



DISSERTAÇÃO DE MESTRADO

**RETROFITTING DO ROBÔ ASEA IRB6-S2 BASEADO EM
TECNOLOGIAS DE COMANDO NUMÉRICO USANDO LINUXCNC**

JUAN SEBASTIAN TOQUICA ARENAS

Brasília, Dezembro de 2016

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA MECÂNICA**

**RETROFITTING DO ROBÔ ASEA IRB6-S2 BASEADO EM
TECNOLOGIAS DE COMANDO NUMÉRICO USANDO LINUXCNC**

JUAN SEBASTIAN TOQUICA ARENAS

**DISSERTAÇÃO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA MECÂNICA DA FACULDADE DE TECNOLOGIA DA
UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM
SISTEMAS MECATRÔNICOS**

APROVADA POR:

Prof. Dr. Alberto J. Alvares, PPMEC/UnB
Orientador

Prof. Dr. André Murilo de Almeida, Gama/UnB
Membro Interno

Prof. Dr. Renan Bonnard, MENESR/França
Membro Externo

BRASÍLIA/DF, 07 DEZEMBRO DE 2016

Toquica Arenas, Juan Sebastian

Retrofitting do robô ASEA IRB6-S2 baseado em tecnologias de comando numérico usando LinuxCNC / JUAN SEBASTIAN TOQUICA ARENAS. –Brasil, 2016.
170 p.

Orientador: Alberto José Alvares

Dissertação (Mestrado) – Universidade de Brasília – UnB

Faculdade de Tecnologia – FT

Programa de Pós-Graduação em Sistemas Mecatrônicos – PPMEC, 2016.

1. Retrofitting. 2. Robô Asea IRB6-S2. 3. Cinemática. 4. LinuxCNC. 5. Arquitetura aberta de controle. I. Alberto José Alvares, orientador. II. Universidade de Brasília. III. Faculdade de Tecnologia.

Agradecimentos

Em primeiro lugar a Deus, por me permitir a oportunidade de estar com vida e honrar seu nome por meio de minhas ações diárias.

Ao meu filho sendo uma infinita motivação diária, sempre confiante em romper os meus limites imaginários com o intuito de atingir meus objetivos sem receios. "A engenharia é como a vida", digo ao meu filho, "é uma arte especial que permite descobrir caminhos nunca antes percorridos, com convicção de que sempre existe mais alguma coisa por fazer. Sendo assim, querido Juan José sempre estará em capacidade de superar seus próprios limites, suas próprias fronteiras geográficas e culturais, seus sonhos, amo a você na distancia"

Ao meu maravilhoso núcleo familiar: minha mãe Clara Inés, meu Pai Efrain e minha irmã Diana Andrea que sempre estiveram e estão ao meu lado, sem se importarem com as condições externas, me contribuindo continuamente ao fortalecimento, mesmo distante, dos laços familiares. Saúdo também à todos os membros de minha família que me apoiaram em meus sonhos.

Agradeço carinhosamente à eterna professora de português, que sem ela não haveria condições de obter a realização desta etapa de minha vida. Sou muito grato à ela ter sempre acreditado em meu potencial, até quando eu duvidava de mim mesmo. Obrigado senhorita Carvalho Bonifácio por estar neste caminho de aprendizagem constante ao meu lado, sempre me protegendo, me contagiando com essa energia de esperança e felicidade.

Aos colegas e amigos Diego Benavides e Luiz Eduardo, que ao longo desta decisão acadêmica nos tornamos maduros com as experiências de vida. Condições estas que nos são necessárias para conquistarmos qualquer meta e cumprir quaisquer sonhos. Reconhecendo a importância da disciplina e paixão.

A todas as pessoas que de alguma maneira, direta e indiretamente, contribuíram para que este sonho pudesse hoje estar concluído, através de uma palavra, um abraço, uma caminhada noturna. Detalhes que marcam as vidas de pessoas que convivem na mesma realidade.

Por último, mas não menos importante, ao professor Alberto Alvares pela oportunidade de levar este projeto até o final, pela paciência quando precisei recuperar o foco na minha estadia no Brasil. O mais importante foi ter aprendido com sua sabedoria que os objetivos só são atingidos com disciplina, sacrifício e objetividade.

Ao CNPq pela bolsa de mestrado que me permitiu manter ao longo dos 24 meses de dedicação exclusiva. Aportando dinamicamente um grão de areia da literatura relacionada com minha paixão, a robótica. Ao PPMEC e especialmente ao professor Edson por me ter dado a oportunidade de fazer concretizar um ideal de vida.

JUAN SEBASTIAN TOQUICA ARENAS

RESUMO

Resumo.

ABSTRACT

Abstract.

SUMÁRIO

RESUMO	i
ABSTRACT	ii
LISTA DE FIGURAS	iv
LISTA DE TABELAS	v
LISTA DE SÍMBOLOS	vii
LISTA DE ABREVIATURAS E ACROGRAMAS	viii
1 Introdução	1
1.1 Contextualização	1
1.2 Definição do Problema.....	1
1.3 Objetivos da Dissertação	1
1.3.1 Objetivo Geral	1
1.3.2 Objetivos Específicos.....	1
1.4 Apresentação do Documento	1
2 Fundamentação Teórica	2
2.1 Introdução.....	2
2.2 MÃl'todos Convencionais de RecuperaÃġÃċo SecundÃqria	2
2.2.1 Conceito e ContextualizaÃġÃċo da RecuperaÃġÃċo SecundÃqria	2
2.2.2 ClassificaÃġÃċo dos MÃl'todos de RecuperaÃġÃċo SecundÃqria	5
2.2.3 MÃl'todos Convencionais.....	5
2.2.4 EficiÃłncia de RecuperaÃġÃċo	5
2.2.5 Aspectos Operacionais da InjeÃġÃċo de Ãqgua	5
2.3 SimulaÃġÃċo NumÃl'rica de ReservatÃşrios	5
2.4 Conceitos de Otimização	5
2.4.1 Definições e Fatos Básicos.....	5
2.4.2 Diferenciabilidade de Funções Escalares.....	8
2.4.3 Diferenciabilidade de Funções Multivariáveis e Campos Vetoriais	9
2.4.4 Convexidade	12
2.4.5 Condições de Otimalidade	14

2.4.6	Principais Algoritmos de Otimização	19
2.4.7	Tratamento de Problemas de Otimização	23
3	Proposta de Metodologia para o <i>Retrofitting</i> de Robôs Industriais	25
3.1	Introdução.....	25
4	Retrofitting do Robô ASEA IRB6-S2	26
4.1	Introdução.....	26
5	Conclusões	27
5.1	Conclusões do Trabalho	27
5.2	Sugestões para Trabalhos Futuros.....	27
	REFERÊNCIAS BIBLIOGRÁFICAS.....	28
	APÊNDICES	29

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SÍMBOLOS

Símbolos Latinos

\mathbb{P}	Problema de Otimização não-linear
\mathbb{R}	Conjunto dos números reais
\mathbb{R}^n	Espaço euclidiano de dimensão n
$\mathbb{R}^{n \times n}$	Espaço das matrizes quadradas de ordem n sobre os números reais
\mathbb{Z}_+	Conjunto dos números inteiros não negativos

Símbolos Gregos

∇	Operador gradiente ou primeira derivada
∇^2	Segunda derivada

Sobrescritos

\cdot	Variação temporal
$-$	Valor ótimo

Notações

$\mathcal{F}(\cdot)$	Conjunto viável (factível)
$H(\cdot)$	Matriz Hessiana
$\mathcal{I}(\cdot)$	Conjunto de Índices
$J(\cdot)$	Matriz Jacobiana
$tr[\cdot]$	Traço de uma matriz
\mathbf{X}	Representação de matriz
\mathbf{x}	Representação de variável em \mathbb{R}^n
$\mathbf{x}^{(n)}$	Iteração de variável

LISTA DE ABREVIATURAS E ACROGRAMAS

BFGS	<i>Broydon-Fletcher-Goldfarb-Shanno</i>
FJ	<i>Fritz-John</i>
KKT	<i>Karush-Kuhn-Tucker</i>
MATLAB	<i>Matrix Laboratory</i>
MRST	<i>MATLAB Reservoir Simulation Toolbox</i>
NPV	<i>Net-Present Value</i>
PSO	<i>Particle Swarm Optimization</i>
SQP	<i>Sequential Quadratic Programming</i>

Capítulo 1

Introdução

1.1 Contextualização

1.2 Definição do Problema

1.3 Objetivos da Dissertação

1.3.1 Objetivo Geral

1.3.2 Objetivos Específicos

1.4 Apresentação do Documento

Capítulo 2

Fundamentação Teórica

2.1 Introdução

Neste capítulo é apresentada a revisão bibliográfica

2.2 Métodos Convencionais de Recuperação Secundária

2.2.1 Conceito e Contextualização da Recuperação Secundária

De acordo com Rosa *et al.*, nas acumulações de petróleo há, na época de sua descoberta, uma dada quantidade de energia, chamada de *energia primária*, cuja grandeza é determinada pelo volume e pela natureza dos fluidos existentes no meio, além dos níveis de pressão e temperatura do reservatório. Quando se dá o processo de produção, parte dessa energia é dissipada por causa da descompressão dos fluidos do reservatório e das resistências que os mesmos encontram ao fluir em direção aos poços produtores — resistências associadas às forças viscosas e capilares presentes no meio poroso. A consequência dessa dissipação de energia primária resulta no decréscimo de pressão do reservatório em sua vida produtiva e, consequentemente, na redução da produtividade dos poços. A quantidade de óleo retirada utilizando-se unicamente a energia do reservatório é denominada *recuperação primária*.

De forma a se minorar os efeitos danosos da dissipação da energia primária, existem duas linhas de ação a serem consideradas:

- Reduzir as resistências viscosas e/ou capilares por meio de métodos especiais, como por exemplo aquecendo a jazida;
- Adicionar suplemento de energia secundária, artificialmente comunicada, através de injeção de fluidos em poços selecionados.

