, ambos com 6 graus de liberdade. Gonçalves (2012) utiliza o método de elementos finitos para determinar uma metodologia para calcular os três primeiros modos de vibrar de uma MEG de pequeno porte. Neste mesmo trabalho é apresentado o modo lateral de vibração próximo a duas vezes a frequência da rede.

Figura 42 – Interface entre estator e carcaça em vermelho, constituída por meio de contra placas (a) e de longarinas (b).

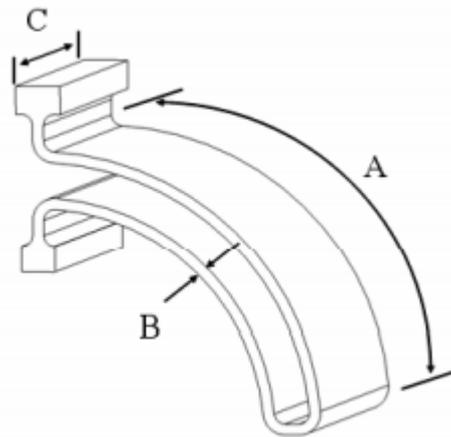


Fonte: Autor.

De forma mais geral, há outros trabalhos realizados em MEGs porém não relacionados com a transmissão de vibração entre estator e carcaça, como o trabalho apresentado por (WAPPLER, 2014). O qual estuda a redução de vibração de uma MEG pequeno porte através da otimização da flange de fixação da mesma. Também há várias patentes relacionadas a MEGs, mas com outras aplicações, como Holter (1993) em que apresenta peças de neoprene entre a carcaça e o trocador de calor. Dessa maneira, reduzindo a transmissão de vibração da carcaça para o trocador de calor que é um componente de menor rigidez. Neste âmbito é possível identificar mecanismos utilizados em outras aplicações que podem ser empregados em máquinas elétricas. Por exemplo, Radle (1997) utiliza juntas isoladoras em furadeiras que também podem ser aplicadas na interface de estator e carcaça de uma MEG.

O desenvolvimento de um mecanismo entre estator e carcaça é proposto para a redução de vibração excitada a duas vezes a frequência da rede. O objetivo do mecanismo é afastar a frequência natural do sistema em relação a frequência de excitação, reduzindo a velocidade de vibração nos mancais. O componente apresentado por Hwang *et al.* (2010) na Fig. 43 serve como referência de conceito para o desenvolvimento do mecanismo proposto. A partir deste perfil o mecanismo é desenvolvido para atenuar a transmissão de vibração entre estator e carcaça, sendo o mesmo instalado em cada extremidade do estator conforme Fig. 44.

Uma MEG que constatou vibração excessiva excitada em duas vezes a frequência da rede durante seu ensaio é utilizada como estudo de caso. Essa MEG possui características conforme a Tab. 6 e a interface entre estator e carcaça é constituída por meio de longarinas. Inicialmente buscou-se a compreensão das características modais e padrões de vibração da solução atual por meio de experimentos em laboratório e posteriormente o ajuste do modelo numérico. Para tal, uma análise modal experimental para determinar os modos de vibrar foi realizada (FU; HE, 2001). Também foi executada uma ODS (*Operating Deflection Shape*

Figura 43 – Mecanismo de (HWANG *et al.*, 2010).

Fonte: (HWANG *et al.*, 2010).

em inglês) na lateral da carcaça para conhecer o padrão de vibração (RICHARDSON, 1997). Para esses procedimentos foram utilizados os equipamentos e *software* apresentados na Tab. 7. Nesses experimentos a máquina é fixada de forma rígida à base através de parafusos e são realizadas as marcações onde são posicionados os transdutores de vibração conforme Fig. 45. O resultado da análise modal experimental apresentou um modo de vibrar torcional em 114 [Hz]. Esse valor é próximo a fonte de excitação, correspondendo ao dobro de 60 [Hz]. O padrão de vibração da lateral da carcaça encontrada na ODS é apresentado na Fig. 46, onde é mostrado um maior deslocamento na parte traseira superior na direção de menor rigidez \bar{x} . Onde a vibração no mancal traseiro atingiu um valor de 4,3 [mm/s], valor acima do estabelecido por norma, neste caso é considerado a NEMA MG-1.

Tabela 6 – Principais características da MEG do estudo de caso.

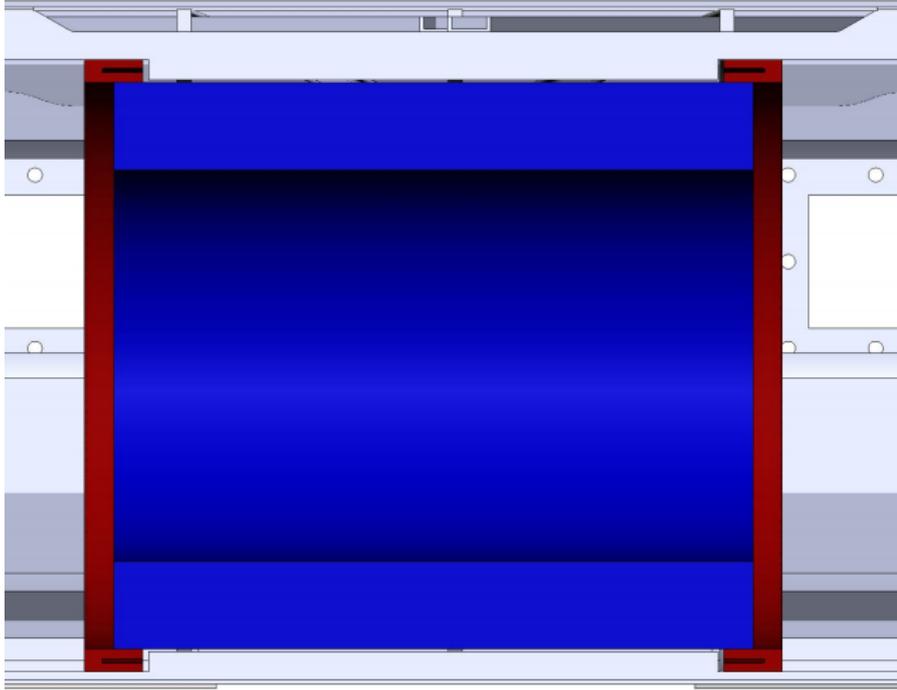
Potência [HP]	7300
Tensão [V]	660
Frequência [Hz]	60
Número de polos	2

Tabela 7 – Equipamentos e *softwares* utilizados para análise modal experimental.

Equipamento/ <i>Software</i>	Descrição
Notebook	7300
<i>Software</i> para coleta de dados	B&K Pulse Labshop v0.0
Módulo para coleta de dados	B&K Modal Analysis MTC Hammer
<i>Software</i> para animação	VIBRANT ME'scopeVES Modal Pro
Acelerômetro	ENDEVCO - modelo 752 A 12, ns. 11494 e sensor 10mV/ms ²
Martelo de impacto	ENDEVCO - modelo 2302-5, ns. 2782 e sensor 1.16mV/N
Tipo de ponta do martelo	Nylon

Um modelo numérico computacional em *software* de elementos finitos é criado para caracterizar o mesmo modo e padrão de vibração. As características e propriedades

Figura 44 – Mecanismo proposto (em vermelho) instalado nas extremidades do estator.



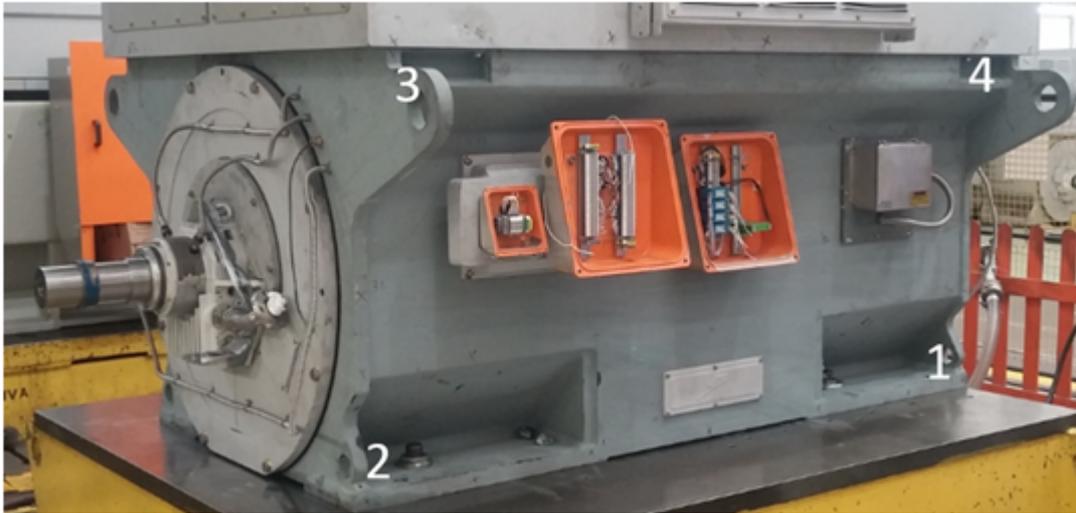
Fonte: Autor.

desse modelo numérico podem ser encontradas em Fonçatti (2018). Após o processamento encontrou-se o mesmo padrão de vibrar e a frequência natural de $114 [Hz]$ como apresentado na Fig. 47. O valor de $114 [Hz]$ é o mesmo apresentado na análise modal experimental, bem como os deslocamentos semelhantes aos da Fig. 46. Esse modelo numérico computacional não contempla o rotor da MEG pois o mesmo vibra de forma independente dos componentes estacionários devido ao filme de óleo dos mancais hidrodinâmicos. Os componentes periféricos como o trocador de calor são representados como pontos de massa, não sendo visíveis na prototipagem virtual. Assim, o modelo numérico é validado, tendo seu comportamento na região de duas vezes a frequência da rede semelhante ao experimentado na análise modal e da ODS.

O modelo numérico computacional da solução atual é utilizado como referência para criar o modelo numérico computacional da solução proposta. A única alteração é o recorte nas longarinas para acomodar o mecanismo proposto, mantendo as características do modelo atual. Esse modelo com o mecanismo proposto é utilizado como função caixa preta na metodologia da Fig. 33 e posteriormente aplicando o método de Kriging para a criação dos modelos substitutos. Dessa forma, é proposto um vetor \mathbf{x} com 3 variáveis de projeto x_1 , x_2 e x_3 conforme apresentado na Fig. 48. Essas variáveis de projeto são definidas por meio de análise geométrica do espaço disponível para a acomodação do mecanismo. O diâmetro interno do mecanismo possui um valor fixo de $980 [mm]$ devido a padronização do diâmetro externo do estator. O diâmetro externo do mecanismo tem seu valor fixo em $1060 [mm]$, mantendo uma simetria radial na altura das longarinas. Na parte interna do mecanismo optou-se por manter dois raios fixos de $2,5 [mm]$, simulando o raio da ferramenta do processo de fabricação. A Tab. 8 apresenta os valores máximos e mínimos para cada variável de projeto do vetor \mathbf{x} .

A faixa mínima de $50 [mm]$ para a variável x_1 é definido pela profundidade máxima

Figura 45 – Principais pontos medidos durante a ODS (*Operating Deflection Shape*).



Fonte: Autor.

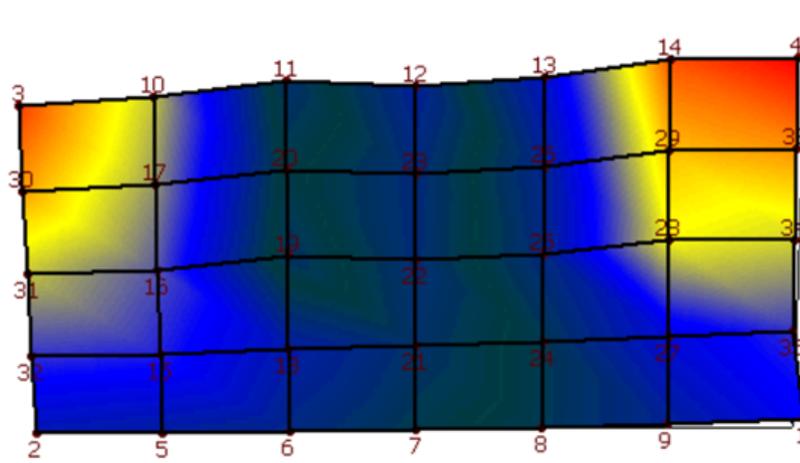
Tabela 8 – Tabela com o valor máximo e mínimo de cada variável de projeto.

Variável	Valor máximo [mm]	Valor mínimo [mm]
x_1	70	50
x_2	1010	990
x_3	1050	1030

que a ferramenta pode alcançar sem ocorrer vibrações excessivas no suporte da mesma. Enquanto a faixa máxima da variável x_2 de 1010 [mm] e a faixa mínima da variável x_3 de 1030 [mm] são definidos para que o espaço mínimo entre as duas partes do mecanismos possibilite a entrada de ferramenta do processo de fabricação.

Além das variáveis de projeto \mathbf{x} é necessário definir a função objetivo e as restrições que vão compor o problema de otimização com modelos substitutos. A função objetivo é definida para atender uma frequência natural alvo $\bar{\omega}$. A restrição da velocidade de vibração nos mancais são de interesse devido aos valores estabelecidos por norma. Como a rigidez do sistema é menor na direção \bar{x} , a velocidade de vibração nos mancais nessa direção é restrita a 1.8 [mm/s]. Outra restrição está relacionada a deflexão do estator na direção \bar{y} , onde deve haver o máximo de 5% do valor do entreferro, evitando assim desbalanceamento eletromagnético entre o estator e rotor. Por último, o material para o mecanismo proposto é aço carbono, desta forma as tensões mecânicas envolvidas são restritas a 220 [MPa] para que não haja deformações permanentes durante a operação. Portanto, podemos definir o problema de otimização do estudo de caso baseado na Eq. (112) como

Figura 46 – Diagrama de cores apresentando o padrão de vibração na frequência de 114 [Hz]. Os pontos 1,2,3 e 4 representam as mesmas posições que os pontos da Fig. 45.

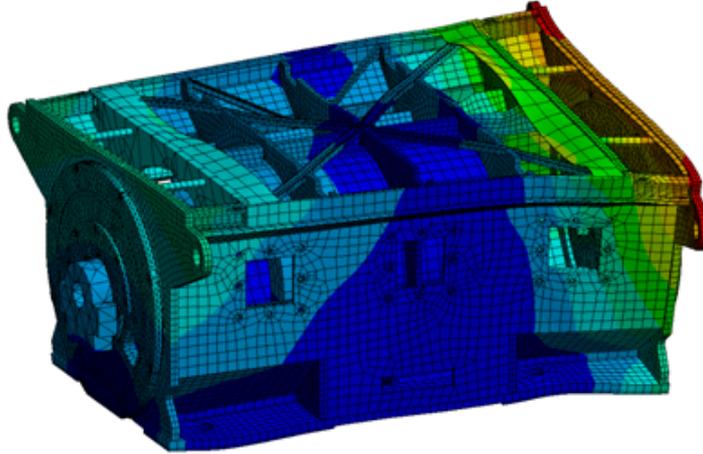


Fonte: Autor.

$$\begin{aligned}
 & \min_{\mathbf{x} \in \mathcal{S}} [\omega(\mathbf{x}) - \bar{\omega}(\mathbf{x})]^2 \\
 & g_1(\mathbf{x}) = \dot{u}_{\bar{x}de}(\mathbf{x}) \leq 1.8 \text{ [mm/s]} \\
 & g_2(\mathbf{x}) = \dot{u}_{\bar{x}nde}(\mathbf{x}) \leq 1.8 \text{ [mm/s]} \\
 & g_3(\mathbf{x}) = u_{\bar{y}est}(\mathbf{x}) \leq 0.25 \text{ [mm]} \\
 & g_4(\mathbf{x}) = \sigma_{de}(\mathbf{x}) \leq 220 \text{ [MPa]} \\
 & g_5(\mathbf{x}) = \sigma_{nde}(\mathbf{x}) \leq 220 \text{ [MPa]} \\
 & 50 \leq x_1 \leq 70 \text{ [mm]} \\
 & 990 \leq x_2 \leq 1010 \text{ [mm]} \\
 & 1030 \leq x_3 \leq 1050 \text{ [mm]}.
 \end{aligned} \tag{119}$$

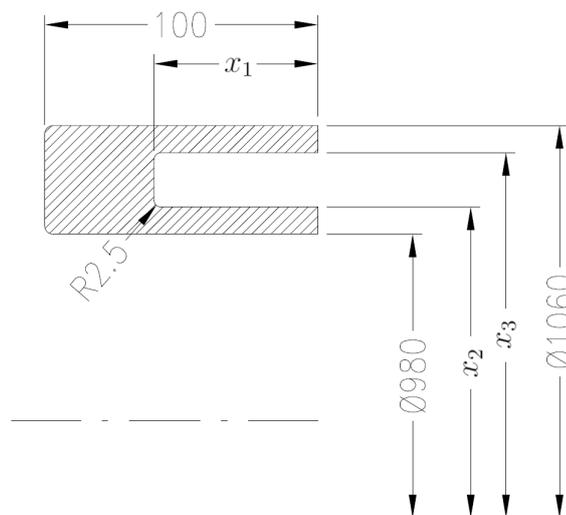
Onde ω representa a frequência natural do sistema a qual desejamos minimizar, $\bar{\omega}$ é a frequência alvo, $\dot{u}_{\bar{x}de}(\mathbf{x})$ é a velocidade de vibração na direção \bar{x} medida no mancal dianteiro da MEG, $\dot{u}_{\bar{x}nde}(\mathbf{x})$ é a velocidade de vibração na direção \bar{x} medida no mancal traseiro da MEG, $u_{\bar{y}est}(\mathbf{x})$ é a deflexão do estator na direção \bar{y} , σ_{de} é a tensão mecânica equivalente no mecanismo proposto localizado na parte dianteira da MEG, σ_{nde} é a tensão mecânica equivalente no mecanismo proposto localizado na parte traseira da MEG e x_1 , x_2 e x_3 são as variáveis de projeto. A frequência alvo $\bar{\omega}$ estabelecida para a função objetivo é de 107.5 [Hz]. Considerando um afastamento de 6% em relação ao valor encontrado na análise modal experimental e no modelo numérico computacional e 11% em relação a frequência de excitação de 120 [Hz].