Quando se suplementa o reservatório com energia transferida artificialmente, ou se empregam meios de incrementar a eficiência da energia primária, a quantidade adicional de óleo produzida é chamada

de *recupera   o secund ria*. Por extens o, todas as opera  es que conduzem   obten  o desse adicional de  leo tamb m s o denominadas *recupera  o secund ria*. Essas opera  es, atualmente, s o implantadas em sua grande maioria o t o cedo quanto poss vel na vida do reservat rio.

  importante ressaltar que h  uma diferen a entre *recupera  o secund ria* e m todos de eleva  o artificial e de estimula  o de po os; estes n o afetam diretamente as energias expulsivas do reservat rio, embora sua aplica  o concorra para economiz -las. As t cnicas de eleva  o artificial e de estimula  o de po os est o mais ligadas ao comportamento dos po os produtores do que ao comportamento do reservat rio como um todo. Contudo, a linha divis ria entre tais m todos e os m todos de *recupera  o secund ria* n o   muito n tida — certos m todos de estimula  o, como a inje  o c lica de vapor, s o usualmente inclu dos entre os m todos de *recupera  o secund ria* (ROSA; CARVALHO; XAVIER, 2006).

Ainda segundo Rosa *et al.*, h  dois objetivos pr ticos b sicos dos m todos de *recupera  o secund ria*:

- **Aumento da efici ncia de recupera  o** — A efici ncia de *recupera  o prim ria*   normalmente baixa; em alguns casos, dependendo das caracter sticas do reservat rio e dos fluidos, ela pode ser at  nula. Em alguns casos, a efici ncia de *recupera  o secund ria* pode passar de 60% em casos bem-sucedidos; contudo, o valor mais frequente dessa efici ncia, nos m todos convencionais, se situa entre 30 e 50%.
- **Acelera  o da produ  o** — O emprego dos m todos de *recupera  o secund ria* busca acelerar a produ  o ou ao menos reduzir a taxa de seu decl nio natural. A *acelera  o* da produ  o resulta em antecipa  o do fluxo de caixa; portanto, h  o aumento de seu valor presente e uma consequente melhoria da economicidade da explora  o do campo ou reservat rio.

Al m dos objetivos b sicos de emprego da *recupera  o secund ria*, Rosa *et al.* citam v rios incentivos ao uso desses m todos, tais como: pre o do petr leo; custos de explora  o, desenvolvimento e produ  o; e avan os tecnol gicos na  rea. Por m, destaca-se que apenas o uso dessas t cnicas n o   suficiente para mitigar todos os males da produ  o de petr leo e do esgotamento das reservas; outras medidas podem e devem ser tomadas, simultaneamente, para aumentar a efici ncia e a rentabilidade da produ  o, tais como:

- **Explora  o de reservas n o convencionais** — Xistos e folhelhos betuminosos, por exemplo, acumulam grandes quantidades de  leo. V rias dessas reservas j  foram encontradas em regi es como Athabasca, no Canad , cintur o do Orinoco, na Venezuela, e o Colorado, nos Estados Unidos. O custo de produ  o nessas reservas   consider vel, mas j  se projetam meios tecnol gicos para reduzir o mesmo. Entre outras reservas n o convencionais de hidrocarbonetos, h  a presen a de g s natural em solu  o existente na  gua de aqu feros; embora a raz o de solubilidade do g s natural na  gua normalmente seja pequena, o imenso volume dos aqu feros permitiria uma produ  o de grandes volumes desse g s. Uma outra

reserva não convencional poderá ser o gás natural proveniente de hidratos localizados no fundo de oceanos e em regiões congeladas da Terra.

- **Estimulação de Poços** De acordo com Thomas *et al.*, a estimulação de poços é um conjunto de atividades realizadas com o objetivo de aumentar o índice de produtividade ou injetividade do poço (THOMAS *et al.*, 2004, p. 166). Os principais métodos de estimulação são: fraturamento hidráulico, em que se cria, através de uma ruptura na rocha-reservatório causada por um elevado gradiente de pressão, um caminho preferencial de alta condutividade, facilitando um fluxo de fluidos do reservatório ao poço (ou vice-versa); e acidificação, onde se injeta um ácido com pressão inferior à pressão de fraturamento da formação, visando remover danos da mesma. Tais métodos contribuem para a aceleração da produção e até, em alguns casos, o aumento da eficiência de recuperação. A aplicação de métodos de estimulação pode, inclusive, ser feita em campos submetidos a operações de recuperação secundária.
- **Uso de poços especiais** — Nas últimas décadas houve um incremento considerável no uso dos chamados *poços especiais*, que possuem como característica marcante a não-verticalidade. Segundo Thomas *et al.*, esses poços são perfurados com várias finalidades, como: controlar um poço em *blowout* por meio de poços de alívio; atingir formações produtoras abaixo de locais inacessíveis, como rios, lagos, cidades, entre outros; desviar a trajetória do poço de acidentes geológicos, como domos salinos e falhas; perfurar vários poços de um mesmo ponto, como é o caso da produção em plataformas marítimas; e desviar poços que tiveram seu trecho final perdido por problemas operacionais (THOMAS *et al.*, 2004, p. 106). O uso desses poços inclinados, horizontais, multilaterais, etc., pode aumentar a velocidade de drenagem do reservatório, ou seja, antecipar a produção, bem como aumentar a eficiência de recuperação através do aumento da eficiência de varrido, por exemplo.
- **Extração de líquidos de gás natural** — A produção de hidrocarbonetos líquidos pode ser aumentada pela instalação de plantas de gasolina natural e de unidades portáteis de extração de líquidos de gás natural.
- **Reestudo de áreas julgadas improdutivas ou antieconômicas** — Mesmo que as reservas mundiais de petróleo sejam limitadas, elas estão longe de terem sido totalmente exploradas; de fato, apenas uma pequena porcentagem da superfície do planeta foi inteiramente explorada. Seja na terra ou no fundo do mar, há ainda perspectivas notáveis fora das áreas hoje em produção; além disso, as estimativas do volume de petróleo que ainda poderá ser descoberto são ainda vagas. É com essa perspectiva que a indústria pode medir as oportunidades que tem à frente no caso de esgotamento das áreas hoje em produção. Portanto, de um modo geral, deve-se pensar sempre na adoção das seguintes medidas, sem danos ao andamento das operações de recuperação secundária: estudar novas áreas; estudar formações mais profundas (o principal exemplo); reestudar áreas consideradas esgotadas ou de produção antieconômica; e investir mais dinheiro, tempo e pessoal em treinamento e pesquisa, visando melhorar os métodos de exploração e produção existentes.

2.2.2 Classificação dos Métodos de Recuperação Secundária

2.2.3 Métodos Convencionais

2.2.4 Eficiência de Recuperação

2.2.5 Aspectos Operacionais da Injeção de Água

2.3 Simulação Numérica de Reservatórios

2.4 Conceitos de Otimização

2.4.1 Definições e Fatos Básicos

Antes de se proceder à análise do problema estudado, fazem-se necessários alguns conceitos básicos da área de otimização e álgebra linear. A presente seção apresenta algumas definições que serão importantes ao percorrer da análise de convexidade e de problemas de otimização. Primeiramente, são dadas algumas definições sobre vetores e matrizes, conforme Aguirre (AGUIRRE, 2015):

Definição 2.4.1.1 Dadas duas variáveis $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^n$, o produto interno entre \mathbf{x} e \mathbf{y} é dado por $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$. Caso este produto seja nulo, os vetores são ditos ortogonais.

Definição 2.4.1.2 Uma matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ é dita semidefinida positiva se $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0, \forall \mathbf{x} \neq \mathbf{0}, \mathbf{x} \in \mathbb{R}^n$. No caso de desigualdade estrita, a matriz \mathbf{A} é dita definida positiva.

Definição 2.4.1.3 Uma matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ é dita semidefinida negativa se $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0, \forall \mathbf{x} \neq \mathbf{0}, \mathbf{x} \in \mathbb{R}^n$. No caso de desigualdade estrita, a matriz \mathbf{A} é dita definida negativa.

Definição 2.4.1.4 Uma matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ é dita indefinida se ela não for semidefinida positiva nem semidefinida negativa.

Uma última definição básica de álgebra linear se refere ao conceito de normas de vetores e matrizes, de acordo com Yang (YANG, 2010):

Definição 2.4.1.5 Seja $\mathbf{v} \in \mathbb{R}^n$. A p -norma ou ℓ_p -norma de \mathbf{v} é dada por

$$\|\mathbf{v}\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{Z}_+ - \{0\}. \quad (2.4.1)$$

Algumas propriedades elementares da norma vetorial devem ser satisfeitas para todo valor de p , entre as quais:

- (a) $\|\mathbf{v}\| \geq 0, \forall \mathbf{v}$;
- (b) $\|\mathbf{v}\| = 0 \Rightarrow \mathbf{v} = \mathbf{0}$;

(c) $\|\alpha \mathbf{v}\| = |\alpha| \|\mathbf{v}\|, \forall \alpha \in \mathbb{R};$

(d) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (Desigualdade triangular).

Algumas normas vetoriais comuns são calculadas tomando-se (2.4.1) com $p = 1$ e $p = 2$, sendo a 2-norma de \mathbf{v} também conhecida como *norma euclidiana*, ou *comprimento* de \mathbf{v} . Neste caso, a norma de \mathbf{v} também pode ser escrita como

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^T \mathbf{v}}. \quad (2.4.2)$$

Um caso especial de norma vetorial ocorre quando $p = \infty$; a ℓ_∞ -norma, ou *norma de Chebyshev* de \mathbf{v} , é dada por

$$\|\mathbf{v}\|_p = v_{max} = \max_{1 \leq i \leq n} |v_i|. \quad (2.4.3)$$

Definição 2.4.1.6 Dada uma matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ qualquer, a sua p -norma, analogamente à Definição 2.4.1.5, pode ser escrita como

$$\|\mathbf{A}\|_p = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{Z}_+ - \{0\}. \quad (2.4.4)$$

Um caso especial da norma matricial ocorre quando se aplica a Equação (2.4.4) com $p = 2$; tal norma é denominada *norma de Frobenius* de \mathbf{A} . Outras normas matriciais populares são tidas como a máxima soma de linhas ou de colunas da matriz ($\|\mathbf{A}\|_1$ e $\|\mathbf{A}\|_\infty$, respectivamente). Deve-se destacar também que todas as propriedades básicas satisfeitas para a norma vetorial devem ser verdadeiras também para a norma matricial.