Figura 47 – Modo de vibrar encontrado pelo modelo numérico computacional.



Fonte: Autor.

Figura 48 – Corte transversal do mecanismo proposto com a localização das variáveis de projeto x_1 , x_2 e x_3 .



Fonte: Autor.

9 RESULTADOS

Para aplicar a metodologia da Fig. 33 para o estudo de caso é necessário a criação de modelos substitutos para as funções da Eq. (119). Para tal, um HCL de 15 pontos de suporte iniciais é criado de forma independente para cada função do problema de otimização, ou seja, cada modelo substituto é iniciado com seu próprio HCL. Os modelos substitutos são criados por meio do método de Kriging, utilizando o parâmetro $p = 2$ e θ sendo definido através da técnica de otimização GA. Enquanto a aquisição de novos pontos de suporte é realizada através da função EI com o orçamento de 20 iterações, conforme demonstrado no Cap. 7. Para a definição da melhor geometria do mecanismo a técnica de otimização PSO é utilizada, sendo executada 5 rodadas com o máximo de iterações $max_{ite} = 100$, obtendo o melhor valor entre elas. Para a Eq. (17) é utilizado $c_1 = c_2 = 1.5$ e w como

$$w = \frac{(w_{min} - w_{max})i_{te}}{max_{ite}}, \quad (120)$$

onde w_{min} e w_{max} representam os valores máximo e mínimo de w , no caso 0.9 e 0.4 respectivamente. A Eq. (120) denota que w decai ao longo das iterações i_{te} , permitindo uma velocidade das partículas maior no início do processo de otimização, sendo reduzida a medida que se aproxima das últimas iterações.

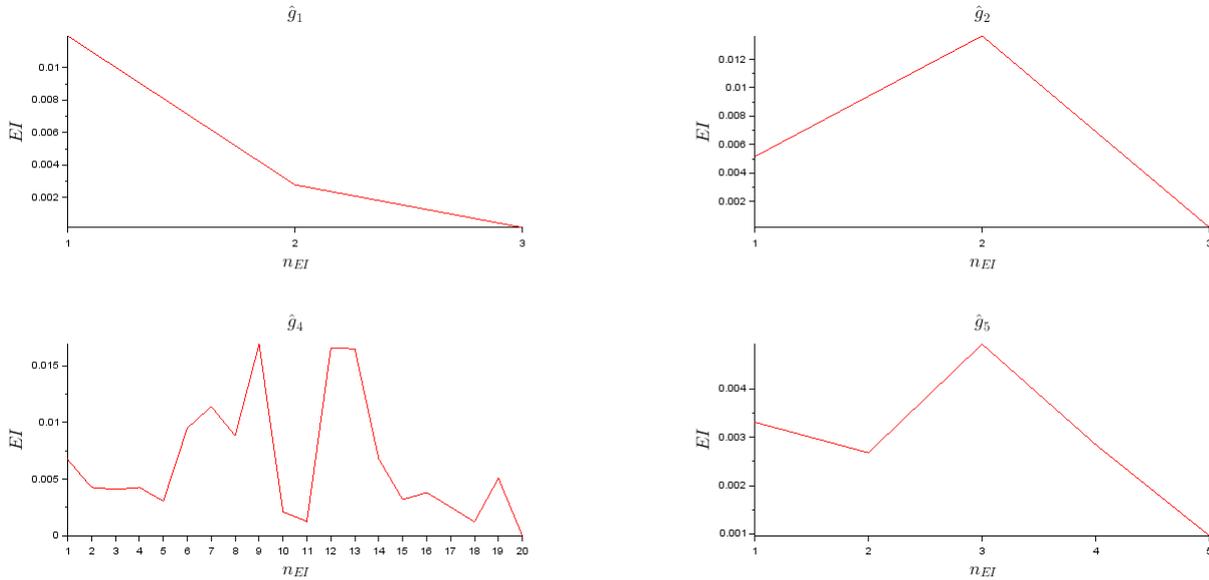
Após a execução da metodologia, a frequência natural encontrada é de 108,57 [Hz], para uma geometria ótima $\mathbf{x}^* = [70, 990, 1050]$. Comparando os valores de \mathbf{x}^* com a Tab. 8 observa-se que o ótimo se encontra nos limites laterais impostos, sendo o valor mínimo para x_2 e os valores máximos para x_1 e x_3 . Devido as restrições laterais conclui-se que a frequência natural alvo não é atendida. Como característica do problema de estudo de caso, as imagens das restrições da Eq. (119) são espaços viáveis, tornando o problema irrestrito.

A função de aquisição EI foi avaliada somente uma vez para a função objetivo, criando um ponto de suporte no ótimo da função \mathbf{x}^* . Esse comportamento mostra sua eficiência como otimizador para problemas de otimização irrestritos. Além disso, a função EI se mostrou como o esperado para todas as funções do problema de otimização com exceção da restrição \hat{g}_4 , como pode ser observado na Tab. 9 e na Fig. 49. Para uma melhor compreensão do comportamento do modelo substituto da função objetivo \hat{y} a Fig. 50 apresenta o resultado no plano $x_1 = 70$ [mm]. A Tab. 9 apresenta os valores das restrições em \mathbf{x}^* , onde n_{EI} é o número de pontos de suporte da função de aquisição EI. Os resultados encontrados pelo algoritmo mostram que o mesmo buscou a menor área transversal do mecanismo proposto, reduzindo sua rigidez.

Tabela 9 – Valor das funções que compõe o problema da Eq. (119) em \mathbf{x}^* e o número de pontos de suporte adquiridos pela função EI.

Função	Valor	n_{EI}
\hat{y}	108.57 [Hz]	1
\hat{g}_1	0.27 [mm/s]	2
\hat{g}_2	0.82 [mm/s]	2
\hat{g}_3	0.031 [mm]	0
\hat{g}_4	0.60 [MPa]	19
\hat{g}_5	0.28 [MPa]	4

Figura 49 – Comportamento da função de aquisição EI para $\hat{g}_1, \hat{g}_2, \hat{g}_4$ e \hat{g}_5 . As outras duas aproximações do problema de otimização foram omitidas pois a participação do EI no número de pontos de suporte é desprezível ou nulo.



Fonte: Autor.

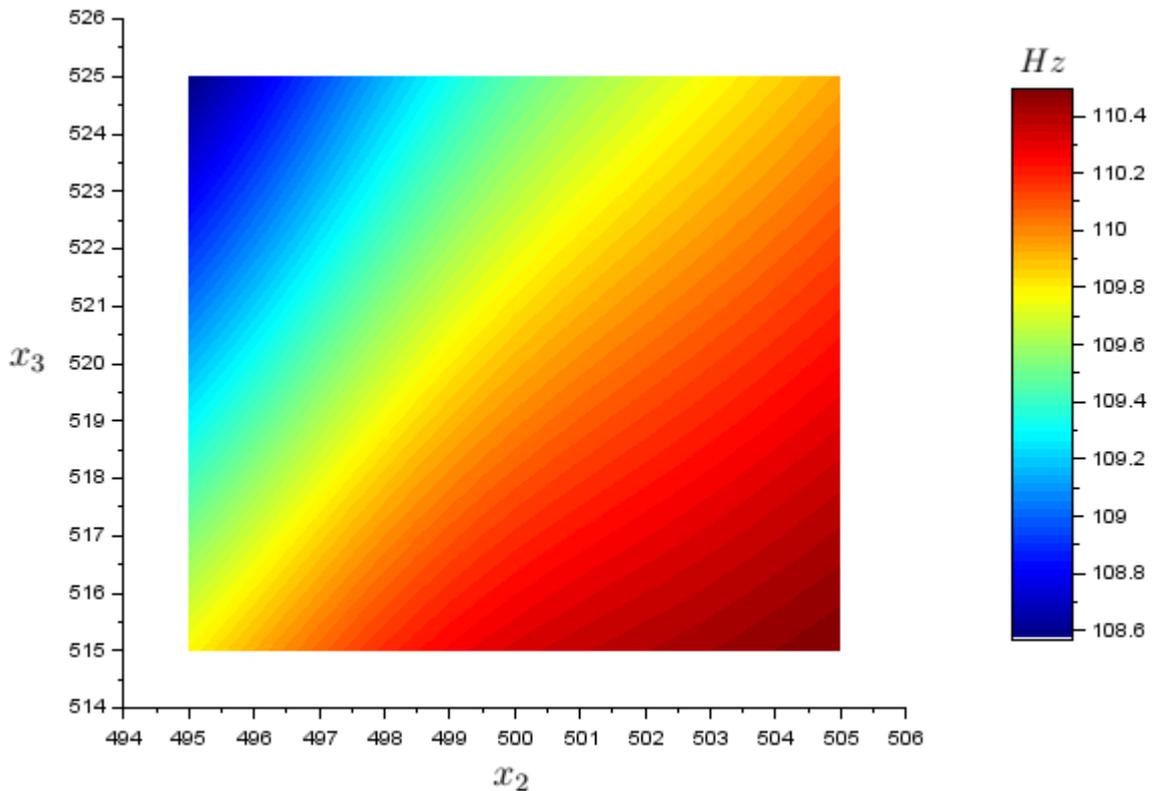
De forma intuitiva, foi verificado via *software* de elementos finitos o resultado para a maior seção transversal do mecanismo, ou seja, para $\mathbf{x} = [50, 1010, 1030]$. A frequência natural para esta condição é de 111.15 [Hz], aumentando a margem em relação a frequência de excitação de 120 [Hz]. Essa análise mostra que a alteração proposta fez reduzir a frequência natural do sistema independente dos valores das variáveis de projeto da geometria do mecanismo.

O aumento da margem fez com que houvesse redução na velocidade vibração dos mancais nas restrições \hat{g}_1 e \hat{g}_2 , obtendo níveis aceitáveis conforme apresentado nas normas do Cap. 2. Considerando que essas funções apresentam um certo comportamento linear o EI atuou somente 2 vezes em direção ao mínimo. Tanto a restrição \hat{g}_1 quanto a restrição \hat{g}_2 possuem o mínimo no valor de \mathbf{x}^* . Tomando como referência o plano de $x_1 = 70$ [mm], a Fig. 51 apresenta o comportamento das restrições \hat{g}_1 e \hat{g}_2 . Como pode ser observado, os valores de velocidades de vibração para \hat{g}_1 são menores que de \hat{g}_2 , isso é provocado pelo fato de o mancal traseiro estar em uma distância axial (direção \bar{z}) maior em relação ao centro do estator, consequentemente causando um momento maior.

A restrição \hat{g}_3 é uma verificação da deflexão do estator. Mesmo a aba inferior do mecanismo possuindo a menor espessura possível este valor foi atendido como esperado. Devido a linearidade dessa restrição o EI não atuou com novos pontos de suporte. A menor deflexão do estator, sendo o valor desejado, ocorre quando a dimensão x_1 é menor e a dimensão x_2 é maior. Mesmo \mathbf{x}^* apresentando valores opostos aos desejáveis, esta restrição foi atendida conforme Tab. 9. Para informação do valor desejado a Fig. 52 apresenta a restrição \hat{g}_3 para o plano $x_1 = 50$ [mm], onde o menor valor está no canto inferior direito, praticamente do lado oposto de \mathbf{x}^* .

Para o plano $x_1 = 70$ [mm], a tensão mecânica das restrições \hat{g}_4 e \hat{g}_5 apresentaram comportamento similar entre elas. Ambas possuem poucos pontos de correlação, dessa

Figura 50 – Comportamento do modelo substituto da função objetivo \hat{y} em função de x_2 e x_3 para $x_1 = 70$ [mm].



Fonte: Autor.

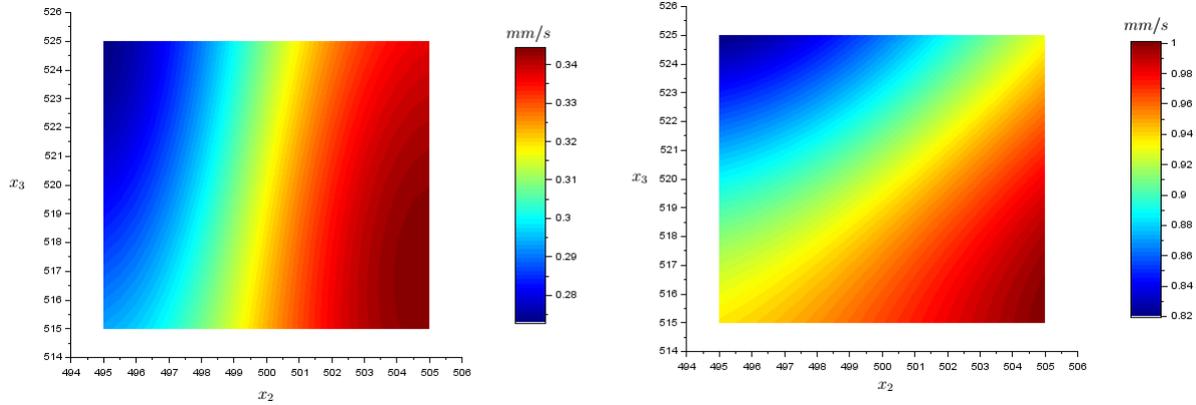
forma tendem aos seus respectivos valores esperados estimados $\hat{\mu}$. Onde $\hat{\mu}$ para g_4 é 0,467 [MPa] e para g_5 é 0,168 [MPa], a Fig. 53 apresenta seus respectivos valores para o plano $x_1 = 70$ [mm]. Os poucos pontos de suporte para a correlação no plano $x_1 = 70$ [mm] estão relacionados que seus ótimos estão distantes de \mathbf{x}^* . Portanto, mesmo o EI adquirindo um número alto de pontos de suporte devido a não-linearidade das funções, essa inserção ocorreu em regiões distantes do ótimo. Entre os pontos de suporte o menor valor de \hat{g}_4 é 0,308 [MPa], para $\mathbf{x} = [52, 1006, 1046]$, enquanto para \hat{g}_5 o menor valor entre os pontos de suporte é de 0,122 [MPa], para $\mathbf{x} = [57, 1004, 1032]$. Ou seja, os pontos de suporte dessas restrições estão distantes do ótimo global do sistema. Os baixos valores de tensão mecânica experimentados pelo mecanismo se deve a própria pressão magnética exercida no entreferro da MEG ser baixa, de acordo com a Eq. (3) é de 1.22 [MPa].

Para verificar o sentido físico dos modelos substitutos, propõe-se modelar o mecanismo conforme Fig. 54, onde $\sigma_r(\nu, t)$ representa a pressão magnética radial oriunda do entreferro conforme Eq. (5). Ainda pode-se reduzir o modelo considerando que a aba inferior do mecanismo seja representada por uma viga engastada conforme mostrado na Fig. 55.

Segundo Rao (2009), a rigidez de uma viga engastada com um grau de liberdade conforme apresentado na Fig. 54 é denotada por

$$k_v = \frac{3E_m I v}{l_v^3}, \quad (121)$$

Figura 51 – Comportamento dos modelos substitutos das funções de restrição \hat{g}_1 (à esquerda) e \hat{g}_2 (à direita), ambas em função de x_2 e x_3 para $x_1 = 70$ [mm].



Fonte: Autor.

onde E_m é o módulo de elasticidade do material em [Pa], I_v é o momento de inércia de área da seção da viga em [m⁴] e l_v é o comprimento da viga em [m], sendo a variável de projeto x_1 . Para o mecanismo proposto a inércia I_v pode ser representada por

$$I_v = \frac{b_v(x_2 - 980)^3}{12}, \quad (122)$$

onde b_v é a largura da viga, a qual é constante. Assim a Eq. (121) torna-se

$$k_v = \frac{3E_m(x_2 - 980)^3}{12x_1^3}, \quad (123)$$

sendo o primeiro termo uma constante. Enquanto a frequência natural de uma viga engastada é calculada por

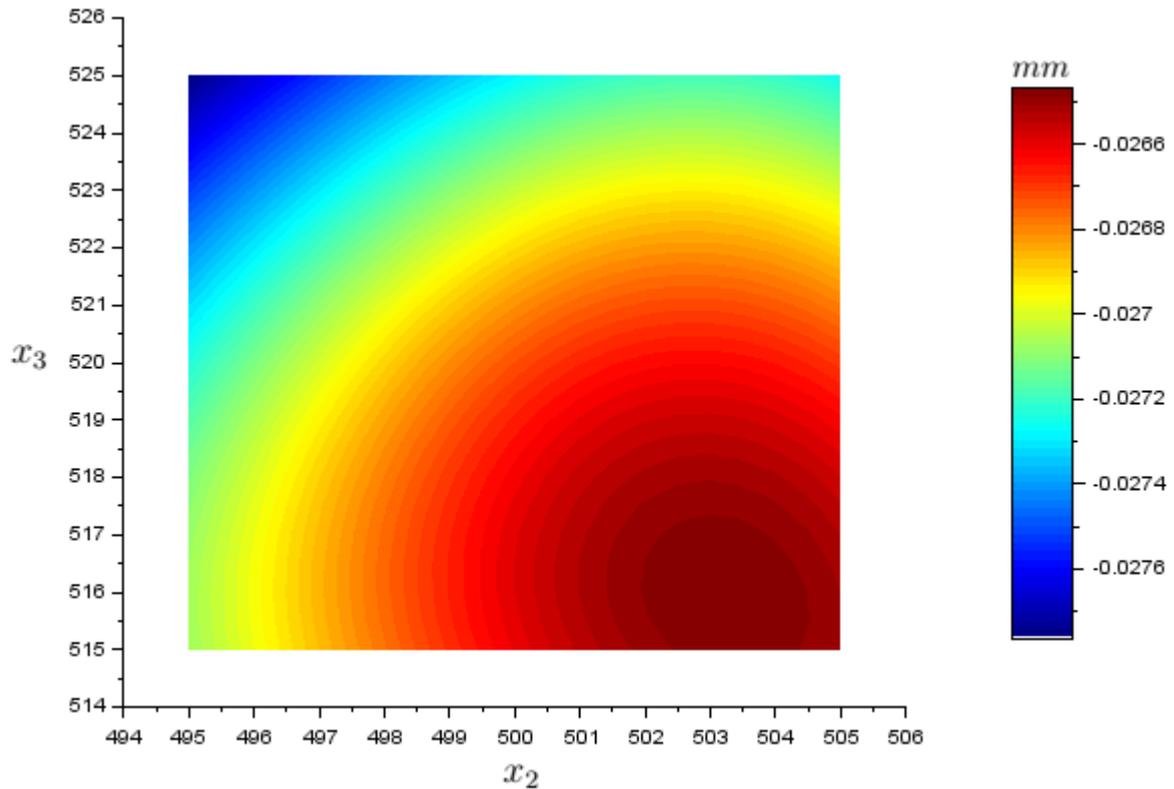
$$\omega_v = \sqrt{\frac{k_v}{m_v}}, \quad (124)$$

onde m_v é a massa da viga, considerada desprezível em relação a toda MEG. Dessa forma, pode-se considerar que a frequência natural do sistema é proporcional a x_1 e x_2 , ou seja,

$$\omega_v \propto \frac{(x_2 - 980)^3}{x_1^3}. \quad (125)$$

A Eq. (125) mostra que a frequência natural do sistema denotada pela função objetivo \hat{g} tende ao mínimo quando x_1 tende ao seu valor máximo e x_2 tende ao seu valor mínimo. Ressalta-se que a Eq. (125) está relacionada a alteração da rigidez entre $k_{s\bar{x}}$ e $k_{f\bar{x}}$, bem como $k_{s\bar{y}}$ e $k_{f\bar{y}}$ apresentadas na Fig. 39, por isso a proporcionalidade com a frequência natural do sistema.