Outros conceitos básicos a serem apresentados são pertencentes à área de otimização. Izmailov e Solodov (IZMAILOV; SOLODOV, 2005) trazem o conceito elementar de problema de otimização e mínimo (ou máximo) de uma função:

Definição 2.4.1.7 Sejam um conjunto $D \subset \mathbb{R}^n$ e uma função $f : D \rightarrow \mathbb{R}$. O problema de se encontrar um minimizador de f em D é escrito como

$$\min f(\mathbf{x}) \text{ sujeito a } \mathbf{x} \in D. \quad (2.4.5)$$

A função f é denominada *função objetivo*; o conjunto D é o conjunto viável do problema, ou conjunto de restrições — seus pontos serão chamados de pontos viáveis. Normalmente, o conjunto de restrições D pode ser definido como

$$D = \{\mathbf{x} \in \Omega \mid h(\mathbf{x}) = 0, g(\mathbf{x}) < 0\} \quad (2.4.6)$$

em que Ω é o conjunto de restrições diretas do problema ($\Omega \subset \mathbb{R}^n$), $h : \Omega \rightarrow \mathbb{R}^l$ e $g : \Omega \rightarrow \mathbb{R}^m$ são as $l \in \mathbb{Z}_+$ restrições de igualdade e $m \in \mathbb{Z}_+$ restrições de desigualdade (também denominadas restrições funcionais). Caso $D = \mathbb{R}^n$, o problema é dito de otimização irrestrita; em caso contrário, se trata de problema com restrições.

Definição 2.4.1.8 Um problema de maximização pode ser escrito como

$$\max f(\mathbf{x}) \text{ sujeito a } \mathbf{x} \in D. \quad (2.4.7)$$

Nota-se que o problema (2.4.7) pode ser reescrito como um problema de minimização equivalente:

$$\min -f(\mathbf{x}) \text{ sujeito a } \mathbf{x} \in D.$$

Visto que resolver um problema de maximização não exige técnicas substancialmente diferentes de um problema de minimização, uma vez que um pode ser reescrito como o outro, serão considerados a partir dessa seção problemas de minimização.

Antes de se prosseguir com os conceitos de minimizador e valor ótimo, são necessários os conceitos de supremo, ínfimo, máximos e mínimos de um conjunto. A definição a seguir é encontrada em Yang (YANG, 2010):

Definição 2.4.1.9 Dado um conjunto $S \in \mathbb{R}$, o número u é denominado limite superior de S se $u \geq x$, $\forall x \in S$. Por consequência, o número β é denominado supremo de S se β é o menor dos limites superiores u de S ($\beta \leq u$, $\forall u$). O supremo de S pode ser denotado por

$$\beta \equiv \sup x \equiv \sup S \equiv \sup(S). \quad (2.4.8)$$

Caso $\beta \in S$, pode-se dizer que β é o valor máximo de S , ou seja,

$$\beta \equiv \max S \equiv \max(S). \quad (2.4.9)$$

De maneira análoga, o número l é denominado limite inferior de S se $l \leq x$, $\forall x \in S$. Por consequência, o número α é denominado ínfimo de S se α é o maior dos limites inferiores l de S ($\alpha \geq l$, $\forall l$). O ínfimo de S pode ser denotado por

$$\alpha \equiv \inf x \equiv \inf S \equiv \inf(S). \quad (2.4.10)$$

Caso $\alpha \in S$, pode-se dizer que α é o valor mínimo de S , ou seja,

$$\alpha \equiv \min S \equiv \min(S). \quad (2.4.11)$$

Algumas propriedades básicas sobre ínfimos e supremos são apresentadas a seguir, de acordo com Yang (YANG, 2010):

$$\inf Q = -\sup(-Q), \quad (2.4.12)$$

$$\sup_{p \in P, q \in Q} (p + q) = \sup(P) + \sup(Q), \quad (2.4.13)$$

$$\sup_{x \in S} (f(x) + g(x)) \leq \sup_{x \in S} (f(x)) + \sup_{x \in S} (g(x)). \quad (2.4.14)$$

Vale notar que os conceitos de supremo e ínfimo apresentados não estendem a conjuntos não-limitados;

por exemplo, seja o conjunto $S = [2, +\infty)$. Verifica-se que, de acordo com a Definição 2.4.1.9, $\inf S = 2$, mas o supremo não existe, ou seja, $\sup S \rightarrow +\infty$. Por este exemplo, conclui-se que ínfimos ou supremos podem não existir em conjuntos não limitados; portanto, para contornar este fato, faz-se a definição de uma extensão dos números reais (YANG, 2010):

$$\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}. \quad (2.4.15)$$

Considerando-se a Equação (2.4.15) e que $\sup(\emptyset) = -\infty$, para qualquer subconjunto de $\bar{\mathbb{R}}$, o supremo e ínfimo sempre existirão, supondo-se $\sup \mathbb{R} = +\infty$ e $\inf \mathbb{R} = -\infty$.

De posse das definições até aqui dadas, é possível definir os conceitos de minimizador e valor ótimo de uma função. As definições a seguir são dadas por Izmailov e Solodov (IZMAILOV; SOLODOV, 2005):

Definição 2.4.1.10 *Dado o problema (2.4.5), diz-se que um ponto $\bar{\mathbf{x}} \in D$ é*

(a) *minimizador global de (2.4.5), se*

$$f(\bar{\mathbf{x}}) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in D; \quad (2.4.16)$$

(b) *minimizador local de (2.4.5), se existe uma vizinhança U de $\bar{\mathbf{x}}$ tal que*

$$f(\bar{\mathbf{x}}) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in D \cap U, \quad (2.4.17)$$

ou, de forma análoga,

$$\exists \epsilon > 0 \bullet f(\bar{\mathbf{x}}) \leq f(\mathbf{x}) \quad \forall \mathbf{x} \in \{\mathbf{x} \in D \mid \|\mathbf{x} - \bar{\mathbf{x}}\| \leq \epsilon\}. \quad (2.4.18)$$

Observa-se que todo minimizador global é também local, mas não de forma recíproca. Se, para $\mathbf{x} \neq \bar{\mathbf{x}}$, a desigualdade (2.4.16) ou (2.4.17) é estrita, $\bar{\mathbf{x}}$ é chamado de *minimizador estrito* (global ou local, respectivamente).

Definição 2.4.1.11 *Seja $\bar{\mathbf{v}} \in [-\infty, +\infty)$ definido por*

$$\bar{\mathbf{v}} = \inf_{\mathbf{x} \in D} f(\mathbf{x}); \quad (2.4.19)$$

neste caso, $\bar{\mathbf{v}}$ é denominado valor ótimo do problema (2.4.5).

2.4.2 Diferenciabilidade de Funções Escalares

A maioria dos algoritmos de otimização e conceitos associados empregam conceitos de diferenciação de funções. Esses conceitos aparecem principalmente no estudo de condições de otimalidade, visto que a análise das derivadas de uma função em um ponto pode, por exemplo, confirmar ou não sua natureza como minimizador (ou maximizador) de uma função.

Inicialmente, serão apresentados conceitos de continuidade e diferenciação básicos para uma função escalar (monovariável), isto é, $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}$.

Definição 2.4.2.1 ¹ Uma função escalar $f(x)$ é considerada contínua em um ponto $c \in \mathbb{R}$ se:

- (a) f é definida em c ($\exists f(c)$);
- (b) $\exists \lim_{x \rightarrow c} f(x)$;
- (c) $\lim_{x \rightarrow c} f(x) = f(c)$.

Diz-se que $f(x)$ é contínua se ela satisfaz as propriedades da Definição 2.4.2.1 em todos os pontos de seu domínio. Caso ela seja contínua em intervalos separados do seu domínio, diz-se que $f(x)$ é *contínua por partes*.

Definição 2.4.2.2 ² Dizemos que a derivada de $f(x)$ escalar em relação a x é equivalente ao limite

$$\frac{d}{dx} f(x) \equiv \frac{df}{dx} \equiv f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}, \quad (2.4.20)$$

desde que tal limite exista.

Caso f possua derivada no ponto x , diz-se que f é *diferenciável* em x ; se f possui derivada em todos os pontos de seu domínio, diz-se simplesmente que f é diferenciável.

Vale ressaltar que a derivada possui algumas propriedades elementares, mostradas a seguir (será utilizada simplesmente a notação f' como redução de $f'(x)$):

- (a) $(f \pm g)' = f' \pm g'$;
- (b) $(fg)' = fg' + f'g$;
- (c) $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$;
- (d) (Regra da Cadeia) $(f \circ g)' \equiv f'(g) = f'(g)g' \equiv (f' \circ g)g'$ ³

Em cursos introdutórios de cálculo, geralmente são apresentados conceitos relativos a testes de sinais de derivadas de uma função $f(x)$ qualquer de maneira a se encontrar pontos especiais, os *pontos críticos* ou *estacionários* e sua natureza. Será visto nas Seções seguintes que esse estudo no fundo é uma aplicação específica das condições de otimalidade em problemas sem restrição; a seguir, os conceitos de continuidade e diferenciação serão estendidos de forma a abranger funções de várias variáveis e também funções multiobjetivo, ou campos vetoriais.