Figura 52 – Comportamento do modelo substituto da função de restrição \hat{g}_3 em função de x_2 e x_3 para $x_1 = 50$ [mm].



Fonte: Autor.

Enquanto a deflexão na restrição \hat{g}_3 é definida como

$$\delta_v = \frac{F_r x_1^3}{3E_m I_v}, \quad (126)$$

substituindo a Eq. (122) na Eq. (126), a deflexão δ_v torna-se

$$\delta_v = \frac{F_r x_1^3 12}{3E_m b_v (x_2 - 980)^3}. \quad (127)$$

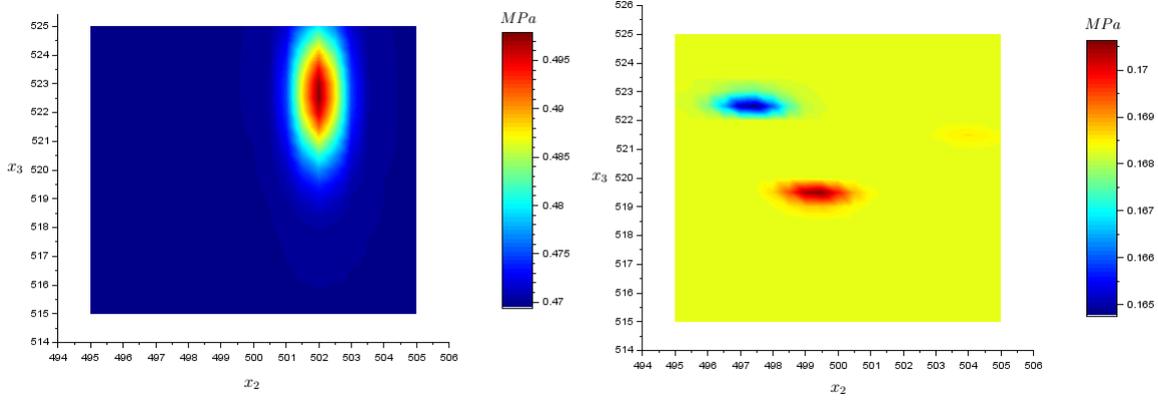
A Eq. (127) demonstra a influência na redução da restrição \hat{g}_3 com o aumento da variável x_2 . Ou seja, a redução da espessura da aba do mecanismo gera uma deflexão maior para o estator, alterando o entreferro da MEG, mas ainda dentro do valor limite estabelecido.

As tensões mecânicas no mecanismo proposto ocorreram no raio interno da aba inferior, conforme Fig. 56. Para a explicação física é possível representar a tensão mecânica envolvida em uma viga engastada como da Fig. 55 por

$$\sigma_v = \frac{F_r x_1}{I_v}. \quad (128)$$

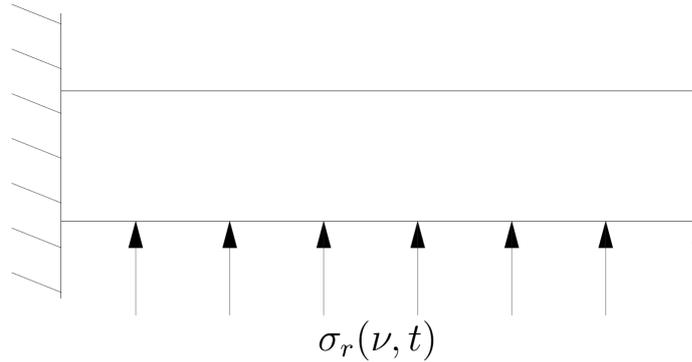
Dessa forma, observa-se que o aumento do comprimento, expresso por x_1 , aumenta a tensão na extremidade engastada da viga. Sendo que esta tensão ocorre de forma alternada nos

Figura 53 – Comportamento do modelo substituto das funções de restrição \hat{g}_4 (à esquerda) \hat{g}_5 (à direita), ambas em função de x_2 e x_3 para $x_1 = 70$ [mm].



Fonte: Autor.

Figura 54 – Modelo da viga com a pressão magnética $\sigma_r(\nu, t)$ distribuída na aba do mecanismo proposto.



Fonte: Autor.

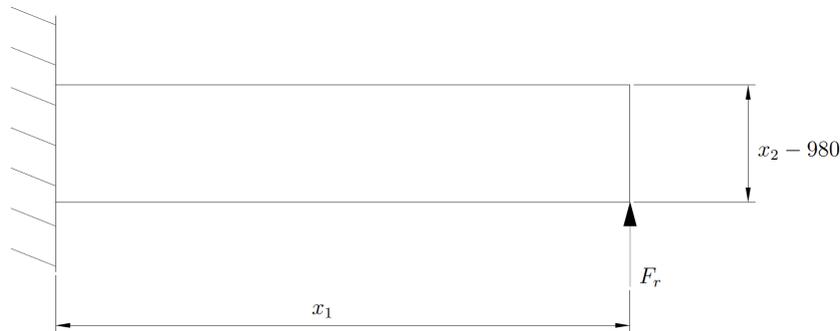
polos da MEG, já que a Eq. (5) está em função de $2p_p\nu$. Assim pode-se afirmar que a tensão mecânica experimentada pelo mecanismo é

$$\sigma_v \propto \frac{12F_r x_1}{(x_2 - 980)^3}. \quad (129)$$

Por fim, uma análise harmônica computacional é executada para a faixa de operação de 0 à 200 [Hz]. Nesse caso, uma excitação com a pressão magnética radial, Eq. 5, é aplicada a uma frequência do dobro da rede no diâmetro interno do estator. O objetivo é realizar um comparativo do comportamento de todo o sistema entre a solução atual e a solução com o mecanismo proposto para as direções \bar{x} e \bar{y} em ambos os mancais. De modo geral, com a implementação do mecanismo houve uma redução da velocidade de vibração em toda a faixa operacional da MEG.

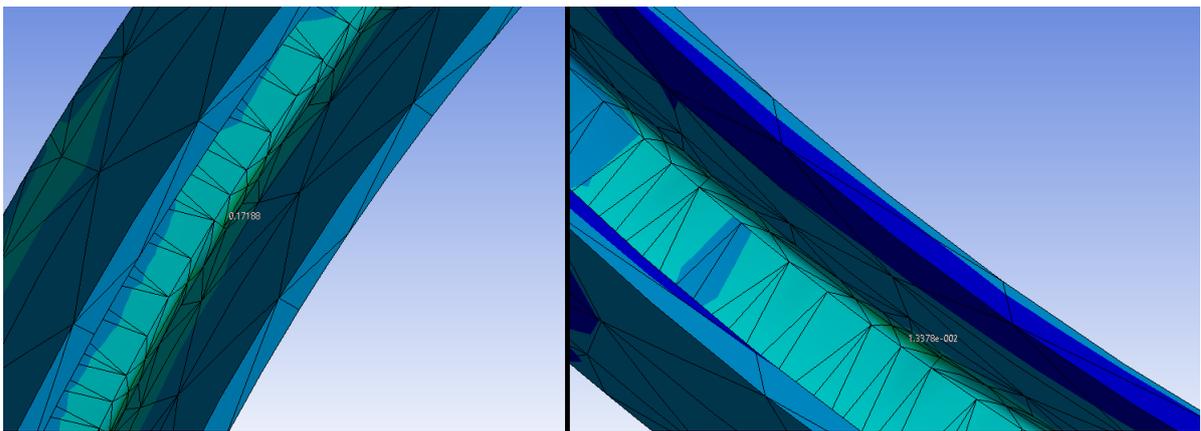
Analisando a Fig. 57, verifica-se na solução atual um pico próximo a 60 [Hz] com valor acima do permitido por norma, o qual é danoso para a operação da máquina pois pode ser excitado pela frequência de alimentação da rede. Na região de 120 [Hz] há picos,

Figura 55 – Modelo da viga com a pressão magnética $\sigma_r(\nu, t)$ concentrada na aba do mecanismo proposto.



Fonte: Autor.

Figura 56 – Tensão mecânica no mecanismo do lado acionado (à esquerda) e do lado não-acionado (à direita).



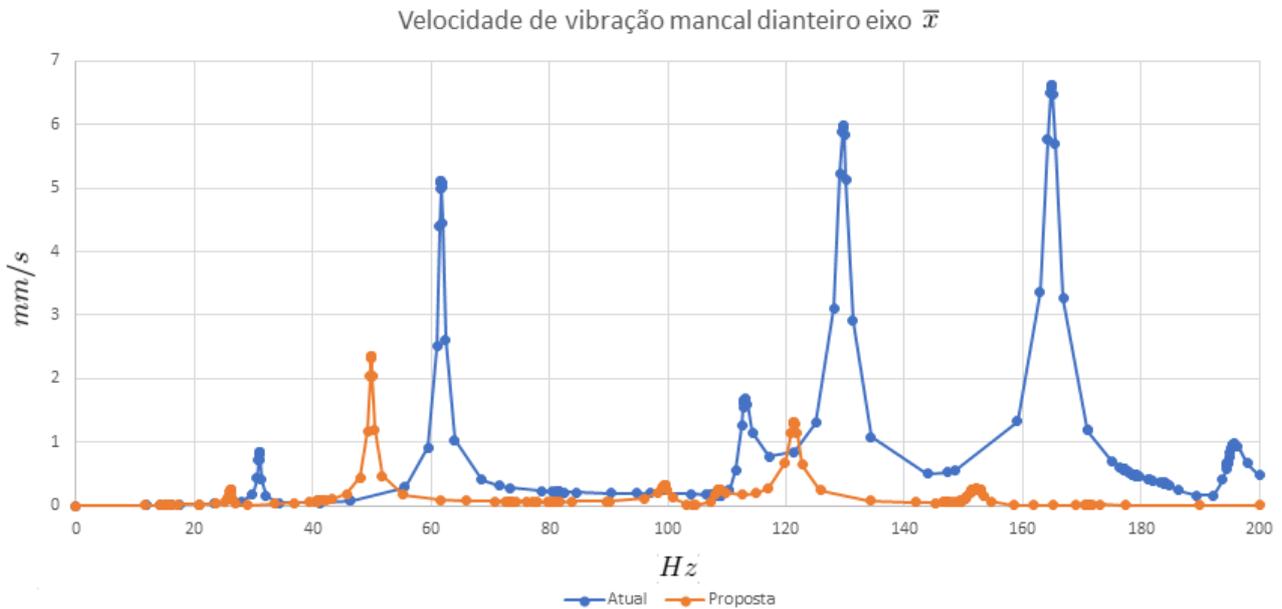
Fonte: Autor.

porém o maior deles está acima desse valor, não sendo um problema para a operação. Para este caso, frequências naturais de ordem superiores não são um problema durante a operação da MEG. Além da redução da velocidade de vibração, o mecanismo fez com que as frequências naturais do sistema tivessem sua frequência reduzida. Com isso a frequência natural encontrada em aproximadamente 130 [Hz] na solução proposta foi transladada para próximo de 120 [Hz]. Mas quando excitada sua velocidade de vibração é menor do que o estabelecido por norma, mesmo podendo ser excitado em duas vezes a frequência da rede de alimentação.

Conforme apresentado na Fig. 58, a velocidade de vibração do mancal dianteiro na direção \bar{y} é maior do que estabelecido por norma na região de 108 [Hz], não se traduzindo em problema para a operação. Mesmo assim, a implementação do mecanismo trouxe uma redução drástica dos valores de velocidade de vibração nessa região, bem como em na região de 120 [Hz]. Também, sua implementação trouxe um pico relevante na região de 80 [Hz], a qual não é danoso para operação da máquina.

O mancal traseiro é o que possui maior velocidade de vibração na região de 120 [Hz], possuindo um modo de vibrar torcional em 114 [Hz], o qual majora o valor na

Figura 57 – Análise harmônica computacional com a velocidade de vibração no mancal dianteiro na direção \bar{x} .

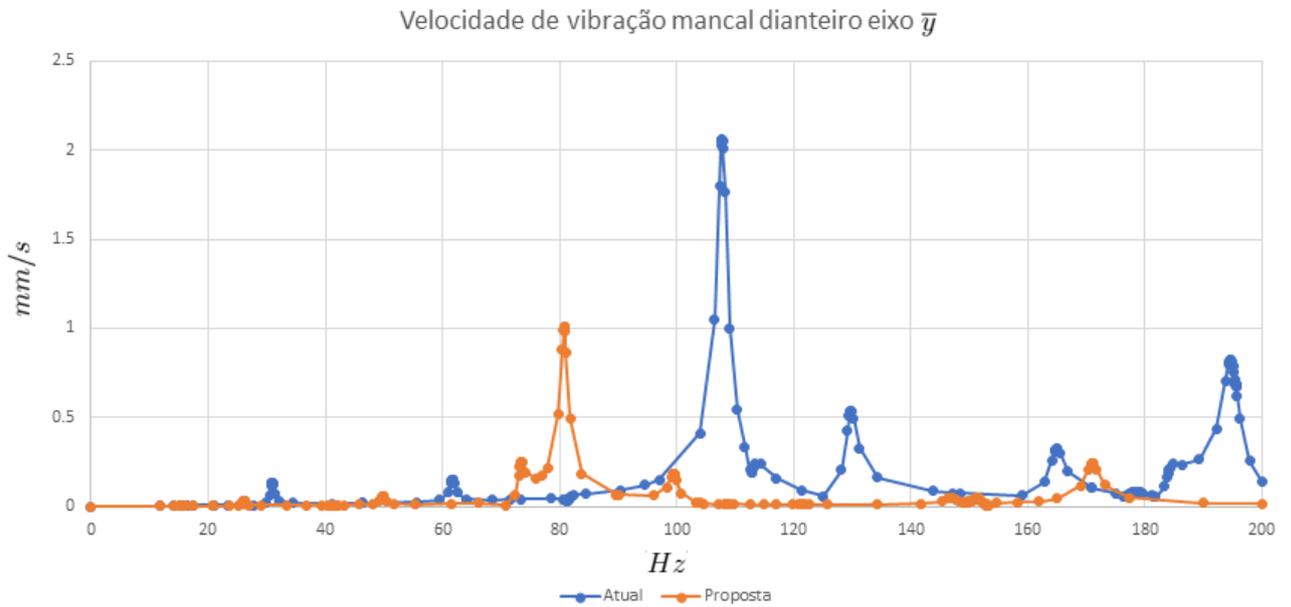


Fonte: Autor.

direção \bar{x} , como mostrado na Fig. 59. Esse fenômeno também pode ser observado para valores na região de 60 [Hz], tendo o maior pico de velocidade de vibração de todos os espectros. A implementação do mecanismo trouxe reduções significativas na região de duas vezes a frequência de operação, eliminando a vibração danosa do sistema atual. A implementação do mecanismo gerou uma frequência natural em 50 [Hz], proveniente da translação da frequência natural em 60 [Hz] da solução atual, por sua vez possui valor maior que o estabelecido por norma. Portanto, por mais que a máquina do estudo de caso é projetada para uma frequência de alimentação de 60 [Hz], não é recomendável aplicar a frequência de 50 [Hz], a qual é padronizada em muitos países.

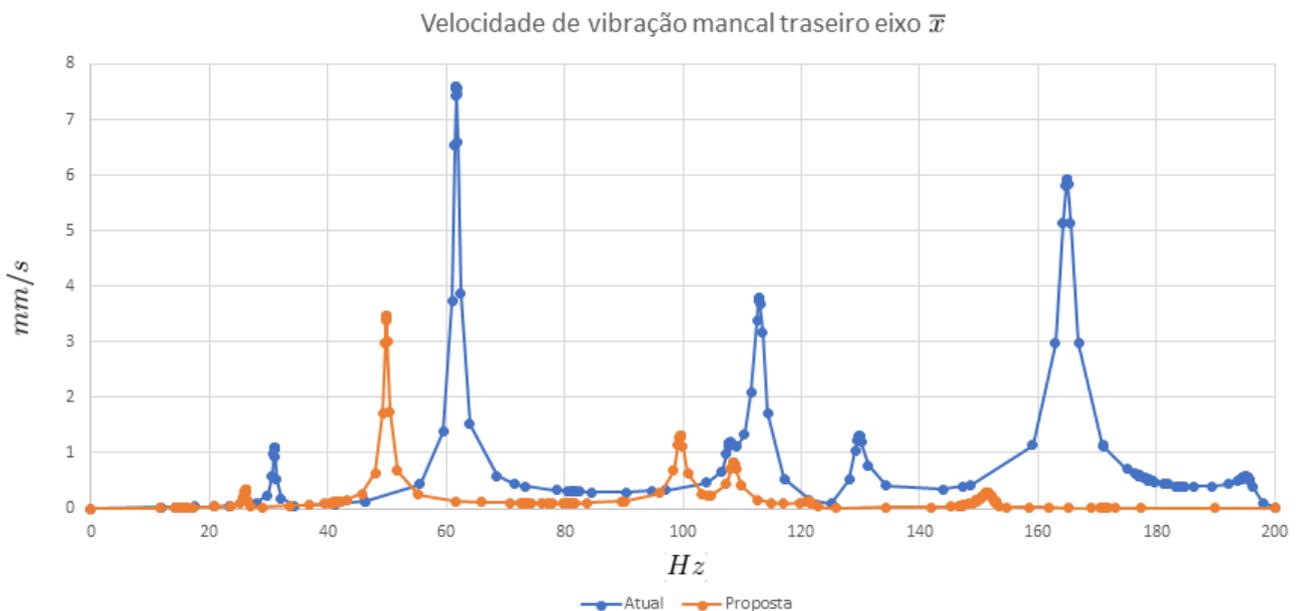
A mesma analogia pode-se fazer entre os mancais quando se trata da direção \bar{y} , pois o comportamento dos modos de vibrar são os mesmos em toda a faixa analisada. Como apresentado na Fig. 60 o mecanismo proposto tem sua melhor performance em termos de velocidade de vibração na região de 120 [Hz], traduzindo em uma redução significativa desses valores.

Figura 58 – Análise harmônica computacional com a velocidade de vibração no mancal dianteiro na direção \bar{y} .



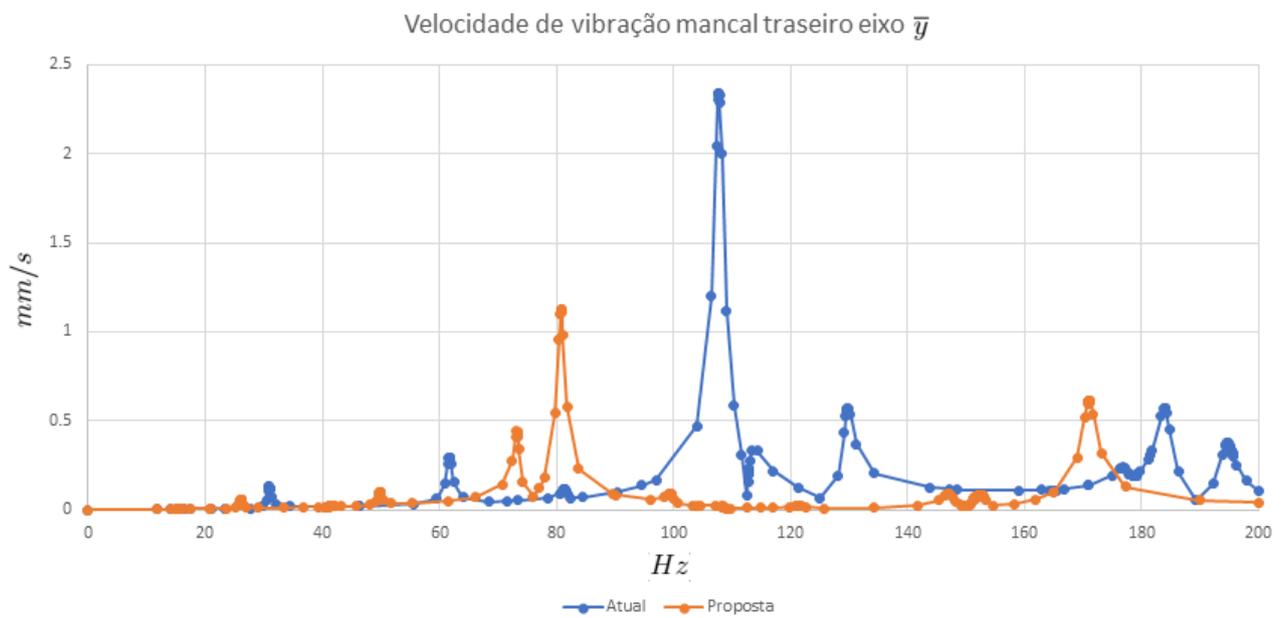
Fonte: Autor.

Figura 59 – Análise harmônica computacional com a velocidade de vibração no mancal traseiro na direção \bar{x} .



Fonte: Autor.

Figura 60 – Análise harmônica computacional com a velocidade de vibração no mancal traseiro na direção \bar{y} .



Fonte: Autor.

10 CONCLUSÃO

Esse trabalho discute uma metodologia de otimização paramétrica utilizando modelos substitutos. Essa metodologia consiste em gerar um problema de otimização através de modelos substitutos de forma independente para as funções objetivo e de restrições, o qual pode ser solucionado através de uma técnica populacional. A metodologia é aplicada em um problema de teste conhecido na literatura, analisando suas vantagens e desvantagens. Posteriormente, a mesma é aplicada no projeto de um mecanismo para a redução de vibração em uma MEG.

A teoria de Kriging é utilizada para definir os modelos substitutos na metodologia discutida nesse trabalho. Esse tipo de modelo substituto mostra uma boa aproximação para funções *black-box* de alto custo computacional. Como exemplo, no estudo de caso, as realizações desse tipo de função são feitas através de um *software* comercial de elementos finitos, onde seu custo é de aproximadamente de 1 hora por ponto de suporte. Dessa forma, os modelos substitutos se mostraram vantajosos para acelerar o processo de otimização, permitindo modelos de rápida e fácil exploração. Porém, alguns cuidados devem ser tomados ao definir os parâmetros iniciais desses modelos substitutos, como os limites laterais do parâmetro θ e o parâmetro p . Na metodologia discutida θ é definido por meio de um Algoritmo Genético e o valor de p é constante, sendo baseado na literatura. Uma dificuldade adicional da metodologia, é a computação da inversa da matriz de correlação Ψ por meio da decomposição de Cholesky, pois em um espaço muito denso de realizações essa matriz pode se tornar singular.

A inicialização dos modelos substitutos ocorre através de um hipercubo latino otimizado, o qual cria pontos de suporte iniciais para a realização de alto custo. Para otimizar o hipercubo latino é utilizada a técnica *maxmin*, sendo uma forma de melhorar a sua distribuição no domínio. Mesmo assim, há um certo custo computacional na criação do HCL otimizado, pois essa técnica exige que se faça 7 avaliações para o parâmetro q . Mesmo o HCL tendo uma boa distribuição de pontos de suporte iniciais, a sua quantidade é definida pelo usuário, e isso pode fazer com que o excesso de pontos leve a um alto custo computacional inicial. Funções desconhecidas são bem aproximadas utilizando um HCL otimizado com 15 pontos de suporte iniciais, conforme resultados observados no problema de teste e apresentado no Cap. 7.

O problema de teste também é utilizado para avaliar um comparativo das funções de aquisições EI, PI, LCB e s^2 com o intuito de conhecer qual se adequa melhor a metodologia discutida. O EI apresentou vantagens em relação ao PI e LCB, necessitando de um número menor de pontos adquiridos e um menor erro entre os valores das variáveis de projeto, função objetivo e restrições em relação aos valores de referência. O PI e o LCB não demonstraram as mesmas vantagens do EI, porém se mostram promissores mesmo dependendo do usuário para definição de constantes, como \mathcal{V} no LCB e o valor alvo para o PI. Em problemas de otimização restritos, a função de aquisição PI pode ser utilizada juntamente com o EI para determinar o ótimo global \mathbf{x}^* (FORRESTER; SOBESTER; KEANE, 2008). Essa abordagem reduz o número de pontos de suporte criados posteriormente ao HCL, e conseqüentemente o número de execuções da função de alto custo computacional. A aquisição de pontos de suporte por s^2 não é vantajoso, atuando nos limites laterais das funções devido sua maior variância. Isso faz com que o número de pontos de suporte seja demasiadamente grande, trazendo um maior custo computacional e não obtendo a mesma acurácia em relação as outras funções de aquisição. Além disso, pode tornar o domínio denso o suficiente a ponto da matriz Ψ ser singular.

Devido ao fato de não termos expressões para as derivadas das funções a serem aproximadas, a solução do problema de otimização deve ser realizada por método de ordem zero. O método PSO é um conveniente otimizador para a metodologia apresentada, sendo dependente dos parâmetros w , c_1 e c_2 definidos pelo usuário. A definição desses parâmetros pode tornar o deslocamento das partículas muito grande, deixando de explorar a região da fronteira entre o espaço viável e inviável. Para evitar isso, o parâmetro w (inércia das partículas) decresce linearmente com o aumento das iterações do PSO. O gerenciamento das restrições através de Deb (2000) é de fácil implementação e eficaz na execução do algoritmo.

Visando a definição do ótimo \mathbf{x}^* , quando a geração de modelos substitutos ocorre com poucos pontos de suporte a metodologia apresenta desvantagens. Isso pode ser observado nas análises dos resultados para o cenário C do problema de teste. Essa definição pode prejudicar a solução do problema, mesmo que reduza o custo computacional da avaliação de pontos de suporte. Essa desvantagem pode ser agravada quando o problema possui restrições ativas, sendo que o ótimo pode estar em uma região longe de pontos de suporte. Destaca-se que mesmo buscando o mínimo das funções objetivo e de restrições através de uma função de aquisição de forma independente, em sua maioria, problemas de otimização com restrições não possuem seu mínimo no mesmo local das funções que o compõe. Isso mostra a importância da distribuição inicial do HCL, determinando pontos de suporte que abrangem todo o domínio de cada função e não somente seu mínimo. Mesmo que essa forma aumente o custo computacional ao inicializar os modelos substitutos.

Em relação ao problema do estudo de caso, sua solução está na fronteira da região viável definida pelas restrições laterais. Isso ocorreu devido as particularidades das funções objetivo e de restrições. Entretanto a metodologia apresentada se mostrou efetiva também para esse tipo de problema. Do ponto de vista do objetivo proposto, o mesmo foi atendido, pois o mecanismo otimizado se mostrou eficiente, atendendo as restrições por toda a faixa de operação. No caso, gerando uma maior margem de separação da frequência natural do sistema em relação a frequência de excitação. Com uma margem maior, reduziu-se os níveis de velocidade de vibração nos mancais, como pode ser visto nos resultados das restrições \hat{g}_1 e \hat{g}_2 . Além disso, as baixas tensões mecânicas experimentadas pelo mecanismo permitem sua fabricação em um aço comercial de baixo custo e fácil usinabilidade.

Mesmo tendo confiança nos modelos substitutos e na solução do problema de otimização, a análise da física do problema deve ser realizada. Assim, o projetista deve procurar compreender se a solução definida pela metodologia está coerente com a física do problema. No estudo de caso apresentado, uma avaliação com equacionamento físico é realizada para o entendimento de cada modelo substituto que compõe o problema de otimização, descrevendo o impacto de cada variável de projeto na física do problema.

Os resultados do problema de teste e do estudo de caso, mostram que a metodologia apresentada é robusta para a solução de problemas de otimização com restrição. Podendo ser aplicada nas mais diversas disciplinas da Engenharia, pois é dependente somente dos modelos substitutos. Auxiliando na tomada de decisão durante a fase de projeto do produto, reduzindo o custo computacional quando *softwares* de elementos finitos são utilizados. Também é possível correlacionar variáveis de problemas multidisciplinares, tornando análises complexas em problemas de otimização de fácil e rápida exploração. Por fim, o mecanismo proposto para o estudo de caso se mostrou eficiente, reduzindo a velocidade de vibração do sistema e atendendo os valores estabelecidos por norma.

10.1 TRABALHOS FUTUROS

Pode-se dividir as sugestões para trabalhos futuros em duas partes. Inicialmente abordando melhorias e sugestões para a metodologia utilizada. Já em uma segunda parte, pode-se citar trabalhos futuros relacionados ao mecanismo proposto e suas questões físicas quando aplicado a uma MEG. Como sugestões para a metodologia apresentada pode-se apontar:

- Aplicar uma metodologia que permita a redução do número de análises da função de alta fidelidade;
- Utilizar outras técnicas de otimização como NES (*Natural Evolution Strategies*);
- Aplicar em um problema com maior número de variáveis de projeto e com maior quantidade de restrições e analisar seu comportamento.

Em relação ao projeto do mecanismo proposto, abaixo seguem algumas sugestões para trabalhos futuros:

- Adicionar a análise de fadiga como restrição, avaliando a tensão mecânica devido a pressão magnética ter comportamento alternado;
- Analisar o comportamento modal da MEG quando há um aumento da pressão magnética, permitindo obter uma maior potência para a mesma geometria elétrica;
- Definir novas geometrias ou materiais, facilitando ainda mais sua fabricação;
- Realizar sua fabricação e a implementação em uma MEG para analisar os resultados e seu comportamento modal.

REFERÊNCIAS

- 1 ANDERSON, Theodore W. **An introduction to multivariate statistical analysis**. [S.l.], 1962. Citado nas pp. 44, 45, 55.
- 2 ANDREWS, Larry C. **Special functions of mathematics for engineers**. [S.l.]: Spie Press, 1998. v. 49. Citado na p. 65.
- 3 API. **API 541 - Form-wound squirrel cage induction motor - 375 kw (500 Horsepower) and larger**. [S.l.: s.n.], 2014. 5 ed. 160p. Citado nas pp. 19, 30–32.
- 4 ARORA, Jasbir. **Introduction to optimum design**. [S.l.]: Elsevier, 2004. Citado nas pp. 33, 35, 70, 72, 121.
- 5 BACCARINI, Lane Maria R. Detecção e diagnóstico de falhas em motores de indução. UFMG, 2005. Citado na p. 26.
- 6 BAUDIN, Michaël; COUVERT, Vincent; STEER, Serge. Optimization in scilab. **Scilab Consortium, INRIA Paris-Rocquencourt**, 2010. Citado na p. 72.
- 7 BOROVKOV, Alexandr A. **Probability theory**. [S.l.]: Springer, 2013. Citado nas pp. 41–43.
- 8 CHAPMAN, Stephen. **Electric machinery fundamentals**. [S.l.]: Tata McGraw-Hill Education, 2005. Citado nas pp. 23, 27.
- 9 COELLO, Carlos A. Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. **Computer methods in applied mechanics and engineering**, Elsevier, v. 191, n. 11-12, p. 1245–1287, 2002. Citado nas pp. 38, 39.
- 10 COSTELLO, Michael J. Understanding the vibration forces in induction motors, p. 67–76, 1990. Citado na p. 26.
- 11 COX, Dennis D.; JOHN, Susan. A statistical method for global optimization. In: IEEE. [PROCEEDINGS] 1992 IEEE International Conference on Systems, Man, and Cybernetics. [S.l.: s.n.], 1992. p. 1241–1246. Citado na p. 65.
- 12 CRESSIE, Noel. **Statistics for spatial data**. [S.l.]: John Wiley & Sons, 2015. Citado nas pp. 47, 51.
- 13 DEB, Kalyanmoy. An efficient constraint handling method for genetic algorithms. **Computer methods in applied mechanics and engineering**, Elsevier, v. 186, n. 2-4, p. 311–338, 2000. Citado nas pp. 39, 40, 72, 98.
- 14 EBERHART, Russell; KENNEDY, James. Particle swarm optimization. In: CITESEER. PROCEEDINGS of the IEEE international conference on neural networks. [S.l.: s.n.], 1995. v. 4, p. 1942–1948. Citado na p. 37.
- 15 ENDERS, Craig K. **Applied missing data analysis**. [S.l.]: Guilford press, 2010. Citado na p. 48.

- 16 FINLEY, William R.; HODOWANEC, Mark M.; HOLTER, Warren G. An analytical approach to solving motor vibration problems, p. 217–232, 1999. Citado nas pp. 26, 27.
- 17 FISHER, Ronald A. On an absolute criterion for fitting frequency curves. **Messenger of Mathematics**, v. 41, p. 155–156, 1912. Citado na p. 48.
- 18 FITZGERALD, Arthur E.; KINGSLEY, Charles; UMANS, Stephen D. **Electric machinery**. [S.l.]: McGraw-Hill New York, 2003. v. 5. Citado na p. 23.
- 19 FONÇATTI, Lucas G. Estudo numérico sobre controle passivo de vibrações em máquinas elétricas rotativas. Universidade Federal do Paraná - UFPR, 2018. Citado nas pp. 79, 82.
- 20 FORRESTER, Alexander; SOBESTER, Andras; KEANE, Andy. **Engineering design via surrogate modelling: a practical guide**. [S.l.]: John Wiley & Sons, 2008. Citado nas pp. 62–65, 97.
- 21 FU, Zhi-Fang; HE, Jimin. **Modal analysis**. [S.l.]: Butterworth-Heinemann, 2001. Citado na p. 80.
- 22 GONÇALVES, Vinicius S. Desenvolvimento de uma metodologia numérica para a predição dos três primeiros modos de vibração de um motor elétrico fixo em base rígida. Florianópolis, 2012. Citado na p. 79.
- 23 HARRIS C. M.; PIERSOL, A. G. **Shock and Vibration Handbook**. [S.l.]: McGraw-Hill, 2002. Citado na p. 79.
- 24 HOLLAND, John H. Adaptation in natural and artificial systems. **Ann Arbor**, 1975. Citado na p. 35.
- 25 HOLTER, Warren G. Vibration reduction for electric motors. US Patent, 1993. Citado na p. 80.
- 26 HWANG, Donghyun *et al.* Vibration transmissibility reduction module with flexure mechanism for personal tools. **Journal of mechanical science and technology**, Springer, v. 24, n. 1, p. 223–226, 2010. Citado nas pp. 80, 81.
- 27 IEC. **IEC 60034-14 - Rotating electrical machines - Part 14: Mechanical vibration of certain machines with shaft heights 56 mm and higher - Measurement, evaluation and limits of vibration severity**. [S.l.: s.n.], 2003. 3 ed. 29p. Citado nas pp. 19, 30, 31.
- 28 JOHNSON, Mark E.; MOORE, Leslie M.; YLVISAKER, Donald. Minimax and maximin distance designs. **Journal of statistical planning and inference**, Elsevier, v. 26, n. 2, p. 131–148, 1990. Citado na p. 62.
- 29 JONES, Donald R. A taxonomy of global optimization methods based on response surfaces. **Journal of global optimization**, Springer, v. 21, n. 4, p. 345–383, 2001. Citado nas pp. 60, 65, 66.
- 30 KARUSH, William. Minima of functions of several variables with inequalities as side constraints. **M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago**, 1939. Citado na p. 33.