2.4.3 Diferenciabilidade de Funções Multivariáveis e Campos Vetoriais

Antes de se proceder à análise de funções que possuem domínio ou contradomínio (ou ambos) multidimensionais, será feita uma diferenciação conceitual entre funções multivariáveis e campos vetoriais (funções multiobjetivo):

¹Ver (THOMAS, 2009a), p. 120

²*Ibid.*, p. 145

³Uma prova dessa regra se encontra em (THOMAS, 2009a), p. 246

Definição 2.4.3.1 Uma aplicação $f(\mathbf{x}) : A \rightarrow B$, $A \subset \mathbb{R}^n$ é denominada *função multivariável* se seu contradomínio é subconjunto dos números reais, isto é, $B \subset \mathbb{R}$.

Definição 2.4.3.2 Uma aplicação $\mathbf{F}(\mathbf{x}) : A \rightarrow B$, $A \subset \mathbb{R}^n$ é denominada *campo vetorial* (ou *função multiobjetivo*) se seu contradomínio é subconjunto de um espaço euclidiano de dimensão $m > 1$, isto é, $B \subset \mathbb{R}^m$.

O conceito de continuidade, nesses casos, pode ser visto como uma extensão da Definição 2.4.2.1⁴; porém, o conceito de diferenciabilidade deve ser largamente estendido, pois agora, ao contrário do caso escalar, o número de direções possíveis é infinito. Em casos em que essas direções correspondem aos vetores da base canônica do espaço do domínio ($\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ com o elemento não-nulo na i -ésima posição), as derivadas de uma função vetorial nessas direções são conhecidas como *derivadas parciais*.

Definição 2.4.3.3 ⁵ Dizemos que a derivada parcial de $f(\mathbf{x})$ em relação à componente x_i é equivalente ao limite

$$\frac{\partial}{\partial x_i} f(\mathbf{x}) \equiv \frac{\partial f}{\partial x_i} \equiv f_{x_i}(\mathbf{x}) = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(\mathbf{x})}{h}, \quad (2.4.21)$$

desde que tal limite exista.

Quando todas as derivadas parciais de uma função $f(\mathbf{x})$ existem, o *gradiente* de f é dado, por definição, como

$$\nabla f := \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \quad (2.4.22)$$

A principal diferença entre funções com domínio multidimensional e as escalares reside no fato de que o número de direções possíveis não é finito; como as derivadas parciais apenas consideram as direções da base do espaço, um novo conceito de derivada é necessário. Para tanto, seja o vetor $\mathbf{d} \in \mathbb{R}^n$ definido como

$$\mathbf{d} = (d_1, d_2, \dots, d_n) = \sum_{i=1}^n d_i \mathbf{e}_i$$

com \mathbf{e}_i sendo o i -ésimo vetor de coordenadas, conceito já discutido anteriormente. Esse vetor \mathbf{d} pode ser entendido como uma *direção* em \mathbb{R}^n . A definição a seguir de derivada direcional se encontra em Guller (GULLER, 2010):

Definição 2.4.3.4 A derivada direcional de f em um ponto \mathbf{x} de seu domínio na direção $\mathbf{d} \in \mathbb{R}^n$ é dada por

$$f'(\mathbf{x}; \mathbf{d}) = \lim_{t \searrow 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}, \quad (2.4.23)$$

desde que o limite exista à medida que $t \geq 0$ se aproxima de 0.

Guller ainda destaca que, além do fato de que $f'(\mathbf{x}; \alpha\mathbf{d}) = \alpha f'(\mathbf{x}; \mathbf{d})$ para $\alpha \geq 0$, a derivada direcional pode ser calculada como mostrado a seguir, se $f'(\mathbf{x}; -\mathbf{d}) = -f'(\mathbf{x}; \mathbf{d})$:

$$f'(\mathbf{x}; \mathbf{d}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x} + t\mathbf{d}) - f(\mathbf{x})}{t}.$$

⁴Ver (THOMAS, 2009b), p. 301

⁵Adaptado de (THOMAS, 2009b), p. 308

Fica claro que, quando a direção \mathbf{d} equivale a algum vetor de coordenadas de \mathbb{R}^n , como \mathbf{e}_i , aplicar a Definição 2.4.3.4 implica em calcular a derivada parcial de f na coordenada x_i , ou seja,

$$f'(\mathbf{x}; \mathbf{e}_i) \equiv \frac{\partial f}{\partial x_i}.$$

Com isso, permite-se concluir que a noção de derivada direcional é uma generalização da derivada parcial para qualquer direção.

Antes de se prosseguir com a análise de diferenciação de funções multivariáveis e campos vetoriais, faz-se necessário apresentar uma notação que será utilizada posteriormente, a *notação de Landau* (“o” pequeno):

Definição 2.4.3.5 (Notação de Landau) ⁶ Segundo a notação de Landau, um vetor pode ser chamado $o(\mathbf{h}) \in \mathbb{R}^n$ se

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|o(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

Definição 2.4.3.6 ⁷ Uma função $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é dita *Gâteaux diferenciável* em $\mathbf{x} \in U$ se a derivada direcional $f'(\mathbf{x}; \mathbf{d})$ existe para toda direção $\mathbf{d} \in \mathbb{R}^n$ e é uma função linear de \mathbf{d} .

Pela definição de diferenciabilidade de Gâteaux e utilizando a definição de direção, a derivada de Gâteaux pode ser calculada como

$$f'(\mathbf{x}; \mathbf{d}) = f'(\mathbf{x}; \sum_{i=1}^n d_i \mathbf{e}_i) = \sum_{i=1}^n d_i f'(\mathbf{x}; \mathbf{e}_i) = \sum_{i=1}^n d_i \frac{\partial f}{\partial x_i} = \langle \mathbf{d}, \nabla f \rangle = \mathbf{d}^T \nabla f. \quad (2.4.24)$$

Definição 2.4.3.7 ⁸ Uma função $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é dita *Fréchet diferenciável* em $\mathbf{x} \in U$ se existe uma função linear $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, $\ell(\mathbf{x}) = \mathbf{l}^T \mathbf{x}$, tal que

$$\lim_{\|\mathbf{h}\| \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \mathbf{l}^T \mathbf{h}}{\|\mathbf{h}\|} = 0.$$

Dizer que uma função f é Fréchet diferenciável em \mathbf{x} equivale a dizer que

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{l}^T \mathbf{h} + o(\mathbf{h}) \quad (2.4.25)$$

Dois fatos importantes podem ser retirados da Equação (2.4.25); o primeiro é que, se tomarmos o limite quando $\mathbf{h} \rightarrow \mathbf{0}$, tem-se que $\lim_{\mathbf{h} \rightarrow \mathbf{0}} f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x})$; isto equivale a dizer que, se f é Fréchet diferenciável em \mathbf{x} , ela é contínua nesse ponto. O outro fato a ser analisado pressupõe que \mathbf{l} pode ser escolhido como ∇f ; neste caso, aplicando-se também a Definição 2.4.1.1, (2.4.25) se torna

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \mathbf{h}^T \nabla f + o(\mathbf{h}). \quad (2.4.26)$$

⁶Ver (GULLER, 2010), p. 6

⁷*Ibid.*

⁸*Ibid.*

Claramente, de acordo com a Equação 2.4.24, o segundo termo é a derivada de Gâteaux de f em \mathbf{x} na direção \mathbf{h} . Portanto, f Fréchet diferenciável em \mathbf{x} implica em f Gâteaux diferenciável em \mathbf{x} .

A segunda derivada de uma função multivariável requer atenção especial: uma vez que a primeira derivada, ou gradiente, é um vetor, a segunda derivada de uma função $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ é uma matriz quadrada de ordem n . Essa matriz é conhecida como *Hessiana*.

Definição 2.4.3.8⁹ A segunda derivada de uma função $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, também conhecida como Hessiana de f , é dada por

$$H(f) \equiv \nabla^2 f \equiv f''(\mathbf{x}) = [h_{ij}], \quad h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}. \quad (2.4.27)$$

Um fato interessante sobre a Hessiana é que, se as segundas derivadas parciais de f existem e são contínuas, tem-se que $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$; neste caso, temos que $H(f) = H^T(f)$, ou seja, tem-se $H(f)$ simétrica. Este fato é largamente utilizado em alguns algoritmos de otimização, particularmente em problemas de programação quadrática.

Todos os conceitos vistos até aqui para funções $f : \mathbb{R}^n \rightarrow \mathbb{R}$ podem ser também aplicados a campos vetoriais $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Os conceitos de diferenciabilidade segundo Gâteaux e Fréchet são análogos aos presentes nas Definições 2.4.3.6 e 2.4.3.7¹⁰. Guller ainda afirma que, se um campo $\mathbf{F}(\mathbf{x})$ é Gâteaux (ou Fréchet) diferenciável em \mathbf{x} , então suas funções componentes $f_i(\mathbf{x})$ são Gâteaux (ou Fréchet) diferenciáveis em \mathbf{x} . A mudança aqui é o cálculo da derivada de \mathbf{F} , agora uma matriz.

Definição 2.4.3.9¹¹ Dado um campo vetorial $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, sua derivada, denominada Jacobiana de \mathbf{F} , é dada por

$$J(\mathbf{F}) = \left[\frac{\partial f_i}{\partial x_j} \right]. \quad (2.4.28)$$

Nota-se, pela definição de Jacobiana, que o conceito de gradiente de uma função multivariável é um caso específico de Jacobiana, assim como uma função multivariável é um tipo especial de campo vetorial, em que o contradomínio é unidimensional.

2.4.4 Convexidade

Segundo Izmailov e Solodov (IZMAILOV; SOLODOV, 2005), o conceito de convexidade é muito importante na teoria de otimização; com noções de convexidade, condições de otimalidade necessárias passam a ser suficientes, ou seja, basta encontrar um ponto estacionário para o problema. Em particular, sob condições de convexidade, todo minimizador local torna-se global. Outra possibilidade possível utilizando convexidade é o uso da teoria da dualidade na sua forma mais completa, ou seja, é possível associar o problema original (primal) a um problema alternativo (dual) que é, em determinadas condições, equivalente ao original e às vezes de mais fácil resolução. Por fim, o conceito de convexidade possibilita o uso de uma das condições de otimalidade mais poderosas aplicadas a problemas com restrições: as condições KKT.

⁹ Adaptado de (YANG, 2010), p. 51

¹⁰ Ver (GULLER, 2010), pp. 8-9

¹¹ Adaptado de (GULLER, 2010), p. 9

Definição 2.4.4.1 Um conjunto $D \subset \mathbb{R}^n$ é dito convexo se, para quaisquer $\mathbf{x} \in D$, $\mathbf{y} \in D$ e $\alpha \in [0, 1]$, tem-se $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in D$.

O ponto $\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}$, $\alpha \in [0, 1]$, é conhecido como a *combinação convexa* de \mathbf{x} e \mathbf{y} , com parâmetro α .

Em termos de convexidade de uma função, as definições a seguir são adaptadas de Izmailov e Solodov¹². Supõe-se, para todas as definições, que o conjunto $D \subset \mathbb{R}^n$ presente nelas é convexo.

Definição 2.4.4.2 Uma função $f : D \rightarrow \mathbb{R}$ é convexa em D se, para quaisquer $\mathbf{x} \in D$, $\mathbf{y} \in D$ e $\alpha \in [0, 1]$, tem-se

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}). \quad (2.4.29)$$

Definição 2.4.4.3 Uma função $f : D \rightarrow \mathbb{R}$ é estritamente convexa em D se, para quaisquer $\mathbf{x} \in D$, $\mathbf{y} \in D$, $\mathbf{x} \neq \mathbf{y}$ e $\alpha \in (0, 1)$, tem-se

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) < \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}). \quad (2.4.30)$$

Definição 2.4.4.4 Uma função $f : D \rightarrow \mathbb{R}$ é fortemente convexa em D , com módulo $\gamma > 0$ se, para quaisquer $\mathbf{x} \in D$, $\mathbf{y} \in D$ e $\alpha \in [0, 1]$, tem-se

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y}) - \gamma\alpha(1 - \alpha)\|\mathbf{x} - \mathbf{y}\|^2. \quad (2.4.31)$$

Vale notar que uma função fortemente convexa é estritamente convexa, e que uma função estritamente convexa é uma função convexa; a recíproca, porém, nem sempre é verdadeira.

Definição 2.4.4.5 Uma função $f : D \rightarrow \mathbb{R}$ é côncava em D se $-f$ for convexa em D .

Definição 2.4.4.6 O problema de otimização (2.4.5) é um problema de minimização convexo quando D é um conjunto convexo e f é uma função convexa.

Uma vez que a função e o conjunto de restrições são convexos, resolver um problema de otimização torna-se menos tortuoso; encontrar um minimizador é garantido. A importância desse fato pode ser vista no teorema a seguir, retirado de Izmailov e Solodov¹³.

Teorema 2.4.4.7 (Teorema de minimização convexa) Todo minimizador local de um problema convexo é um minimizador global, e o conjunto de minimizadores é convexo. Além disso, se a função objetivo for estritamente convexa, só há no máximo um minimizador.

Uma vez que resolver um problema de maximização é análogo a resolver um problema de minimização, o teorema acima pode ser adaptado com o conceito de concavidade visto na Definição 2.4.4.6: o problema a ser resolvido neste caso se trata de maximização de uma função côncava num conjunto convexo.

Tentar determinar a convexidade de uma função por meio da Definição 2.4.4.2 pode-se tornar um trabalho árduo; Yang apresenta uma alternativa para se determinar a convexidade de uma função, dada a seguir.

¹²Ver (IZMAILOV; SOLODOV, 2005), pp. 66-70

¹³*Ibid.*, p. 69. Ver prova em *Ibid.*, pp. 69-70.

Teorema 2.4.4.8 ¹⁴ Uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é convexa se sua matriz Hessiana é semidefinida positiva em todos os pontos de D .

Guller destaca, porém, que o teorema acima é válido apenas se a função f é duas vezes Fréchet diferenciável; além disso, afirma que, se a Hessiana é definida positiva, a função f é estritamente convexa¹⁵.

Um outro resultado importante obtido com análise convexa se refere aos conceitos de diferenciabilidade de Gâteaux e Fréchet. O teorema a seguir destaca essa relação.

Teorema 2.4.4.9 ¹⁶ Seja $C \subset \mathbb{R}^n$ convexo com interior não-vazio ($\text{int}(C) \neq \emptyset$) e $f : C \rightarrow \mathbb{R}$ uma função convexa. Se todas as derivadas parciais de f existem para um ponto \mathbf{x} no interior de C , f é Fréchet diferenciável em \mathbf{x} . Mais ainda, se f é Gâteaux diferenciável em \mathbf{x} , f é Fréchet diferenciável em \mathbf{x} .

Nota-se que a convexidade suprime a distinção entre os conceitos de diferenciabilidade de Gâteaux e Fréchet vistos anteriormente; isso resulta em uma melhor análise de derivadas da função atingida, visto que obter derivadas de Fréchet normalmente é um processo mais difícil que a diferenciação segundo Gâteaux.

A presente Seção se dedicou a abordar os conceitos mais elementares de análise convexa, que foram analisados no presente estudo; Guller, Izmailov e Solodov apresentam outros elementos dessa análise, como os teoremas de separação, que são aplicados à noção de dualidade. Tendo em vista este fato, algumas condições de otimalidade foram reunidas, considerando ou não análise convexa. As semelhanças e diferenças entre tais condições serão vistas na Seção a seguir.

2.4.5 Condições de Otimalidade

As condições de otimalidade podem ser vistas como condições que devem ser satisfeitas para que um ponto dado seja minimizador de uma função, ou condições que garantem que um ponto é minimizador da função; tais condições são denominadas, respectivamente, *condições necessárias de otimalidade* e *condições suficientes de otimalidade*. Existem várias maneiras de apresentar essas condições; contudo, para efeitos deste estudo, serão apresentadas somente condições para problemas irrestritos e problemas com restrições de igualdade e desigualdade.

Definição 2.4.5.1 Um problema de minimização irrestrita é aquele cuja forma é

$$\min f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n. \quad (2.4.32)$$

Analogamente, um problema de maximização irrestrita pode ser escrito como

$$\max f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n. \quad (2.4.33)$$

Um fato importante sobre a existência de um minimizador global de um problema pode ser mostrado pelo teorema a seguir:

¹⁴ Adaptado de (YANG, 2010), pp. 56-57.

¹⁵ Ver Teorema 4.28 em (GULLER, 2010), p. 99.

¹⁶ Ver escrita original e prova em (GULLER, 2010), p. 101.

Teorema 2.4.5.2 (Teorema de Weierstrass) ¹⁷ *Seja $f : K \rightarrow \mathbb{R}$ uma função contínua e K um espaço de medida compacto. Logo, existe um ponto $\bar{\mathbf{x}} \in K$ que é minimizador global de f em K , isto é,*

$$f(\bar{\mathbf{x}}) \leq f(\mathbf{x}), \forall \mathbf{x} \in K.$$

As condições de otimalidade a serem apresentadas a seguir são relevantes para o problema irrestrito (2.4.32)¹⁸:

Teorema 2.4.5.3 (Condição Necessária de Primeira Ordem) *Seja uma função $f : U \rightarrow \mathbb{R}$ Gâteaux diferenciável em um conjunto aberto $U \subseteq \mathbb{R}^n$. Um mínimo local é também ponto crítico de f , ou seja,*

$$\mathbf{x} \text{ é mínimo local} \Rightarrow \nabla f(\mathbf{x}) = 0.$$

Corolário 2.4.5.4 *Seja uma função $f : U \rightarrow \mathbb{R}$ definida em um conjunto aberto $U \subseteq \mathbb{R}^n$. Se $\mathbf{x} \in U$ é um minimizador local de f e existe a derivada direcional $f'(\mathbf{x}; \mathbf{d})$ para alguma direção $\mathbf{d} \in \mathbb{R}^n$, então $f'(\mathbf{x}; \mathbf{d}) \geq 0$.*

Teorema 2.4.5.5 (Condição Necessária de Segunda Ordem) *Seja uma função $f : U \rightarrow \mathbb{R}$ duas vezes Gâteaux diferenciável em um conjunto aberto $U \subseteq \mathbb{R}^n$, e com segundas derivadas parciais contínuas ($f \in C^2$). Se $\mathbf{x} \in U$ é um minimizador local de f , então sua Hessiana $H(f(\mathbf{x}))$ é semidefinida positiva.*

Teorema 2.4.5.6 (Condições Suficientes de Segunda Ordem) *Seja uma função $f : U \rightarrow \mathbb{R}$ tal que $f \in C^2$ em um conjunto aberto $U \subseteq \mathbb{R}^n$. Portanto:*

- (a) *Se $\nabla f(\mathbf{x}) = 0$ e $H(f(\mathbf{x}))$ é definida positiva para $\mathbf{x} \in U$, então \mathbf{x} é minimizador local estrito de f ;*
- (b) *Se o conjunto U é convexo, $H(f)$ é semidefinida positiva em U e $\nabla f(\mathbf{x}) = 0$, então \mathbf{x} é minimizador global de f ;*
- (c) *Se $\nabla f(\mathbf{x}) = 0$ e $H(f(\mathbf{x}))$ é indefinida para $\mathbf{x} \in U$, então \mathbf{x} é um ponto de sela de f .*

Agora, seja um problema de otimização em sua forma geral

$$\begin{aligned} \min f(\mathbf{x}) \\ \text{s. a.} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, r \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \dots, m \end{aligned} \quad (2.4.34)$$

Tal problema é também conhecido como *programa não-linear* ou *matemático*. Ele será denotado por \mathbb{P} .