- 31 KRAUSE, Paul *et al.* **Analysis of electric machinery and drive systems**. [S.l.]: John Wiley & Sons, 2013. v. 75. Citado na p. 23.
- 32 KRIGE, Daniel G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Southern African Institute of Mining and Metallurgy**, Southern African Institute of Mining e Metallurgy, v. 52, n. 6, p. 119–139, 1951. Citado na p. 57.
- 33 KUHN, Harold W.; TUCKER, Albert W. Nonlinear programming. In: PROCEEDINGS of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950. Berkeley e Los Angeles: University of California Press, 1951. p. 481–492. Citado na p. 33.
- 34 KUSHNER, Harold J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise, 1964. Citado na p. 65.
- 35 LAFFOON, C. M.; ROSE, B. A. Special problems of two-pole turbine generators. **Transactions of the American Institute of Electrical Engineers, IEEE**, v. 59, n. 1, p. 30–34, 1940. Citado na p. 79.
- 36 MACKAY, David J. C. **Information theory, inference and learning algorithms**. [S.l.]: Cambridge university press, 2003. Citado na p. 53.
- 37 MATHERON, Georges. Principles of geostatistics. **Economic geology**, Society of Economic Geologists, v. 58, n. 8, p. 1246–1266, 1963. Citado na p. 57.
- 38 MCKAY, Michael D.; BECKMAN, Richard J.; CONOVER, William J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. **Technometrics**, Taylor & Francis, v. 21, n. 2, p. 239–245, 1979. Citado na p. 62.
- 39 MEZURA-MONTES, Efrén; COELLO, Carlos A. Constraint-handling in nature-inspired numerical optimization: past, present and future. **Swarm and Evolutionary Computation**, Elsevier, v. 1, n. 4, p. 173–194, 2011. Citado na p. 39.
- 40 MOCKUS, Jonas; TIESIS, Vytautas; ZILINSKAS, Antanas. The application of Bayesian methods for seeking the extremum. **Towards global optimization**, v. 2, n. 117-129, p. 2, 1978. Citado na p. 66.
- 41 MORRIS, Max D.; MITCHELL, Toby J. Exploratory designs for computational experiments. **Journal of statistical planning and inference**, Elsevier, v. 43, n. 3, p. 381–402, 1995. Citado nas pp. 62–64.
- 42 NAU, Sebastião L. Análise da influência da excentricidade do rotor sobre o ruído acústico de origem magnética gerado por motores de indução trifásicos. UFSC, SC, 2007. Citado na p. 27.
- 43 NEMA. **NEMA MG-1 - Motors and generators**. [S.l.: s.n.], 2016. 711p. Citado nas pp. 19, 30, 31.

- 44 PENNIMAN, A. L.; TAYLOR, H. D. Suppression of magnetic vibration and noise of two-pole turbine generators. **Electrical Engineering**, IEEE, v. 60, n. 6, p. 283–288, 1941. Citado na p. 79.
- 45 PYRHONEN, Juha; JOKINEN, Tapani; HRABOVCOVA, Valeria. **Design of rotating electrical machines**. [S.l.]: John Wiley & Sons, 2013. Citado na p. 24.
- 46 RADLE, Patrick J. Power tool with vibration isolated handle. US Patent, 1997. Citado na p. 80.
- 47 RAO, Singiresu. **Vibrações Mecânicas. Quarta**. [S.l.]: São Paulo: Pearson Prentice Hall, 2009. Citado na p. 89.
- 48 RAO, Singiresu. **Engineering Optimization : Theory and Practice**. [S.l.]: New Age International Pvt. Ltd., 2010. Citado na p. 35.
- 49 RICHARDSON, Mark H. Is it a mode shape, or an operating deflection shape? **Sound and Vibration**, [Bay Village, Ohio, Acoustical Publications, inc.], v. 31, n. 1, p. 54–67, 1997. Citado na p. 81.
- 50 SACKS, Jerome *et al.* Design and analysis of computer experiments. **Statistical science**, JSTOR, p. 409–423, 1989. Citado na p. 65.
- 51 SIMPSON, Timothy W. *et al.* Kriging models for global approximation in simulation-based multidisciplinary design optimization. **AIAA journal**, v. 39, n. 12, p. 2233–2241, 2001. Citado na p. 57.
- 52 SUTTHIPHORNSOMBAT, Banharn. **Prediction of torque and radial forces in permanent magnet synchronous machines using field reconstruction method**. 2010. Tese (Doutorado) – The University of Texas at Arlington. Citado na p. 28.
- 53 TAMPION, Ariadne Ann. **Double-frequency stator core vibration in large two-pole turbogenerators**. 1990. Tese (Doutorado) – University of Southampton. Citado na p. 28.
- 54 TAYLOR, James I. The vibration analysis handbook. Vibration Consultants, 2005. Citado na p. 26.
- 55 VALAVI, Mostafa; NYSVEEN, Arne; NILSSEN, Robert. Characterization of radial magnetic forces in low-speed permanent magnet wind generator with non-overlapping concentrated windings, p. 2943–2948, 2012. Citado na p. 28.
- 56 WAPPLER, Gilberto Paulo. Otimização paramétrica de um flange de motor elétrico. Universidade do Estado de Santa Catarina, 2014. Citado nas pp. 24, 37, 80.
- 57 WILLIAMS, Christopher K.; RASMUSSEN, Carl E. **Gaussian processes for machine learning**. [S.l.]: MIT press Cambridge, MA, 2006. v. 2. Citado na p. 52.

APÊNDICE A – FUNÇÃO DENSIDADE DE PROBABILIDADE MULTIVARIADA

A função densidade de probabilidade multivariada gaussiana da Eq. (42) é proveniente da sua forma univariada da Eq. (27), a qual pode ser reescrita como

$$f_Z(z) = q \exp \left[-\frac{1}{2}(z - \mu)r(z - \mu) \right], \quad (\text{A.1})$$

onde r é uma constante positiva e q é determinada para satisfazer a condição

$$\int_{-\infty}^{\infty} f_Z(z) dz = 1. \quad (\text{A.2})$$

Considerando \mathbf{u} o vetor contendo as N médias μ do vetor \mathbf{z} em uma distribuição multivariada e \mathbf{R} uma matriz positiva definida e simétrica na forma

$$\mathbf{R} = \begin{bmatrix} r_{11} & \cdots & r_{1N} \\ \vdots & \ddots & \vdots \\ r_{N1} & \cdots & r_{NN} \end{bmatrix}, \quad (\text{A.3})$$

pode-se representar a Eq. (A.1) em uma forma multivariada como

$$f_{\mathbf{z}}(\mathbf{z}) = Q \exp \left[-\frac{1}{2}(\mathbf{z} - \mathbf{u})^T \mathbf{R}(\mathbf{z} - \mathbf{u}) \right]. \quad (\text{A.4})$$

Sendo Q a ser definido para satisfazer

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} Q \exp \left[-\frac{1}{2}(\mathbf{z} - \mathbf{u})^T \mathbf{R}(\mathbf{z} - \mathbf{u}) \right] dz_N \dots dz_1 = 1. \quad (\text{A.5})$$

Para solucionar a integral da Eq. (A.5), é admitido que existe uma matriz \mathbf{P} não singular e real que diagonaliza \mathbf{R} na forma

$$\mathbf{P}^T \mathbf{R} \mathbf{P} = \mathbf{I}, \quad (\text{A.6})$$

onde \mathbf{I} é a matriz identidade $N \times N$. Também é definida a relação

$$\mathbf{z} - \mathbf{u} = \mathbf{P} \mathbf{y}, \quad (\text{A.7})$$

assim o termo exponencial da Eq. (A.5) torna-se

$$(\mathbf{z} - \mathbf{u})^T \mathbf{R}(\mathbf{z} - \mathbf{u}) = \mathbf{y}^T \mathbf{P}^T \mathbf{R} \mathbf{P} \mathbf{y} = \mathbf{y}^T \mathbf{y}. \quad (\text{A.8})$$

Dessa forma, a Eq. (A.4) fica em função de \mathbf{y} na forma

$$f_{\mathbf{Y}}(\mathbf{y}) = Q \exp \left[-\frac{1}{2} \mathbf{y}^T \mathbf{y} \right] \mathbf{J}(\mathbf{y}), \quad (\text{A.9})$$

onde $\mathbf{J}(\mathbf{y})$ é o Jacobiano devido a troca da variável \mathbf{z} por \mathbf{y} , definido por

$$\mathbf{J}(\mathbf{y}) = abs \left| \frac{\partial \mathbf{z}}{\partial \mathbf{y}} \right| = abs |\mathbf{P}|, \quad (\text{A.10})$$

onde abs representa o operador absoluto caracterizando $f_{\mathbf{Y}}(\mathbf{y}) \geq 0$. Da Eq. (A.6) pode-se resolver $\mathbf{J}(\mathbf{y}) = (\sqrt{\mathbf{R}})^{-1}$, tornando a Eq. (A.9) em

$$f_{\mathbf{Y}}(\mathbf{y}) = Q \frac{1}{\sqrt{\mathbf{R}}} \exp \left[-\frac{1}{2} \mathbf{y}^T \mathbf{y} \right], \quad (\text{A.11})$$

a qual sua integral é unitária, logo

$$f_{\mathbf{Y}}(\mathbf{y}) = Q \frac{1}{\sqrt{\mathbf{R}}} \prod_{i=1}^N \left[\int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} y_i^2 \right) dy_i \right] = 1. \quad (\text{A.12})$$

A integral da Eq. (A.12) é conhecida como integral de Euler-Poisson e seu resultado é $\sqrt{2\pi}$, logo

$$\begin{aligned} Q \frac{1}{\sqrt{\mathbf{R}}} \prod_{i=1}^N \sqrt{2\pi} &= 1 \\ Q \frac{1}{\sqrt{\mathbf{R}}} \sqrt{2\pi}^N &= 1 \\ Q &= \frac{\sqrt{\mathbf{R}}}{(2\pi)^{\frac{N}{2}}}. \end{aligned} \quad (\text{A.13})$$

Portanto a estrutura da função densidade de probabilidade multivariada para \mathbf{z} é

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{\sqrt{\mathbf{R}}}{(2\pi)^{\frac{N}{2}}} \exp \left[-\frac{1}{2} (\mathbf{z} - \mathbf{u})^T \mathbf{R} (\mathbf{z} - \mathbf{u}) \right], \quad (\text{A.14})$$

sendo ainda necessário provar que o vetor \mathbf{u} contém a esperança de \mathbf{z} e \mathbf{R} tenha relação com a matriz de variâncias-covariâncias \mathbf{C} .

Remanejando a Eq. (A.7) obtém-se o valor esperado do vetor aleatório \mathbf{Z} como

$$E[\mathbf{Z}] = \mathbf{P}E[\mathbf{Y}] + \mathbf{u}. \quad (\text{A.15})$$

O valor esperado $E[\mathbf{Y}]$ é computado utilizando a Eq. (39), onde

$$E[\mathbf{Y}] = \frac{1}{(2\pi)^{\frac{N}{2}}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y_i \exp \left[-\frac{1}{2} y_i^2 \right] dy_N \cdots dy_1, \quad (\text{A.16})$$

onde a integral é simétrica em y_i , logo $E[\mathbf{Y}] = 0$. Portanto, pode-se constatar que \mathbf{u} contém a esperança matemática do vetor aleatório \mathbf{Z} como

$$\begin{aligned}
E[\mathbf{Z}] &= \mathbf{P}\cdot\mathbf{0} + \mathbf{u} \\
E[\mathbf{Z}] &= \mathbf{u}.
\end{aligned} \tag{A.17}$$

Para definir \mathbf{R} , computa-se a covariância do vetor aleatório \mathbf{Z} utilizando a Eq. (41) na forma

$$\begin{aligned}
Cov[Z_i, Z_j] &= E[(\mathbf{P}\mathbf{Y} + \mathbf{u} - (\mathbf{P}E[\mathbf{Y}] + \mathbf{u}))(\mathbf{P}\mathbf{Y} + \mathbf{u} - \mathbf{P}E[\mathbf{Y}] + \mathbf{u})^T] \\
Cov[Z_i, Z_j] &= E[(\mathbf{P}\mathbf{Y})(\mathbf{P}\mathbf{Y})^T] \\
Cov[Z_i, Z_j] &= E[\mathbf{P}\mathbf{Y}\mathbf{Y}^T\mathbf{P}] \\
Cov[Z_i, Z_j] &= \mathbf{P}E[\mathbf{Y}\mathbf{Y}^T]\mathbf{P}.
\end{aligned} \tag{A.18}$$

Considerando que $E[\mathbf{Y}\mathbf{Y}^T]$ pode ser representado por $E[Y_i Y_j]$, assim

$$E[Y_i, Y_j] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y_i y_j \prod_{h=1}^N \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_h^2\right) \right] dy_N \dots dy_1. \tag{A.19}$$

Quando $i = j$ a Eq. (A.19) torna-se

$$\begin{aligned}
E[Y_i^2] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y_i^2 \exp\left(-\frac{1}{2}y_i^2\right) dy_i \prod_{h=1, h \neq i}^N \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_h^2\right) dy_h \right] \\
E[Y_i^2] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y_i^2 \exp\left(-\frac{1}{2}y_i^2\right) dy_i \\
E[Y_i^2] &= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} E[Y_i^2] = 1.
\end{aligned}$$

Por sua vez, quando $i \neq j$, o valor esperado $E[Y_i, Y_j]$ é

$$\begin{aligned}
E[Y_i, Y_j] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y_i \exp\left(-\frac{1}{2}y_i^2\right) dy_i \\
&\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y_j^2 \exp\left(-\frac{1}{2}y_j^2\right) dy_j \cdot \prod_{h=1, h \neq i, j}^N \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y_h^2\right) dy_h \right] \\
E[Y_i, Y_j] &= 0.0.1 \\
E[Y_i, Y_j] &= 0.
\end{aligned} \tag{A.20}$$

Logo, $E[\mathbf{Y}\mathbf{Y}^T] = \mathbf{I}$, assim $Cov[Z_i, Z_j]$ torna-se

$$Cov[Z_i, Z_j] = \mathbf{P}\mathbf{I}\mathbf{P}^T = \mathbf{P}\mathbf{P}^T. \tag{A.21}$$

Considerando a inversa em ambos os lados da Eq. (A.6) na forma

$$\begin{aligned}
(\mathbf{P}^T\mathbf{R}\mathbf{P})^{-1} &= \mathbf{I}^{-1} \\
\mathbf{P}^{-1}\mathbf{R}\mathbf{P} &= \mathbf{I}
\end{aligned} \tag{A.22}$$

e multiplicando por \mathbf{P} a esquerda e \mathbf{P}^T a direita obtém-se

$$\begin{aligned}\mathbf{P}\mathbf{P}^{-1}\mathbf{R}^{-1}\mathbf{P}^{-T}\mathbf{P}^T &= \mathbf{P}\mathbf{I}\mathbf{P}^T \\ \mathbf{R}^{-1} &= \mathbf{P}\mathbf{P}^T,\end{aligned}\tag{A.23}$$

concluindo que

$$\mathbf{C} = Cov[Z_i, Z_j] = \mathbf{R}^{-1}.\tag{A.24}$$

Finalmente, substituindo a Eq. (A.24) na Eq. (A.14) e considerando a Eq. (A.17), o vetor aleatório \mathbf{Z} pode ser associado a uma distribuição normal multivariada $\mathbf{Z} \sim \mathcal{N}(E[\mathbf{Z}], \mathbf{C})$, conforme a Eq. (42).