Definição 2.4.5.7 *O conjunto*

$$\mathcal{F}(\mathbb{P}) = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0, 1 \leq i \leq r, 1 \leq j \leq m\}$$

é denominado conjunto factível de \mathbb{P} .

¹⁷Ver teorema e prova em (GULLER, 2010), p. 33

¹⁸Essas condições, juntamente com suas provas, se encontram em (GULLER, 2010), pp. 35-39.

Deve-se ressaltar que se $g_i(\bar{\mathbf{x}}) < 0$ implica que essa i -ésima condição de desigualdade não influi na determinação da condição de minimizador local de $\bar{\mathbf{x}}$; tal condição é denominada *inativa*. Sendo assim, a Definição a seguir trata das restrições de desigualdade *ativas* do problema \mathbb{P} .

Definição 2.4.5.8 Uma restrição $g_i(\mathbf{x})$ é dita ativa se $g_i(\mathbf{x}) = 0$, $\mathbf{x} \in \mathcal{F}(\mathbb{P})$. O conjunto

$$\mathcal{I}(\mathbf{x}) := \{i \mid g_i(\mathbf{x}) = 0\}$$

é chamado de conjunto de índices das restrições ativas de \mathbb{P} .

Definição 2.4.5.9 Um ponto factível $\bar{\mathbf{x}} \in \mathcal{F}(\mathbb{P})$ é minimizador local do problema \mathbb{P} se for minimizador de f numa vizinhança factível de $\bar{\mathbf{x}}$, ou seja,

$$\exists \epsilon > 0 \bullet f(\bar{\mathbf{x}}) \leq f(\mathbf{x}), \forall \mathbf{x} \in \mathcal{F}(\mathbb{P}) \cap \bar{B}_\epsilon(\bar{\mathbf{x}}), \quad (2.4.35)$$

em que $\bar{B}_\epsilon(\bar{\mathbf{x}})$ é uma bola aberta de centro $\bar{\mathbf{x}}$ e raio ϵ . Caso $\bar{\mathbf{x}}$ satisfaça

$$f(\bar{\mathbf{x}}) \leq f(\mathbf{x}), \forall \mathbf{x} \in \mathcal{F}(\mathbb{P}), \quad (2.4.36)$$

o ponto é um minimizador global do problema \mathbb{P} .

Vale ressaltar que, utilizando os sinais de desigualdade apropriados, as definições de minimizadores podem ser modificadas para o caso de maximizadores.

Para se dar prosseguimento à análise de condições de otimalidade sobre o problema (2.4.34), será definido o operador *Lagrangiano* do problema (também conhecido como multiplicadores de Lagrange):

Definição 2.4.5.10 A função

$$\mathcal{L}(\mathbf{x}, \lambda, \mu) := \lambda_0 f(\mathbf{x}) + \sum_{i=1}^r \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^m \mu_j h_j(\mathbf{x}) \quad (\lambda_i \geq 0, 0 \leq i \leq r) \quad (2.4.37)$$

é chamada de *Lagrangiana fraca* do problema \mathbb{P} . Caso $\lambda_0 = 1$, a função é simplesmente chamada de *Lagrangiana*.

A Lagrangiana é utilizada como base para as condições a seguir:

Teorema 2.4.5.11 (Condições de Fritz-John) Se $\bar{\mathbf{x}}$ é minimizador local de (2.4.34), então existem multiplicadores $(\lambda, \mu) := (\lambda_0, \lambda_1, \dots, \lambda_r, \mu_1, \mu_2, \dots, \mu_m)$, não todos nulos, com $(\lambda_0, \dots, \lambda_r) \geq 0$, em que

$$\nabla_{\mathbf{x}} \mathcal{L}(\bar{\mathbf{x}}, \lambda, \mu) = 0, \quad (2.4.38)$$

$$\lambda_i \geq 0, g_i(\bar{\mathbf{x}}) \leq 0, \lambda_i g_i(\bar{\mathbf{x}}) = 0, i = 1, 2, \dots, r. \quad (2.4.39)$$

Ressalta-se que a Lagrangiana nessas condições está em sua versão fraca.

Utilizando-se a noção de $\mathcal{I}(\mathbf{x})$, vista na Definição 2.4.5.8, e utilizando-se o gradiente do Lagrangiano,

pode-se reescrever (2.4.38) na forma

$$\lambda_0 \nabla f(\bar{\mathbf{x}}) + \sum_{j \in \mathcal{I}(\bar{\mathbf{x}})} \lambda_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{j=1}^m \mu_j \nabla h_j(\bar{\mathbf{x}}) = 0. \quad (2.4.40)$$

Uma aplicação interessante das condições FJ reside no fato de que agora é possível escrever *condições suficientes de primeira ordem* de otimalidade:

Teorema 2.4.5.12 (Condições Suficientes de Primeira Ordem) *Seja $\bar{\mathbf{x}}$ uma solução viável para o problema \mathbb{P} em (2.4.34), satisfazendo as condições FJ, em que a primeira condição é usada conforme (2.4.40). Se a totalidade dos vetores*

$$\lambda_0 \nabla f(\bar{\mathbf{x}}), \{ \lambda_i \nabla g_i(\bar{\mathbf{x}}) \}_{i \in \mathcal{I}(\bar{\mathbf{x}})}, \{ \nabla h_j(\bar{\mathbf{x}}) \}_1^m$$

formam uma base de \mathbb{R}^n , então $\bar{\mathbf{x}}$ é minimizador local de \mathbb{P} .

A importância do teorema de Fritz-John reside no fato de que ele sempre é aplicável nos pontos minimizadores locais. Porém, há casos em que λ_0 pode ser 0, um fato estranho visto que significa que a função objetivo não teria influência nas condições de otimalidade de primeira ordem. Portanto, são necessárias suposições adicionais sobre o problema \mathbb{P} de maneira que essa possibilidade não seja alcançada. Tais suposições que garantem $\lambda_0 > 0$ ($\lambda_0 = 1$, de fato) são denominadas *qualificação das restrições*, e as condições de otimalidade resultantes são chamadas *condições de Karush-Kuhn-Tucker (KKT)*.

Corolário 2.4.5.13 (Condições de Karush-Kuhn-Tucker) *Se os vetores*

$$\{ \nabla g_i(\bar{\mathbf{x}}), i \in \mathcal{I}(\bar{\mathbf{x}}), \nabla h_j(\bar{\mathbf{x}}), j = 1, \dots, m \}$$

são linearmente independentes, então $\lambda_0 > 0$ e, portanto, utilizando a função Lagrangiana não-fraca,

$$\nabla_{\mathbf{x}} \mathcal{L}(\bar{\mathbf{x}}, \lambda, \mu) = 0, \quad (2.4.41)$$

$$\lambda_i \geq 0, g_i(\bar{\mathbf{x}}) \leq 0, \lambda_i g_i(\bar{\mathbf{x}}) = 0, i = 1, 2, \dots, r, \quad (2.4.42)$$

$$h_j(\bar{\mathbf{x}}) = 0, j = 1, 2, \dots, m. \quad (2.4.43)$$

O problema das condições KKT reside no fato de que elas falham em pontos que, ao se aplicar as condições FJ, tem-se $\lambda_0 = 0$, o que significa que a função objetivo não entra nas condições de otimalidade, o contrário do que é esperado. É importante então, ao se considerar as condições KKT, identificar, dado o problema \mathbb{P} em (2.4.34), condições adicionais sobre a função objetivo f e principalmente sobre as restrições de desigualdade g_i e de igualdade h_j . Serão apresentadas, a seguir, condições e necessárias para a existência das condições KKT:

Teorema 2.4.5.14 *Seja um ponto $\bar{\mathbf{x}}$ FJ para o problema \mathbb{P} . As condições KKT se aplicam a $\bar{\mathbf{x}}$ se e somente se*

$$\begin{aligned} & \{ \mathbf{d} \mid \langle \nabla f(\bar{\mathbf{x}}), \mathbf{d} \rangle < 0 \} \cap \{ \mathbf{d} \mid \langle \nabla g_i(\bar{\mathbf{x}}), \mathbf{d} \rangle \leq 0, i \in \mathcal{I}(\bar{\mathbf{x}}) \} \\ & \cap \{ \mathbf{d} \mid \langle \nabla h_j(\bar{\mathbf{x}}), \mathbf{d} \rangle = 0, j = 1, \dots, m \} = \emptyset. \end{aligned} \quad (2.4.44)$$

Corolário 2.4.5.15 (Restrições Lineares e Côncavas) *Seja $\bar{\mathbf{x}}$ minimizador local de \mathbb{P} . As condições KKT se aplicam a $\bar{\mathbf{x}}$ se as restrições ativas $\{g_i\}_{i \in \mathcal{I}(\bar{\mathbf{x}})}$ forem funções côncavas numa vizinhança convexa de $\bar{\mathbf{x}}$ e as restrições de igualdade $\{h_j\}_1^m$ forem funções afins em \mathbb{R}^n .*

Em particular, as condições KKT se aplicam a todos os minimizadores locais se todas as restrições g_i e h_j forem funções afins, ou seja,

$$g_i(\mathbf{x}) = \langle \mathbf{a}_i, \mathbf{x} \rangle + \alpha_i, \quad h_j(\mathbf{x}) = \langle \mathbf{b}_j, \mathbf{x} \rangle + \beta_j.$$

Teorema 2.4.5.16 (Mangasarian-Fromovitz) *Seja um ponto $\bar{\mathbf{x}}$ FJ para o problema \mathbb{P} . Se os gradientes das restrições de igualdade $\{\nabla h_j(\bar{\mathbf{x}})\}_1^m$ forem linearmente independentes e se existir uma direção \mathbf{d} tal que*

$$\langle \nabla g_i(\bar{\mathbf{x}}), \mathbf{d} \rangle < 0, \quad i \in \mathcal{I}(\bar{\mathbf{x}}), \quad \langle \nabla h_j(\bar{\mathbf{x}}), \mathbf{d} \rangle = 0, \quad j = 1, \dots, m, \quad (2.4.45)$$

então as condições KKT são satisfeitas em $\bar{\mathbf{x}}$

Uma das qualificações de restrições mais antigas e conhecidas é a *qualificação de restrições de Slater*, quando as restrições são convexas.