APÊNDICE B – VALOR ESPERADO E COVARIÂNCIA CONDICIONAL

Para determinar uma distribuição Gaussiana multivariada condicional deve-se obter os parâmetros de valor esperado e da covariância condicional conforme as Eqs. (44) e (45). Define-se um vetor aleatório composto por \mathbf{Z}_1 e \mathbf{Z}_2 na forma $\mathbf{Z} = \{\mathbf{Z}_1, \mathbf{Z}_2\}^T$, onde cada componente desse vetor tem seu valor esperado $E[\mathbf{Z}] = \{E[\mathbf{Z}_1], E[\mathbf{Z}_2]\}$ e uma matriz de variâncias-covariâncias conforme a Eq. (43). Assim, pode-se expandir o termo exponencial da Eq. (42) na forma

$$\supset = \left(-\frac{1}{2} \begin{bmatrix} \mathbf{Z}_1 - E[\mathbf{Z}_1] \\ \mathbf{Z}_2 - E[\mathbf{Z}_2] \end{bmatrix}^T \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{Z}_1 - E[\mathbf{Z}_1] \\ \mathbf{Z}_2 - E[\mathbf{Z}_2] \end{bmatrix} \right). \quad (\text{B.1})$$

Podendo ser desenvolvido como

$$\begin{aligned} \supset &= \left(-\frac{1}{2} \begin{bmatrix} \mathbf{Z}_1 - E[\mathbf{Z}_1] \\ \mathbf{Z}_2 - E[\mathbf{Z}_2] \end{bmatrix}^T \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_1 - E[\mathbf{Z}_1] \\ \mathbf{Z}_2 - E[\mathbf{Z}_2] \end{bmatrix} \right) \\ \supset &= -\left[\frac{1}{2} (\mathbf{Z}_1 - E[\mathbf{Z}_1])^T \mathbf{H}_{11} (\mathbf{Z}_1 - E[\mathbf{Z}_1]) + \frac{1}{2} (\mathbf{Z}_1 - E[\mathbf{Z}_1])^T \mathbf{H}_{12} (\mathbf{Z}_2 - E[\mathbf{Z}_2]) + \right. \\ &\left. + \frac{1}{2} (\mathbf{Z}_2 - E[\mathbf{Z}_2])^T \mathbf{H}_{21} (\mathbf{Z}_1 - E[\mathbf{Z}_1]) + \frac{1}{2} (\mathbf{Z}_2 - E[\mathbf{Z}_2])^T \mathbf{H}_{22} (\mathbf{Z}_2 - E[\mathbf{Z}_2]) \right], \quad (\text{B.2}) \end{aligned}$$

onde \mathbf{H} e seus subíndices representam a inversa de \mathbf{C} particionada conforme a Eq. (E.2).

Utilizando a propriedade de complemento dos quadrados, a qual diz

$$\frac{1}{2} \check{\mathbf{h}}^T \check{\mathbf{A}} \check{\mathbf{h}} + \check{\mathbf{b}}^T \check{\mathbf{h}} + \check{c} = \frac{1}{2} (\check{\mathbf{h}} + \check{\mathbf{A}}^{-1} \check{\mathbf{b}})^T \check{\mathbf{A}} (\check{\mathbf{h}} + \check{\mathbf{A}}^{-1} \check{\mathbf{b}}) + \check{c} - \frac{1}{2} \check{\mathbf{b}}^T \check{\mathbf{A}}^{-1} \check{\mathbf{b}}, \quad (\text{B.3})$$

onde $\check{\mathbf{A}}$ é uma matriz simétrica e não-singular, define-se

$$\check{\mathbf{h}} = \mathbf{Z}_1 - E[\mathbf{Z}_1], \quad (\text{B.4})$$

$$\check{\mathbf{A}} = \mathbf{H}_{11}, \quad (\text{B.5})$$

$$\check{\mathbf{b}} = \mathbf{H}_{21} (\mathbf{Z}_2 - E[\mathbf{Z}_2]) \quad (\text{B.6})$$

e

$$\check{c} = \frac{1}{2} (\mathbf{Z}_2 - E[\mathbf{Z}_2])^T \mathbf{H}_{22} (\mathbf{Z}_2 - E[\mathbf{Z}_2]). \quad (\text{B.7})$$

Utilizando a Eq. (B.3) juntamente com as Eqs. (B.4-B.7) é possível escrever uma distribuição multivariada para \mathbf{Z}_1 condicionado a \mathbf{Z}_2 como

$$\begin{aligned}
f_{\mathbf{z}}(\mathbf{Z}_1|\mathbf{Z}_2) = \frac{1}{\mathfrak{C}} \exp \left[- \left(\frac{1}{2}(\mathbf{Z}_1 - E[\mathbf{Z}_1] + \mathbf{H}_{11}^{-1}\mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2]))^T \mathbf{H}_{11}(\mathbf{Z}_1 - E[\mathbf{Z}_1]) + \right. \right. \\
\left. \left. + \mathbf{H}_{11}^{-1}\mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2])) + \frac{1}{2}(\mathbf{Z}_2 - E[\mathbf{Z}_2])^T \mathbf{H}_{22}(\mathbf{Z}_2 - E[\mathbf{Z}_2]) - \right. \right. \\
\left. \left. - \frac{1}{2}(\mathbf{Z}_2 - E[\mathbf{Z}_2])^T \mathbf{H}_{21} \mathbf{H}_{11}^{-1} \mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2]) \right) \right], \tag{B.8}
\end{aligned}$$

onde \mathfrak{C} é uma constante independente de \mathbf{Z}_1 . Considerando somente os termos dependentes de \mathbf{Z}_1 chega-se à

$$\begin{aligned}
f_{\mathbf{z}}(\mathbf{Z}_1|\mathbf{Z}_2) = \frac{1}{\mathfrak{C}} \exp \left[- \left(\frac{1}{2}(\mathbf{Z}_1 - E[\mathbf{Z}_1] + \mathbf{H}_{11}^{-1}\mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2]))^T \mathbf{H}_{11}(\mathbf{Z}_1 - E[\mathbf{Z}_1]) + \right. \right. \\
\left. \left. + \mathbf{H}_{11}^{-1}\mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2])) \right) \right]. \tag{B.9}
\end{aligned}$$

Nota-se que $E[\mathbf{Z}_1|\mathbf{Z}_2]$ pode ser escrita como

$$E[\mathbf{Z}_1|\mathbf{Z}_2] = E[\mathbf{Z}_1] - \mathbf{H}_{11}^{-1}\mathbf{H}_{21}(\mathbf{Z}_2 - E[\mathbf{Z}_2]), \tag{B.10}$$

onde o termo $\mathbf{H}_{11}^{-1}\mathbf{H}_{21}$ torna-se $-\mathbf{C}_{12}\mathbf{C}_{22}^{-1}$ pela Eq. (E.13), assim

$$E[\mathbf{Z}_1|\mathbf{Z}_2] = E[\mathbf{Z}_1] + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}(\mathbf{Z}_2 - E[\mathbf{Z}_2]). \tag{B.11}$$

Enquanto a matriz de variâncias-covariâncias pode ser escrita como

$$C[\mathbf{Z}_1|\mathbf{Z}_2] = \mathbf{H}_{11}^{-1} = \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{12}^T. \tag{B.12}$$

APÊNDICE C – DERIVADA DA FUNÇÃO DENSIDADE DE PROBABILIDADE UNIVARIADA EM RELAÇÃO A μ E A σ^2

A maximização (ou minimização) da função verossimilhança da Eq. (57) no Capítulo 4 consiste em encontrar os parâmetros definido por $\hat{\boldsymbol{\kappa}} = \{\hat{\mu}, \hat{\sigma}^2\}$. Iniciando pela definição de $\hat{\mu}$ temos

$$\frac{\partial \ln L}{\partial \mu} = \frac{\partial}{\partial \mu} \left(-\frac{n}{2} \ln(2\pi\sigma^2) \right) + \frac{\partial}{\partial \mu} \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 \right), \quad (\text{C.1})$$

onde o primeiro termo não depende de μ e tem sua derivada nula, portando

$$\frac{\partial \ln L}{\partial \mu} = \frac{\partial}{\partial \mu} \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 \right). \quad (\text{C.2})$$

Considerando a derivada parcial para o termo dependente, é possível reorganizar na forma

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^n \frac{\partial}{\partial \mu} (z_i - \mu)^2, \quad (\text{C.3})$$

tal que

$$\frac{\partial}{\partial \mu} (z_i - \mu)^2 = 2(z_i - \mu)(-1), \quad (\text{C.4})$$

obtendo

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu). \quad (\text{C.5})$$

Igualando a Eq. (C.5) à zero como

$$\begin{aligned} 0 &= \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu) \\ 0 &= \sum_{i=1}^n z_i - \sum_{i=1}^n \mu \\ n\mu &= \sum_{i=1}^n z_i, \end{aligned} \quad (\text{C.6})$$

defina-se o valor esperado que maximiza a Eq. (57) na forma

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n z_i \quad (\text{C.7})$$

Derivando parcialmente $\ln L$ em relação a σ^2 , temos

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{\partial}{\partial \sigma^2} \left(-\frac{n}{2} \ln(2\pi\sigma^2) \right) + \frac{\partial}{\partial \sigma^2} \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 \right), \quad (\text{C.8})$$

onde o primeiro termo pode ser escrito como

$$-\frac{n}{2} \frac{\partial}{\partial \sigma^2} \ln (2\pi\sigma^2) = -\frac{n}{2} \frac{1}{\sigma^2}, \quad (\text{C.9})$$

obtendo

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\partial}{\partial \sigma^2} \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 \right). \quad (\text{C.10})$$

Enquanto o segundo termo pode ser rearranjado para que

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{1}{2} \frac{\partial}{\partial \sigma^2} \sigma^{-2} \sum_{i=1}^n (z_i - \mu)^2, \quad (\text{C.11})$$

tal que

$$\frac{\partial}{\partial \sigma^2} \sigma^{-2} \sum_{i=1}^n (z_i - \mu)^2 = -\frac{1}{\sigma^4} \sum_{i=1}^n (z_i - \mu)^2, \quad (\text{C.12})$$

obtendo

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{1}{2} \left(-\frac{1}{\sigma^4} \sum_{i=1}^n (z_i - \mu)^2 \right). \quad (\text{C.13})$$

Pode-se rearranjar a Eq. (C.13) isolando o termo $\frac{1}{2\sigma^2}$ e igualando a mesma a zero na forma

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{1}{2\sigma^2} \left(-n + \frac{1}{\sigma^2} \sum_{i=1}^n (z_i - \mu)^2 \right) = 0, \quad (\text{C.14})$$

onde

$$n\sigma^2 = \sum_{i=1}^n (z_i - \mu)^2. \quad (\text{C.15})$$

Finalmente definindo a variância que maximiza a Eq. (57) como

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \mu)^2. \quad (\text{C.16})$$

**APÊNDICE D – DERIVADA DA FUNÇÃO DENSIDADE DE
PROBABILIDADE MULTIVARIADA EM RELAÇÃO A μ E A σ^2**

No Capítulo 6 é necessário obter os parâmetros $\hat{\boldsymbol{\kappa}} = \{\hat{\boldsymbol{\mu}}, \hat{\sigma}^2\}$, derivando parcialmente a Eq. (88) de forma independente em relação a $\hat{\boldsymbol{\kappa}}$. Iniciando pela definição de $\hat{\boldsymbol{\mu}}$ temos

$$\frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\mu}} = -\frac{\partial}{\partial \boldsymbol{\mu}} \frac{n}{2} \ln(2\pi) - \frac{\partial}{\partial \boldsymbol{\mu}} \frac{n}{2} \ln(\sigma^2) - \frac{\partial}{\partial \boldsymbol{\mu}} \frac{1}{2} \ln |\boldsymbol{\Psi}| - \frac{\partial}{\partial \boldsymbol{\mu}} \frac{(\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})}{2\sigma^2}, \quad (\text{D.1})$$

onde somente o último termo é dependente de $\boldsymbol{\mu}$, portanto

$$\frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\mu}} = -\frac{\partial}{\partial \boldsymbol{\mu}} \frac{(\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})}{2\sigma^2}. \quad (\text{D.2})$$

Considerando a derivada parcial dentro do termo dependente na forma

$$\frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\mu}} = -\frac{1}{2\sigma^2} \frac{\partial}{\partial \boldsymbol{\mu}} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu}) \quad (\text{D.3})$$

e desenvolvendo sua multiplicação obtemos

$$\begin{aligned} \frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\mu}} &= -\frac{1}{2\sigma^2} \frac{\partial}{\partial \boldsymbol{\mu}} (\mathbf{y}\boldsymbol{\Psi}^{-1} - \mathbf{1}\boldsymbol{\mu}\boldsymbol{\Psi}^{-1})(\mathbf{y} - \mathbf{1}\boldsymbol{\mu}) \\ \frac{\partial \ln \mathcal{L}}{\partial \boldsymbol{\mu}} &= -\frac{1}{2\sigma^2} \frac{\partial}{\partial \boldsymbol{\mu}} (\mathbf{y}^2 \boldsymbol{\Psi}^{-1} - \mathbf{y}^T \boldsymbol{\Psi}^{-1} \mathbf{1}\boldsymbol{\mu} - \mathbf{1}^T \boldsymbol{\mu} \boldsymbol{\Psi}^{-1} \mathbf{y} + \mathbf{1}\boldsymbol{\mu}^2 \boldsymbol{\Psi}^{-1} \mathbf{1}). \end{aligned} \quad (\text{D.4})$$

Computando a derivada parcial e igualando à zero resulta em

$$\begin{aligned} 0 &= -\frac{1}{2\sigma^2} (\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{y} + \mathbf{1}^T \boldsymbol{\mu} \boldsymbol{\Psi}^{-1} \mathbf{1}) \\ 0 &= -(\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{y} + \mathbf{1}^T \boldsymbol{\mu} \boldsymbol{\Psi}^{-1} \mathbf{1}). \end{aligned} \quad (\text{D.5})$$

Finalmente, isolando o termo $\boldsymbol{\mu}$, a média estimada $\hat{\boldsymbol{\mu}}$ é

$$\hat{\boldsymbol{\mu}} = \frac{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{y}}{\mathbf{1}^T \boldsymbol{\Psi}^{-1} \mathbf{1}}. \quad (\text{D.6})$$

A derivada parcial da Eq. (88) em relação a σ^2 é

$$\frac{\partial \ln \mathcal{L}}{\partial \sigma^2} = -\frac{\partial}{\partial \sigma^2} \frac{n}{2} \ln(2\pi) - \frac{\partial}{\partial \sigma^2} \frac{n}{2} \ln(\sigma^2) - \frac{\partial}{\partial \sigma^2} \frac{1}{2} \ln |\boldsymbol{\Psi}| - \frac{\partial}{\partial \sigma^2} \frac{(\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})}{2\sigma^2}, \quad (\text{D.7})$$

onde somente o segundo e o último termo são dependentes de σ^2 , portanto

$$\frac{\partial \ln \mathcal{L}}{\partial \sigma^2} = -\frac{\partial}{\partial \sigma^2} \frac{n}{2} \ln(\sigma^2) - \frac{\partial}{\partial \sigma^2} \frac{(\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})}{2\sigma^2}. \quad (\text{D.8})$$

Alocando a derivada parcial dentro dos termos dependentes na forma

$$\frac{\partial \ln \mathcal{L}}{\partial \mu} = -\frac{n}{2} \frac{\partial}{\partial \sigma^2} \ln(\sigma^2) - \frac{1}{2} (\mathbf{y} - \mathbf{1}\mu)^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\mu) \frac{\partial}{\partial \sigma^2} \frac{1}{\sigma^2}, \quad (\text{D.9})$$

onde igualando a zero e computando a derivada parcial resulta em

$$0 = -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2} (\mathbf{y} - \mathbf{1}\mu)^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\mu), \quad (\text{D.10})$$

Por fim, organizando os termos e isolando σ^2 , a variância estimada $\hat{\sigma}^2$ é definida por

$$\hat{\sigma}^2 = \frac{(\mathbf{y} - \mathbf{1}\hat{\mu})^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{n}. \quad (\text{D.11})$$

APÊNDICE E – MATRIZ INVERSA PARTICIONADA

Para determinarmos a predição no método de Kriging, conforme o Cap. 6, deve-se realizar a inversão da matriz $\tilde{\Psi}$ na Eq. (96), satisfazendo a relação $\tilde{\Psi} \cdot \tilde{\Psi}^{-1} = \mathbf{I}$.