Corolário 2.4.5.17 (Slater) *Seja o problema \mathbb{P} em (2.4.34), com as restrições de desigualdade convexas, as restrições de desigualdade afins e um minimizador local $\bar{\mathbf{x}}$. Se existe um ponto viável \mathbf{x}_0 tal que $g_i(\mathbf{x}_0) < 0$, $i \in \mathcal{I}(\bar{\mathbf{x}})$, então as condições KKT são satisfeitas em $\bar{\mathbf{x}}$.*

Todas as condições apresentadas até aqui são condições de otimalidade de primeira ordem para o problema (2.4.34)¹⁹. A seguir, serão apresentadas as condições de segunda ordem para (2.4.34).

Primeiramente, denota-se por $\nabla_{\mathbf{x}}^2 \mathcal{L}(\bar{\mathbf{x}}, \lambda, \mu)$ a Hessiana do Lagrangiano em relação a \mathbf{x} do problema \mathbb{P} :

$$\nabla_{\mathbf{x}}^2 \mathcal{L}(\bar{\mathbf{x}}, \lambda, \mu) = \nabla^2 f(\bar{\mathbf{x}}) + \sum_{i=1}^r \lambda_i \nabla^2 g_i(\bar{\mathbf{x}}) + \sum_{j=1}^m \mu_j \nabla^2 h_j(\bar{\mathbf{x}}).$$

Teorema 2.4.5.18 (Condições Necessárias de Segunda Ordem) *Seja $\bar{\mathbf{x}}$ um minimizador local do problema \mathbb{P} satisfazendo as condições KKT com multiplicadores $\bar{\lambda}$ e $\bar{\mu}$. Se os gradientes das condições ativas*

$$\nabla g_i(\bar{\mathbf{x}}), \quad i \in \mathcal{I}(\bar{\mathbf{x}}), \quad \nabla h_j(\bar{\mathbf{x}}), \quad j = 1, \dots, m$$

são linearmente independentes, então $\nabla_{\mathbf{x}}^2 \mathcal{L}(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu})$ deve ser semidefinida positiva no subespaço linear dado por

$$M = (\text{span} \{ \nabla g_i(\bar{\mathbf{x}}), i \in \mathcal{I}(\bar{\mathbf{x}}), \nabla h_j(\bar{\mathbf{x}}), j = 1, \dots, m \})^\perp,$$

isto é, se uma direção \mathbf{d} satisfaz

$$\langle \mathbf{d}, \nabla g_i(\bar{\mathbf{x}}) \rangle = 0, \quad i \in \mathcal{I}(\bar{\mathbf{x}}), \quad \langle \mathbf{d}, \nabla h_j(\bar{\mathbf{x}}) \rangle = 0, \quad j = 1, \dots, m,$$

então $\langle \nabla_{\mathbf{x}}^2 \mathcal{L}(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}) \mathbf{d}, \mathbf{d} \rangle \geq 0$.

¹⁹Mais detalhes e provas dessas condições se encontram em (GULLER, 2010), pp. 211-220.

Teorema 2.4.5.19 (Condições Suficientes de Segunda Ordem) *Seja $\bar{\mathbf{x}}$ um minimizador local do problema \mathbb{P} satisfazendo as condições KKT com multiplicadores $\bar{\lambda}$ e $\bar{\mu}$. Se*

$$\langle \nabla_{\bar{\mathbf{x}}}^2 \mathcal{L}(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu}) \mathbf{d}, \mathbf{d} \rangle > 0 \quad (2.4.46)$$

para todo $\mathbf{d} \neq \mathbf{0}$ tal que

$$\begin{aligned} \langle \mathbf{d}, \nabla g_i(\bar{\mathbf{x}}) \rangle &\leq 0, \quad i \in \mathcal{I}(\bar{\mathbf{x}}), \\ \langle \mathbf{d}, \nabla g_i(\bar{\mathbf{x}}) \rangle &= 0, \quad i \in \mathcal{I}(\bar{\mathbf{x}}) \text{ e } \bar{\lambda}_i > 0, \\ \langle \mathbf{d}, \nabla h_j(\bar{\mathbf{x}}) \rangle &= 0, \quad j = 1, \dots, m, \end{aligned} \quad (2.4.47)$$

então $\bar{\mathbf{x}}$ é minimizador local estrito de \mathbb{P} e existem uma constante $c > 0$ e uma bola $\bar{B}_\epsilon(\bar{\mathbf{x}})$ tal que

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + c\|\mathbf{x} - \bar{\mathbf{x}}\|^2, \quad \forall \mathbf{x} \in \bar{B}_\epsilon(\bar{\mathbf{x}}) \text{ viável.} \quad (2.4.48)$$

Corolário 2.4.5.20 *Seja $\bar{\mathbf{x}}$ um minimizador local do problema \mathbb{P} satisfazendo as condições KKT com multiplicadores $\bar{\lambda}$ e $\bar{\mu}$. Se $\bar{\lambda}_i > 0, \forall i \in \mathcal{I}(\bar{\mathbf{x}})$ e a Hessiana $\nabla_{\bar{\mathbf{x}}}^2 \mathcal{L}(\bar{\mathbf{x}}, \bar{\lambda}, \bar{\mu})$ for definida positiva no subespaço*

$$\{\mathbf{d} \mid \langle \mathbf{d}, \nabla g_i(\bar{\mathbf{x}}) \rangle = 0, \quad i \in \mathcal{I}(\bar{\mathbf{x}}), \quad \langle \mathbf{d}, \nabla h_j(\bar{\mathbf{x}}) \rangle = 0, \quad j = 1, \dots, m\},$$

então $\bar{\mathbf{x}}$ é minimizador local estrito de \mathbb{P} .

As condições de otimalidade até aqui apresentadas foram além daquelas para problemas irrestritos; agora, além de procurar minimizar a função objetivo, deve-se ter cuidado com as restrições. Nota-se que é conveniente que essas restrições sejam de natureza específica, de maneira a tornar o problema \mathbb{P} viável em termos de resolução²⁰.

2.4.6 Principais Algoritmos de Otimização

O propósito desta Seção é oferecer uma visão geral dos principais algoritmos de otimização vistos na literatura; uma análise detalhada dos mesmos é excessiva do ponto de vista deste relatório de estudo dirigido; tal análise demanda até várias produções literárias, conforme destaca Yang (YANG, 2010).

Segundo Yang, resolver um problema de otimização pode ser comparado a uma caça ao tesouro: imaginemos que estamos caçando um tesouro em uma cordilheira, com limite de tempo. Em um extremo, não temos ideia de onde começar a procurar e estamos de olhos vendados; isto resulta em uma busca aleatória, que não é tão eficiente quanto poderíamos esperar. Por outro lado, temos ideia de que o tesouro se encontra no ponto mais alto da cordilheira; isto nos leva a buscar um caminho direto. Na maioria dos casos, estamos entre os dois extremos: não estamos de olhos vendados, mas não sabemos por onde começar a procurar. Uma vez que é ineficiente andar em passos aleatórios, na prática andamos seguindo algumas pistas, olhando de forma razoavelmente aleatória, mas com um propósito por trás; esta é a essência de vários algoritmos de otimização modernos.

De maneira geral, algoritmos de otimização podem ser classificados como *determinísticos* ou *estocásticos*. Em alguns algoritmos determinísticos, a noção de gradiente é utilizada; são os algoritmos *baseados*

²⁰Todas as condições até aqui e suas provas se encontram em (GULLER, 2010), pp. 230-235.

em gradiente. Quando a função objetivo apresenta descontinuidades, tais algoritmos tendem a falhar; sendo assim, há algoritmos determinísticos que não utilizam gradiente; são os algoritmos *livres de gradiente*.

Os algoritmos estocásticos são divididos em dois tipos, geralmente: *heurísticos* e *meta-heurísticos*; embora a diferença entre esses tipos seja pequena. De maneira geral, *heurística* significa “descoberta ou busca por tentativa e erro”. Boas soluções podem ser encontradas deste modo em tempo razoável; não é garantido, porém, que a solução ótima seja encontrada. Isso é vantajoso quando se quer não uma solução ótima dificilmente atingível, mas uma boa solução que possa ser encontrada com razoável facilidade.

Na linha dos algoritmos heurísticos, a adição de aleatorização e buscas locais gera a classe dos chamados algoritmos *meta-heurísticos* — “meta-” aqui quer dizer “além de”. De forma geral, os algoritmos meta-heurísticos possuem melhor desempenho do que seus equivalentes heurísticos.

Antes de se proceder à descrição dos principais algoritmos de otimização, revisitaremos a definição de um problema de otimização com restrições, apresentado como \mathbb{P} :

$$\begin{aligned} \min f(\mathbf{x}) \\ \text{s. a.} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, r \quad (\mathbb{P}). \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \dots, m \end{aligned} \quad (2.4.49)$$

Duas classes de problemas especiais devem ser mencionadas a respeito do problema \mathbb{P} :

- **Programação Linear** Caso especial do problema \mathbb{P} quando a função objetivo é uma forma afim, isto é,

$$f(\mathbf{x}) = \mathbf{q}^T \mathbf{x} + c;$$

- **Programação Quadrática** Caso especial do problema \mathbb{P} quando a função objetivo é uma forma quadrática, isto é,

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{q}^T \mathbf{x} + c,$$

Com \mathbf{H} simétrica. Em caso contrário, faz-se a aproximação

$$\tilde{\mathbf{H}} := \frac{1}{2}(\mathbf{H} + \mathbf{H}^T),$$

que é sempre simétrica.