Considerando a matriz $\mathbf{F}_{n \times n}$ particionada como

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \quad (\text{E.1})$$

e sua inversa $\mathbf{H} = \mathbf{F}^{-1}$, também particionada na forma

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}, \quad (\text{E.2})$$

onde cada submatriz possui as seguintes dimensões

- F_{11} e H_{11} são $p \times p$;
- F_{22} e H_{22} são $q \times q$;
- $F_{12} = F_{12}^T$ e $H_{12} = H_{12}^T$ são $p \times q$;

e $p + q = n$. Logo

$$\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \mathbf{I}, \quad (\text{E.3})$$

onde \mathbf{I} é a matriz identidade $n \times n$. A multiplicação das matrizes \mathbf{F} e \mathbf{H} na Eq. (E.3) é computada na forma

$$\begin{bmatrix} F_{11}H_{11} + F_{12}H_{12}^T & F_{11}H_{12} + F_{12}H_{22} \\ F_{12}^T H_{11} + F_{22}H_{12}^T & F_{12}^T H_{12} + F_{22}H_{22} \end{bmatrix} = \begin{bmatrix} I_{11} & 0 \\ 0 & I_{22} \end{bmatrix}, \quad (\text{E.4})$$

onde do sistema de equações é obtido

$$\begin{aligned} F_{11}H_{11} + F_{12}H_{12}^T &= I_{11} \\ H_{11} &= F_{11}^{-1} - F_{11}^{-1}F_{12}H_{12}^T, \end{aligned} \quad (\text{E.5})$$

$$\begin{aligned} F_{11}H_{12} + F_{12}H_{22} &= 0 \\ H_{12} &= -F_{11}^{-1}F_{12}H_{22}, \end{aligned} \quad (\text{E.6})$$

$$\begin{aligned} F_{12}^T H_{11} + F_{22}H_{12}^T &= 0 \\ H_{12}^T &= -F_{22}^{-1}F_{12}^T H_{11} \end{aligned} \quad (\text{E.7})$$

e

$$\begin{aligned} F_{12}^T H_{12} + F_{22} H_{22} &= I_{22} \\ H_{22} &= F_{22}^{-1} - F_{22}^{-1} F_{12}^T H_{12}. \end{aligned} \quad (\text{E.8})$$

Substituindo a Eq. (E.7) na Eq. (E.5) temos

$$\begin{aligned} H_{11} &= F_{11}^{-1} + F_{11}^{-1} F_{12} F_{22}^{-1} F_{12}^T H_{11} \\ F_{11}^{-1} &= (I_{11} - F_{11}^{-1} F_{12} F_{22}^{-1} F_{12}^T) H_{11} \\ I_{11} &= (F_{11} - F_{12} F_{22}^{-1} F_{12}^T) H_{11} \\ H_{11} &= (F_{11} - F_{12} F_{22}^{-1} F_{12}^T)^{-1}. \end{aligned} \quad (\text{E.9})$$

De forma análoga, substituindo a Eq. (E.6) na Eq. (E.8) temos

$$\begin{aligned} H_{22} &= F_{22}^{-1} + F_{22}^{-1} F_{12}^T F_{11}^{-1} F_{12} H_{22} \\ F_{22}^{-1} &= (I_{22} - F_{22}^{-1} F_{12}^T F_{11}^{-1} F_{12}) H_{22} \\ I_{22} &= (F_{22} - F_{12}^T F_{11}^{-1} F_{12}) H_{22} \\ H_{22} &= (F_{22} - F_{12}^T F_{11}^{-1} F_{12})^{-1} \end{aligned} \quad (\text{E.10})$$

Calculando os termos cruzados, deve-se substituir a Eq. (E.10) na Eq. (E.6), onde

$$H_{12} = -F_{11}^{-1} F_{12} (F_{22} - F_{12}^T F_{11}^{-1} F_{12})^{-1}. \quad (\text{E.11})$$

Enquanto para H_{12}^T , é necessário substituir a Eq. (E.9) na Eq. (E.7), logo

$$H_{12}^T = -F_{22}^{-1} F_{12}^T (F_{11} - F_{12} F_{22}^{-1} F_{12}^T)^{-1}. \quad (\text{E.12})$$

Finalmente a inversa de uma matriz particionada como da Eq. (E.1) é

$$\mathbf{H} = \mathbf{F}^{-1} = \begin{bmatrix} (F_{11} - F_{12} F_{22}^{-1} F_{12}^T)^{-1} & -F_{11}^{-1} F_{12} (F_{22} - F_{12}^T F_{11}^{-1} F_{12})^{-1} \\ -F_{22}^{-1} F_{12}^T (F_{11} - F_{12} F_{22}^{-1} F_{12}^T)^{-1} & (F_{22} - F_{12}^T F_{11}^{-1} F_{12})^{-1} \end{bmatrix}. \quad (\text{E.13})$$

Considerando inverter a matriz $\tilde{\Psi}$, onde $F_{11} = \Psi$, $F_{12} = \psi$ e $F_{22} = 1$, substitui-se esses valores na Eq. (E.13), portanto

$$\tilde{\Psi}^{-1} = \begin{bmatrix} (\Psi - \psi 1 \psi^T)^{-1} & -\Psi^{-1} \psi (1 - \psi^T \Psi^{-1} \psi)^{-1} \\ -\psi^T (\Psi - \psi 1 \psi^T)^{-1} & (1 - \psi^T \Psi^{-1} \psi)^{-1} \end{bmatrix}, \quad (\text{E.14})$$

podendo ser organizada na forma

$$\tilde{\Psi}^{-1} = \begin{bmatrix} \Psi^{-1} + \Psi^{-1} \psi (1 - \psi^T \Psi^{-1} \psi)^{-1} \psi^T \Psi^{-1} & -\Psi^{-1} \psi (1 - \psi^T \Psi^{-1} \psi)^{-1} \\ -(1 - \psi^T \Psi^{-1} \psi)^{-1} \psi^T \Psi^{-1} & (1 - \psi^T \Psi^{-1} \psi)^{-1} \end{bmatrix}. \quad (\text{E.15})$$

APÊNDICE F – DESENVOLVIMENTO MATEMÁTICO PARA A PREDIÇÃO DE KRIGING

A predição da Eq. (97) é obtida partindo da forma concentrada da função verossimilhança. Assim, pode-se expandir a Eq. (96) na forma

$$\ln(\mathcal{L}(\hat{y}|\mathbf{X}, \mathbf{y}, \hat{\mu}, \hat{\sigma}^2, \theta, p)) = \frac{- \begin{Bmatrix} \mathbf{y} - \mathbf{1}\hat{\mu} \\ \hat{y} - \hat{\mu} \end{Bmatrix}^T \begin{bmatrix} \mathbf{\Psi} & \boldsymbol{\psi} \\ \boldsymbol{\psi}^T & 1 \end{bmatrix}^{-1} \begin{Bmatrix} \mathbf{y} - \mathbf{1}\hat{\mu} \\ \hat{y} - \hat{\mu} \end{Bmatrix}}{2\hat{\sigma}^2}, \quad (\text{F.1})$$

onde $\tilde{\Psi}^{-1}$ é solucionada conforme Apêndice E. Substituindo a Eq. (E.15) na Eq. (F.1) e isolando os termos com \hat{y} temos a função logaritmo concentrada de verossimilhança aumentada na forma

$$\ln[\mathcal{L}(\hat{y})] = \left(\frac{-1}{2\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu})^2 + \left(\frac{\boldsymbol{\psi}^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu}). \quad (\text{F.2})$$

Derivando a Eq. (F.2) em função de \hat{y} e igualando a zero, tal que

$$\begin{aligned} \frac{d \ln[\mathcal{L}(\hat{y})]}{d\hat{y}} &= 0 \\ \frac{d \ln[\mathcal{L}(\hat{y})]}{d\hat{y}} &= \left(\frac{-2}{2\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) \hat{y} - \left(\frac{-2}{2\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) \hat{\mu} + \left(\frac{\boldsymbol{\psi}^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) \hat{y} = 0 \\ \ln[\mathcal{L}(\hat{y})] &= \left(\frac{-1}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) (\hat{y} - \hat{\mu}) + \left(\frac{\boldsymbol{\psi}^T \mathbf{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu})}{\hat{\sigma}^2(1 - \boldsymbol{\psi}^T \mathbf{\Psi}^{-1} \boldsymbol{\psi})} \right) = 0. \end{aligned} \quad (\text{F.3})$$

Finalmente, isolando \hat{y} , chega-se na predição da Eq. (97).

APÊNDICE G – DESENVOLVIMENTO MATEMÁTICO PARA O EI

Uma forma de adquirir novos pontos para modelos substitutos é através do método EI (*Expected Improvement*) da Eq. (110), a qual é proveniente da Eq. (109). Sendo a solução para sua integral através da substituição das variáveis

$$\mathcal{Z} = \frac{y - \hat{y}(\mathbf{x})}{s} \quad (\text{G.1})$$

e

$$\mathcal{J} = \frac{y_{min} - \hat{y}(\mathbf{x})}{s}. \quad (\text{G.2})$$

Da Eq. (G.1) pode-se obter a derivada em relação a variável de interesse y na forma

$$dy = sd\mathcal{Z}, \quad (\text{G.3})$$

enquanto da parcela positiva da Eq. (104) pode-se obter a derivada $dI = -dy$, onde substituindo a Eq. (G.3) chega-se a

$$dI = -sd\mathcal{Z}. \quad (\text{G.4})$$

Utilizando as Eqs. (G.1) e (G.2) pode-se reescrever a parcela positiva da Eq. (104) para uma relação de I como

$$I = y_{min} - y = s\mathcal{J} + \hat{y}(\mathbf{x}) - y = s\mathcal{J} - s\mathcal{Z} = s(\mathcal{J} - \mathcal{Z}), \quad (\text{G.5})$$

Portanto, considera-se a relação

$$I = \begin{cases} s(\mathcal{J} - \mathcal{Z}) & \text{se } \mathcal{Z} < \mathcal{J} \\ 0 & \text{caso contrário} \end{cases} \quad (\text{G.6})$$

para evitar uma melhora negativa e substituindo as Eqs. (G.1), (G.2) e (G.4) na Eq. (109) obtém-se

$$\begin{aligned} E(I) &= \frac{1}{\sqrt{2\pi s^2}} \int_{-\infty}^{\mathcal{J}} s(\mathcal{J} - \mathcal{Z}) \exp\left(-\frac{\mathcal{Z}^2}{2}\right) sd\mathcal{Z} \\ &= s \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathcal{J}} \mathcal{J} \exp\left(-\frac{\mathcal{Z}^2}{2}\right) d\mathcal{Z} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathcal{J}} \mathcal{Z} \exp\left(-\frac{\mathcal{Z}^2}{2}\right) d\mathcal{Z} \right]. \end{aligned} \quad (\text{G.7})$$

Onde calculando a última integral dentro dos colchetes da Eq. (G.7) obtemos

$$= s \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathcal{J}} \mathcal{J} \exp\left(-\frac{\mathcal{Z}^2}{2}\right) d\mathcal{Z} + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mathcal{J}^2}{2}\right) \right], \quad (\text{G.8})$$

podendo ser reduzida na forma

$$E(I) = s[\mathcal{J}\mathcal{H}(\mathcal{J}) + \mathcal{I}(\mathcal{J})] \quad (\text{G.9})$$

ou dependente do vetor de variáveis aleatórias \mathbf{x} como na Eq. (110).

APÊNDICE H – RESULTADOS DO PROBLEMA EXEMPLO DA METODOLOGIA

Para avaliar a metodologia proposta no Cap. 7 é utilizado um problema de otimização de uma treliça de 3 barras encontrado em Arora (2004). Assim, foi definido 3 cenários para selecionar a melhor função de aquisição que se adequa a metodologia. Foram realizados 5 testes para o cenário A, B e C, os quais iniciam a metodologia com hipercubo latino com 15, 10 e 5 pontos de suportes iniciais respectivamente. Tomando as considerações da Tab. 3, a solução ótima é $x_1^* = 6,31$ [in^2] e $x_2^* = 3,26$ [in^2], enquanto o valor da função objetivo y é de 21,11 [lb]. Os valores das restrições g_1 , g_2 e g_3 em \mathbf{x}^* são 5000 [psi], 3663,22 [psi] e -1338 [psi] respectivamente, onde a restrição g_1 está ativa.

As Tabs. 10 até 14 apresentam os resultados obtidos para os 5 testes para cada cenário. Onde \hat{x}_1^* em [in^2], \hat{x}_2^* em [in^2], $\hat{y}(\hat{\mathbf{x}}^*)$ em [lb], $\hat{g}_1(\hat{\mathbf{x}}^*)$ em [psi], $\hat{g}_2(\hat{\mathbf{x}}^*)$ em [psi] e $\hat{g}_3(\hat{\mathbf{x}}^*)$ em [psi] são os valores ótimos, da função objetivo e das restrições respectivamente encontrados utilizando o metodologia proposta. $y(\hat{\mathbf{x}}^*)$ em [lb], $g_1(\hat{\mathbf{x}}^*)$ em [psi], $g_2(\hat{\mathbf{x}}^*)$ em [psi] e $g_3(\hat{\mathbf{x}}^*)$ em [psi] são as funções de referência do problema de otimização aplicando os valores ótimos ($\hat{\mathbf{x}}^*$) encontrados pela metodologia proposta. $d_{\hat{x}_1^*}$, $d_{\hat{x}_2^*}$, $d_{\hat{y}(\hat{\mathbf{x}}^*)}$, $d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$, $d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$ e $d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$ representam a diferença em [%] entre os valores das variáveis de projeto, da função objetivo e das restrições em relação aos respectivos valores de referência. $d_{y(\hat{\mathbf{x}}^*)}$, $d_{g_1(\hat{\mathbf{x}}^*)}$, $d_{g_2(\hat{\mathbf{x}}^*)}$ e $d_{g_3(\hat{\mathbf{x}}^*)}$ denotam a diferença em [%] das funções de referência em ($\hat{\mathbf{x}}^*$) em relação a elas mesmas em (\mathbf{x}^*). Finalmente, $n_{\hat{y}(\hat{\mathbf{x}}^*)}$, $n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$, $n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$ e $n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$ representam o número de pontos adicionais cada função de aquisição acrescentou ao modelo de Kriging.

Tabela 10 – Resultados do teste 1.

Variável	Cenário A - 15 pontos					Cenário B - 10 pontos					Cenário C - 5 pontos				
	EI	PI	LCB	s^2		EI	PI	LCB	s^2		EI	PI	LCB	s^2	
\hat{x}_1^*	6.35	6.29	6.92	6.16		6.64	6.66	6.64	5.88		6.41	6.17	7.13	7.18	
$d_{\hat{x}_1^*}$	0.645669	-0.30207	8.82948	-2.41883		4.98494	5.27027	4.98494	-7.29592		1.575663	-2.25284	11.51473	12.13092	
\hat{x}_2^*	3.35	3.8	2.22	3.4		2.26	1.66	2.09	4.35		3.22	3.92	1	1	
$d_{\hat{x}_2^*}$	2.686567	14.21053	-46.8468	4.117647		2.26	1.66	2.09	4.35		-1.24224	16.83673	-226	-226	
$\hat{y}(\hat{\mathbf{x}}^*)$	21.32	21.27	21.63	20.85		21.04	20.65	20.9	20.99		21.33	21.41	20.97	21.34	
$d_{\hat{y}(\hat{\mathbf{x}}^*)}$	1.031895	0.799248	2.450301	-1.19904		-0.28517	-2.17918	-0.0002	-0.52406		1.078293	1.447922	-0.61993	1.124649	
$n_{\hat{y}(\hat{\mathbf{x}}^*)}$	1	2	1	0		1	2	1	0		1	2	1	5	
$y(\hat{\mathbf{x}}^*)$	21.1	20.78	20.89	21.1		21.1	21.07	21.07	21.1		21.07	21.09	21.04	21.08	
$d_{y(\hat{\mathbf{x}}^*)}$	0	-1.53994	-1.00527	0		0	-0.14238	-0.14238	0		-0.14238	-0.04742	-0.28517	-0.09488	
$\hat{g}_1(\hat{\mathbf{x}}^*)$	4999.99	5000	4999.99	5000		4999.99	5000	4999.99	4999.99		4999.99	4999.99	5000	5000	
$d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	-0.0002	0	-0.0002	0		-0.0002	0	-0.0002	-0.0002		-0.0002	-0.0002	0	0	
$n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	3	2	6	20		4	2	3	20		6	5	8	20	
$g_1(\hat{\mathbf{x}}^*)$	5051.91	5124.75	5312.97	4949.83		5011.52	4972.53	5005.1	5011.03		5066.28	5084.73	5024.37	5121.56	
$d_{g_1(\hat{\mathbf{x}}^*)}$	1.027532	2.434265	5.890679	-1.01357		0.22987	-0.55244	0.101896	-17.1803		1.308258	1.666362	0.485036	2.373496	
$\hat{g}_2(\hat{\mathbf{x}}^*)$	3475.04	3206.83	3945.83	3693.41		4089.64	4507.05	4221.72	3126.14		3613.35	3493.2	3205.03	4431.19	
$d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	-5.41519	-14.2318	7.162245	0.817402		10.42683	18.72245	13.22921	-17.1803		-1.38016	-4.86717	-14.296	17.33101	
$n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	5	3	4	20		5	1	3	20		6	4	5	20	
$g_2(\hat{\mathbf{x}}^*)$	3533.2	3404.78	3525.97	3721.65		3631.82	3618.95	3591.89	326.09		3627.44	3770.8	5290.09	4314.98	
$d_{g_2(\hat{\mathbf{x}}^*)}$	-3.67995	-7.59051	-3.89255	1.570003		-0.86458	-1.22328	-1.98586	-1023.38		-0.98637	2.852975	30.75316	15.10459	
$\hat{g}_3(\hat{\mathbf{x}}^*)$	-1451.06	-1417.5	-1906.17	-1516.17		-1743.6	-768.08	-902.09	-1798.6		-1362.34	-1499.55	-1275.09	-353.02	
$d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	7.759845	5.576014	29.78276	11.72098		23.23583	-74.2605	-48.3732	1.481683		1.752866	10.74256	-4.96985	-279.146	
$n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	3	2	3	20		4	2	1	20		2	1	1	20	
$g_3(\hat{\mathbf{x}}^*)$	-1433.58	-1342.61	-1740.56	-1438.78		-1542.32	-1101.27	-1213.28	-1358.59		-1469.91	-1469.64	-1469.69	-1530.84	
$d_{g_3(\hat{\mathbf{x}}^*)}$	6.635137	0.309099	23.10176	6.972574		-1542.32	-1101.27	-1213.28	-1358.59		8.942724	8.925995	8.929094	12.56696	

Tabela 11 – Resultados do teste 2.

Variável	15 pontos					10 pontos					5 pontos				
	EI	PI	LCB	s^2		EI	PI	LCB	s^2		EI	PI	LCB	s^2	
\hat{x}_1^*	6	5.95	5.9	6.85		6.33	6.31	5.7	6.33		5.95	6.35	6	5.99	
$d_{\hat{x}_1^*}$	-5.15	-6.03361	-6.9322	7.89781	0.331754	0.015848	-10.6842	0.331754	-6.03361	0.645669	-5.15	-5.32554			
\hat{x}_2^*	4.02	4.16	4.16	1.07	3.31	3.35	4.57	3.21	4.41	1.26	4.12	4.51			
$d_{\hat{x}_2^*}$	18.90547	21.63462	21.63462	-204.673	1.510574	2.686567	28.66521	-1.55763	26.0771	-158.73	20.87379	27.71619			
$\hat{y}(\hat{\mathbf{x}}^*)$	20.99	21.02	20.84	20.45	21.22	21.19	20.69	21.12	21.35	19.45	21.2	21.47			
$d_{\hat{y}(\hat{\mathbf{x}}^*)}$	-0.52406	-0.38059	2.136465	-3.17848	0.565504	0.424729	-1.98163	0.094697	1.17096	-8.48329	0.471698	1.7233335			
$n_{\hat{y}(\hat{\mathbf{x}}^*)}$	1	3	1	0	1	3	1	1	1	3	1	1			
$y(\hat{\mathbf{x}}^*)$	21.1	21.17	21.11	21.1	21.1	21.08	21.07	21.1	21.39	21.49	21.34	21.1			
$d_{y(\hat{\mathbf{x}}^*)}$	0	0.330657	0.047371	0	0	-0.09488	-0.14238	0	1.355774	1.814798	1.124649	0			
$\hat{g}_1(\hat{\mathbf{x}}^*)$	4999.99	5000	4999.99	4999.99	5000	4999.99	5000	4999.99	4999.99	4999.99	5000	4999.99			
$d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	-0.0002	0	-0.0002	-0.0002	0	-0.0002	0	-0.0002	-0.0002	-0.0002	0	-0.0002			
$n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	4	2	3	20	6	1	6	20	5	3	4	20			
$g_1(\hat{\mathbf{x}}^*)$	5013.92	5022.6	5037.15	5105.79	5030.99	5029.46	5095.46	5006.67	5033.91	4740.92	5000.42	5098.79			
$d_{g_1(\hat{\mathbf{x}}^*)}$	0.277627	0.449966	0.73752	2.071961	0.615982	0.585749	1.873432	0.133222	0.673631	-5.46476	0.008399	1.937519			
$\hat{g}_2(\hat{\mathbf{x}}^*)$	3400.65	3340.76	3354.09	4633.92	3727.88	3746.28	3341.12	4035.51	3341.57	5730.2	3607.85	3309.11			
$d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	-7.72117	-9.65229	-9.21651	20.94771	1.734498	2.217133	-9.64048	9.225352	-9.62571	36.07169	-1.53471	-10.701			
$n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	3	1	3	20	2	1	2	20	5	3	5	20			
$g_2(\hat{\mathbf{x}}^*)$	3625.02	3592.29	3605.04	3619.41	3765.04	3799.92	3769.8	4027.55	3951.17	4033.31	4169.56	3701.27			
$d_{g_2(\hat{\mathbf{x}}^*)}$	-1.05379	-1.97451	-1.61385	-1.21042	2.704354	3.597444	2.827206	9.045946	7.287715	9.175838	12.14373	1.028026			
$\hat{g}_3(\hat{\mathbf{x}}^*)$	-1690.12	-1720.49	-1737.62	-397.29	-1226.82	-1100.93	-1730.07	-1171.86	-1239.29	-4614.86	-4883.13	-1930.97			
$d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	20.80681	22.20472	22.97165	-236.897	-9.09995	-21.5754	22.6355	-14.2167	-8.00216	70.99674	72.59012	30.68458			
$n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	1	1	1	20	2	2	1	20	12	3	3	20			
$g_3(\hat{\mathbf{x}}^*)$	-1356.34	-1368.42	-1367.68	-1479.72	-1223.61	-1111.7	-1176	-1191.74	-628.57	-4563.5	-5566.74	-1608.55			
$d_{g_3(\hat{\mathbf{x}}^*)}$	1.318254	2.189386	2.136465	9.546401	-9.38616	-20.3976	-13.8146	-12.3114	-112.937	70.67032	75.95613	16.7909			

Tabela 12 – Resultados do teste 3.

Variável	Cenário A - 15 pontos					Cenário B - 10 pontos					Cenário C - 5 pontos						
	EI	PI	LCB	s^2	EI	PI	LCB	s^2	EI	PI	LCB	s^2	EI	PI	LCB	s^2	
\hat{x}_1^*	6.28	6.26	6.1	6.09	6.02	5.98	5.87	6.32	6.92	6.45	6.32	6.93	6.92	6.45	5.51	6.93	
$d_{\hat{x}_1^*}$	-0.46178	-0.78275	-3.426229508	-3.596059113	-4.80066	-5.50167	-7.47871	0.174051	8.82948	2.186047	-14.5009	8.961039	8.82948	2.186047	-14.5009	8.961039	
\hat{x}_2^*	3.33	3.37	3.54	3.91	4.17	4.29	4.57	3.02	1	1	2.09	1	1	2.09	1	1	
$d_{\hat{x}_2^*}$	2.102102	3.264095	7.90960452	16.62404092	21.82254	24.00932	28.66521	-7.94702	-226	-226	-55.9809	-226	-226	-55.9809	-226	-226	
$\hat{y}(\hat{\mathbf{x}}^*)$	21.1	21.1	20.8	21.16	21.2	21.18	21.19	20.9	19.73	16.67	17.54	20.6	19.73	16.67	17.54	20.6	
$d_{\hat{y}(\hat{\mathbf{x}}^*)}$	0	0	-1.442307692	0.283553875	0.471698	0.377715	0.424729	-0.95694	-6.94374	-26.5747	-20.2965	-2.42718	-6.94374	-26.5747	-20.2965	-2.42718	
$n_{\hat{y}(\hat{\mathbf{x}}^*)}$	1	2	1	0	1	3	1	1	1	5	1	5	1	5	1	5	
$y(\hat{\mathbf{x}}^*)$	21.1	21.11	21.1	21.1	21.1	21.08	21.1	21.1	21.18	20.99	21.23	21.1	21.18	20.99	21.23	21.1	
$d_{y(\hat{\mathbf{x}}^*)}$	0	0.047370914	0	0	0	-0.09488	0	0	0.377715	-0.52406	0.612341	0	0.377715	-0.52406	0.612341	0	
$\hat{g}_1(\hat{\mathbf{x}}^*)$	4999.99	4999.99	5000	4999.93	4999.99	4999.99	5000	4999.99	5000	5000	5000	5000	5000	5000	5000	5000	5000
$d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	-0.0002	-0.0002	0	-0.00140002	-0.0002	-0.0002	0	-0.0002	0	0	0	0	0	0	0	0	0
$n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	4	3	6	20	5	4	4	20	5	4	10	20	5	3	10	20	20
$g_1(\hat{\mathbf{x}}^*)$	5000.97	4998.13	4977.73	5020.5	5040.41	5044.55	5065.71	4947.95	4957.3	4752.44	5245.42	5065.61	4957.3	4752.44	5245.42	5065.61	5065.61
$d_{g_1(\hat{\mathbf{x}}^*)}$	0.019396	-0.037413993	-0.447392687	0.408325864	0.80172	0.883131	1.297153	-1.05195	-0.86136	-5.20911	4.678748	1.295204	-0.86136	-5.20911	4.678748	1.295204	1.295204
$\hat{g}_2(\hat{\mathbf{x}}^*)$	3649.39	3632.09	3632.2	3383.49	3349.51	3305.79	3210.67	3947.6	3809.36	5850.51	5106.07	4822.51	3809.36	5850.51	5106.07	4822.51	4822.51
$d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	-0.37897	-0.857082286	-0.854027862	-8.267498943	-9.36585	-10.8122	-14.0952	7.203871	-666.861	37.38631	28.25754	24.03914	-666.861	37.38631	28.25754	24.03914	24.03914
$n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	4	2	3	20	4	2	6	20	4	2	5	20	4	2	5	20	20
$g_2(\hat{\mathbf{x}}^*)$	3676.18	3671.93	3715.41	3565.72	3685.96	3709.62	3757.43	3797.57	3809.36	3799.92	3905.57	3919.81	3809.36	3799.92	3905.57	3919.81	3919.81
$d_{g_2(\hat{\mathbf{x}}^*)}$	0.35254	0.237204958	1.404690196	-2.734370618	0.616936	1.250802	2.507299	3.537789	3.83634	3.597444	6.20524	6.545981	3.83634	3.597444	6.20524	6.545981	6.545981
$\hat{g}_3(\hat{\mathbf{x}}^*)$	-1234.84	-1229.84	-1374.8	-1559.9	-1672.81	-1637.42	-1837.39	-1238.42	-779.34	-2175.85	-5536.46	-491.67	-779.34	-2175.85	-5536.46	-491.67	-491.67
$d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	-8.39137	-8.832043193	2.64329357	14.19578178	19.98733	18.25799	27.15428	-8.07803	-71.7428	38.48565	75.82462	-172.227	-71.7428	38.48565	75.82462	-172.227	-172.227
$n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	1	1	1	20	2	2	1	20	2	2	6	20	2	2	6	20	20
$g_3(\hat{\mathbf{x}}^*)$	-1202.99	-1175.16	-1199.61	-1345.04	-1211.47	-1192.4	-1202.69	-1305.46	-716.09	-1991.22	-6183.01	-1282.71	-716.09	-1991.22	-6183.01	-1282.71	-1282.71
$d_{g_3(\hat{\mathbf{x}}^*)}$	-11.2611	-13.89598012	-11.57459508	0.489204782	-10.4823	-12.2492	-11.2889	-2.52784	-86.9123	32.78191	78.35261	-4.34627	-86.9123	32.78191	78.35261	-4.34627	-4.34627

Tabela 13 – Resultados do teste 4.

Variável	Cenário A - 15 pontos					Cenário B - 10 pontos					Cenário C - 5 pontos				
	EI	PI	LCB	s^2		EI	PI	LCB	s^2		EI	PI	LCB	s^2	
\hat{x}_1^*	6.2	6.05	6.08	6.33		6.75	6.92	6.72	6.35		7.1	6.13	5.3	5.26	
$d_{\hat{x}_1^*}$	-1.75806	-4.28099	-3.76645	0.331754	6.5333333	8.82948	8.82948	6.116071	0.645669	11.14085	-2.92007	-19.0377	-19.943		
\hat{x}_2^*	3.32	3.68	3.07	3.14	2.24	1.42	1.42	2.37	3.23	1	2.3	1	5.55		
$d_{\hat{x}_2^*}$	1.807229	11.41304	-6.18893	-3.82166	-45.5357	-129.577	-129.577	-37.5527	-0.92879	-226	-41.7391	-226	41.26126		
$\hat{y}(\hat{\mathbf{x}}^*)$	20.87	20.85	20.32	21.05	21.34	20.57	20.57	21.3	21.21	20.94	19.75	15.73	20.46		
$d_{\hat{y}(\hat{\mathbf{x}}^*)}$	-1.10206	-1.19904	-3.83858	-0.2375	1.124649	-2.57657	-2.57657	0.938967	0.518623	-0.76409	-6.83544	-34.1386	-3.12805		
$n_{\hat{y}(\hat{\mathbf{x}}^*)}$	1	2	1	0	1	3	3	1	1	1	2	1	5		
$y(\hat{\mathbf{x}}^*)$	21.1	21.19	21.13	21.1	21.1	20.93	20.93	20.98	21.1	21.29	21.42	21.34	21.1		
$d_{y(\hat{\mathbf{x}}^*)}$	0	0.424729	0.141978	0	0	-0.81223	-0.81223	-0.57197	0	0.892438	1.493931	1.124649	0		
$\hat{g}_1(\hat{\mathbf{x}}^*)$	5000	5000	5000	4999.99	4999.99	4999.99	4999.99	4999.99	4999.99	5000	4999.99	5000	4999.99		
$d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	0	0	0	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	0	-0.0002	0	-0.0002		
$n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	3	2	3	20	4	3	3	3	20	5	3	5	20		
$g_1(\hat{\mathbf{x}}^*)$	4953.43	4943.93	4867.35	4987.3	5090.59	5049.99	5049.99	5112.66	5027.38	4978.14	4818.04	4892.81	5236.11		
$d_{g_1(\hat{\mathbf{x}}^*)}$	-0.94016	-1.13412	-2.7253	-0.25465	1.779558	0.989903	0.989903	2.20355	0.544618	-0.43912	-3.77664	-2.19077	4.509264		
$\hat{g}_2(\hat{\mathbf{x}}^*)$	3615.83	3482.25	3683.57	3685.72	4060.83	4388.26	4388.26	4022.52	3680.63	4453.86	4263.98	4976.95	3011.532		
$d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	-1.31063	-5.19693	0.552453	0.610464	9.791348	16.52227	16.52227	8.932212	0.473017	17.75179	14.08918	26.39629	-21.6397		
$n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	5	3	4	20	3	1	1	3	20	5	4	3	20		
$g_2(\hat{\mathbf{x}}^*)$	3611.44	3591.2	3544.42	3638.68	3731.54	3757.55	3757.55	3755.67	3682.83	3794.14	3817.21	3835.03	3608.72		
$d_{g_2(\hat{\mathbf{x}}^*)}$	-1.43378	-2.00546	-3.35175	-0.67442	1.830879	2.510412	2.510412	2.461611	0.532471	3.450584	4.034098	4.480017	-1.51023		
$\hat{g}_3(\hat{\mathbf{x}}^*)$	-1731.02	-1976.4	-2694.61	-1374.47	-817.64	-1380.45	-1380.45	-564.88	-1213.7	-671.98	-2943.24	-3652.87	-2396.52		
$d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	22.67796	32.27788	50.32825	2.619919	-63.698	3.041762	3.041762	-136.946	-10.2797	-99.1815	54.52427	63.35867	44.14985		
$n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	3	2	5	20	2	3	3	1	20	11	4	4	20		
$g_3(\hat{\mathbf{x}}^*)$	-1650.61	-1751.32	-2275.35	-1407.66	-909.14	-1977.41	-1977.41	-947.23	-1235.06	-637.09	-2468.56	-4543.5	-1174.2		
$d_{g_3(\hat{\mathbf{x}}^*)}$	18.91119	23.57422	41.17564	4.91596	-47.2226	32.31247	32.31247	-41.3025	-8.37206	-110.09	45.77973	70.54121	-13.9891		

Tabela 14 – Resultados do teste 5.

Variável	Cenário A - 15 pontos					Cenário B - 10 pontos					Cenário C - 5 pontos				
	EI	PI	LCB	s^2		EI	PI	LCB	s^2		EI	PI	LCB	s^2	
\hat{x}_1^*	6.32	6.19	6.29	6.33		6.56	6.88	6.55	5.97		6.96	6.82	6.89	6.29	
$d_{\hat{x}_1^*}$	0.174051	-1.92246	-0.30207	0.331754		3.82622	8.299419	3.679389	-5.67839		9.353448	7.492669	8.432511	-0.30207	
\hat{x}_2^*	3.1	3.57	3.26	3.1		2.52	1.24	2.33	4.16		1	1	1	3.32	
$d_{\hat{x}_2^*}$	-5.16129	8.683473	0	-4.15335		-29.3651	-162.903	-39.9142	21.63462		-226	-226	-226	1.807229	
$\hat{y}(\hat{\mathbf{x}}^*)$	20.98	21.12	21.1	21.04		21.23	20.87	20.86	21.07		21.09	20.04	20.19	21.12	
$d_{\hat{y}(\hat{\mathbf{x}}^*)}$	-0.57197	0.094697	0	-0.28517		0.612341	-1.10206	-1.15053	-0.14238		-0.04742	-5.28942	-4.50718	0.094697	
$n_{\hat{y}(\hat{\mathbf{x}}^*)}$	1	3	1	0		1	3	1	0		1	1	1	1	
$y(\hat{\mathbf{x}}^*)$	21.1	21.12	21.13	21.1		21.09	21.16	21.08	21.1		21.39	21.42	21.44	21.1	
$d_{y(\hat{\mathbf{x}}^*)}$	0	0.094697	0.141978	0		-0.04742	0.283554	-0.09488	0		1.355774	1.493931	1.585821	0	
$\hat{g}_1(\hat{\mathbf{x}}^*)$	4999.99	4999.99	4999.99	5000		4999.99	5000	5000	4999.96		4999.99	5000	5000	5000	
$d_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	-0.0002	-0.0002	-0.0002	0		-0.0002	0	0	-0.0008		-0.002	0	0	4.112974	
$n_{\hat{g}_1(\hat{\mathbf{x}}^*)}$	4	2	5	2		4	2	4	20		5	4	6	20	
$g_1(\hat{\mathbf{x}}^*)$	4971.38	4993.66	4993.12	4986.13		5013.05	4984.5	4979.95	5005.99		4954.82	4910.6	5004.5	5004.26	
$d_{g_1(\hat{\mathbf{x}}^*)}$	-0.5757	-0.12696	-0.13779	-0.27817		0.260321	-0.31096	-0.40261	0.119657		-0.91184	-1.82055	0.089919	0.085127	
$\hat{g}_2(\hat{\mathbf{x}}^*)$	3669.13	3549.07	3547.83	3692.07		3776.46	4692.69	3582.62	3200.03		3962.34	4426.27	4392.27	3820.35	
$d_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	-0.0002	-3.21634	-3.25241	0.781404		2.998575	21.93774	11.42785	-14.474		7.549075	17.23912	16.59848	4.112974	
$n_{\hat{g}_2(\hat{\mathbf{x}}^*)}$	4	3	3	20		5	2	3	20		5	3	4	20	
$g_2(\hat{\mathbf{x}}^*)$	3598.27	3549.07	3544.42	3638.62		3631.92	3621.75	3582.62	3625		3877.36	4100.82	4150.49	3865.34	
$d_{g_2(\hat{\mathbf{x}}^*)}$	-1.80503	-3.21634	-3.35175	-0.67608		-0.8618	-1.14503	-2.24975	-1.05434		5.52283	10.67104	11.74006	5.229035	
$\hat{g}_3(\hat{\mathbf{x}}^*)$	-1163.76	-1164.6	-1222.87	-1364.89		-1519.55	-684.76	-938	-1700.66		-1638.75	-3673.25	-3865.77	-1476.88	
$d_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	-15.0117	-14.9287	-9.45235	1.93642		11.91734	-95.4641	-42.693	21.29761		18.32433	63.56197	65.37663	9.372461	
$n_{\hat{g}_3(\hat{\mathbf{x}}^*)}$	1	1	1	20		4	2	1	20		9	3	3	20	
$g_3(\hat{\mathbf{x}}^*)$	-1212.17	-1164.6	-1215.99	-1402.74		-1581.86	-1110.69	-1201.43	-1362.75		-503.66	-4611.45	-5367.4	-1457.24	
$d_{g_3(\hat{\mathbf{x}}^*)}$	-10.4185	-14.9287	-10.0716	4.58246		15.38695	-20.5071	-11.4056	1.782425		-165.747	70.97529	75.06316	8.151025	