Ressalta-se que esses casos especiais são facilmente resolvidos sob dadas condições — por exemplo, a matriz \mathbf{H} (ou $\tilde{\mathbf{H}}$) ser semidefinida positiva. Para outras classes de problemas, serão apresentados vários algoritmos. Tais algoritmos podem ser encontrados alguns em (GULLER, 2010) e outros em (YANG, 2010):

- **Método de Newton:** Assume que o problema \mathbb{P} é irrestrito e a função objetivo é continuamente diferenciável. Originalmente aplicado à solução de equações, o Método de Newton também pode ser aplicado em otimização no sentido de que ele é utilizado para resolver a condição de otimalidade necessária de primeira ordem (ver Teorema 2.4.5.3). É um algoritmo iterativo cujos passos são dados

por

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} - H^{-1}(f(\mathbf{x}^{(n)}))f(\mathbf{x}^{(n)}), \quad (2.4.50)$$

tomando-se um ponto viável inicial $\mathbf{x}^{(0)}$. H^{-1} é a matriz Hessiana inversa da função objetivo; uma dificuldade desse algoritmo reside no condicionamento numérico resultante da aplicação de sucessivas inversões dessa matriz, além do cálculo dessas inversões em si.

- **Método do Gradiente Descendente:** Busca obter o menor valor da função objetivo a partir de um ponto $\mathbf{x}^{(0)}$. Toma-se um passo $\alpha(i) > 0$ para cada iteração, com o cuidado de situá-lo em casos que as iterações aproximem-se adequadamente do ponto desejado. Cada iteração é dada por

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} - \alpha^{(n)} \left\| \nabla f(\mathbf{x}^{(n)}) \right\|_2^2.$$

- **Método Simplex:** Utilizado em problemas de programação linear, foi introduzido por George Dantzig em 1947. Funciona da seguinte forma: assume-se que os pontos extremos do problema são conhecidos, ou se determina esses pontos para se checar a existência de solução viável. Com esses pontos conhecidos, é trivial determinar o ponto de ótimo utilizando relações algébricas e a função objetivo. Se o teste de otimalidade falha, um ponto extremo adjacente é testado. O algoritmo para em caso de encontro de uma solução viável ou quando se trata de um problema ilimitado²¹.
- **Método de Penalidade:** Utilizado em problemas de forma geral, como (2.4.49). A ideia é definir uma função de penalidade a ser minimizada de maneira que o problema de otimização sobre ela seja irrestrito. Normalmente, essa função de penalização é dada por

$$\Pi(\mathbf{x}, \mu_i, \nu_j) = f(\mathbf{x}) + \sum_{i=1}^r \mu_i g_i(\mathbf{x}) + \sum_{j=1}^m \nu_j h_j(\mathbf{x}), \quad (2.4.51)$$

em que $\mu_i \gg 1$ e $\nu_j \geq 0$ grandes o suficiente para se garantir uma boa qualidade da solução a ser encontrada. Porém, um método mais geral para transformar um problema com restrições num problema irrestrito é utilizar ferramentas como as condições FJ e KKT.

- **Algoritmo BFGS:** É um tipo de algoritmo quase-Newton utilizado na resolução de problemas de otimização irrestritos com função objetivo não-linear. A ideia é aproximar a Hessiana da função por uma matriz $\mathbf{B}^{(n)}$. A ideia geral se resume em utilizar as equações

$$\begin{aligned} \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \beta^{(k)} \mathbf{s}^{(k)}, \\ \mathbf{u}^{(k)} &= \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \quad \mathbf{v}^{(k)} = \nabla f(\mathbf{x}^{(k+1)}) - \nabla f(\mathbf{x}^{(k)}), \\ \mathbf{B}^{(k+1)} &= \mathbf{B}^{(k)} + \frac{\mathbf{v}^{(k)} \mathbf{v}^{(k)T}}{\mathbf{v}^{(k)T} \mathbf{v}^{(k)}} - \frac{(\mathbf{B}^{(k)} \mathbf{u}^{(k)}) (\mathbf{B}^{(k)} \mathbf{u}^{(k)})^T}{\mathbf{u}^{(k)T} \mathbf{B}^{(k)} \mathbf{u}^{(k)}}, \end{aligned}$$

para iterar a matriz \mathbf{B} de maneira a se buscar o ótimo desejado.

- **Algoritmo Nelder-Mead:** Desenvolvido por J. A. Nelder e R. Mead em 1965, possui a ideia de se buscar a solução de um problema de otimização por meio de operações sobre figuras n -dimensionais conhecidas como *simplex*. Um *simplex* nada mais é do que uma generalização do triângulo para todas

²¹Ver algoritmo detalhado em (YANG, 2010), pp 70-75.

dimensionais; ou seja, um simplex n -dimensional é definido pelo fecho convexo (o menor conjunto convexo contendo todos os pontos dados) de $n + 1$ pontos distintos. Tal simplex pode ser refletido, expandido, contraído ou reduzido de forma a conter, eventualmente, a solução do problema. Este fato deu o apelido de “Algoritmo da Ameba” para este método.

- **Programação Quadrática Sequencial (SQP):** É outro método bastante difundido na literatura. Consiste em utilizar sucessivas aplicações de Programação Quadrática para se encontrar a solução do problema original, considerando-se que a função objetivo pode ser aproximada por uma expansão de Taylor de 2ª ordem. O problema a ser resolvido em cada iteração se torna

$$\begin{aligned} \min \quad & \frac{1}{2} \mathbf{s}^T \nabla^2 \mathcal{L}(\mathbf{x}^{(k)}) \mathbf{s} + \nabla f(\mathbf{x}^{(k)})^T \mathbf{s} + f(\mathbf{x}^{(k)}) \quad \text{s. a.} \\ & \nabla g_i(\mathbf{x}^{(k)}) \mathbf{s} + g_i(\mathbf{x}^{(k)}) \leq 0, \quad i = 1, \dots, r \quad (\mathbb{P}^*), \\ & \nabla h_j(\mathbf{x}^{(k)}) \mathbf{s} + h_j(\mathbf{x}^{(k)}) = 0, \quad j = 1, \dots, m \end{aligned} \quad (2.4.52)$$

em que $\nabla^2 \mathcal{L}(\mathbf{x}^{(k)})$ é a Hessiana do Lagrangiano da função objetivo aplicada em $\mathbf{x}^{(k)}$. Nota-se que, em \mathbf{s} , as restrições do problema \mathbb{P}^* são todas afins; portanto vale o Corolário 2.4.5.15, ou seja, as condições KKT se aplicam a todos os minimizadores locais de \mathbb{P}^* . Ao se resolver \mathbb{P}^* , o ponto $\bar{\mathbf{s}}^{(k)}$ encontrado é então utilizado para atualizar a solução do problema original, utilizando-se um fator de correção α :

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \bar{\mathbf{s}}^{(k)}. \quad (2.4.53)$$

Uma vez que é dispendioso calcular a Hessiana do Lagrangiano a cada iteração, uma alternativa é aproximá-la utilizando uma aproximação BFGS, já discutida anteriormente. É de interesse que, a cada iteração, a função objetivo de \mathbb{P}^* tenha Hessiana semidefinida positiva, o que garante sua convexidade e existência de solução para o problema de otimização a cada iteração.

- **Algoritmos meta-heurísticos:** Como discutido anteriormente, são algoritmos de tentativa e erro que utilizam buscas aleatórias, mas com uma predição de passos inerente à estrutura do algoritmo para evitar o uso da chamada “força bruta”. Normalmente, esses algoritmos se baseiam em fenômenos naturais, visto que já foi observada a eficiência de fatos naturais em resultados considerados ótimos. Alguns algoritmos nessa classe são²²:

Algoritmos Genéticos: Se baseiam em princípios da genética, notadamente a teoria da seleção natural de Charles Darwin;

Simulated Annealing: É baseado em uma mimetização do processo de recozimento de um material em que um metal resfria e é congelado em um estado cristalino de energia mínima e maior tamanho dos cristais de maneira a reduzir defeitos em estruturas metálicas;

Algoritmo das Formigas: Simula o comportamento geral de uma colônia de formigas na natureza. Particularmente, é uma analogia à comunicação desses animais por meio de feromônios, em casos de busca e encontro de comida, por exemplo;

Algoritmo das Abelhas: Bem similar ao Algoritmo das Formigas, tomando-se por base colônias de abelhas; além da análise de feromônios, há a análise de “danças sinalizadoras” desses animais sinalizando alguns tipos de ocorrências;

²²Todos os algoritmos se encontram em (YANG, 2010), pp. 173-229.

Particle Swarm (PSO): Utiliza o comportamento de grandes grupos de animais como peixes e aves. É mais simples que algoritmos genéticos e os de abelhas e formigas por não utilizar noções como *crossover* genético ou feromônios; usa apenas randomização e comunicação geral entre os elementos considerados;

Harmony Search: Baseado na música, e sua tentativa de se encontrar um estado de harmonia; a harmonia musical pode ser comparada ao ótimo em um problema de otimização. Os passos podem ser entendidos como um refinamento do musicista;

Algoritmo dos Vaga-Lumes: Assim como os Algoritmos das Formigas e das Abelhas, utiliza um comportamento animal como base de seus passos. Neste caso, utiliza o fato de que vaga-lumes usam sinais luminosos distintos para várias formas de comunicação, como busca de parceiros, encontro de presas e fuga de predadores.

2.4.7 Tratamento de Problemas de Otimização

Segundo Reklaitis (REKLAITIS, 1983), para que seja possível se aplicar as técnicas e algoritmos de otimização descritos até aqui em problemas concretos de engenharia, é necessário:

- Definir os limites do sistema a ser otimizado;
- Definir um critério de classificação das soluções candidatas para se determinar qual a “melhor”;
- Selecionar as variáveis do sistema que serão utilizadas para caracterizar ou identificar pontos candidatos;
- Definir um modelo que expressará o modo com que as variáveis relacionadas se relacionam.

Definir os limites do sistema significa que definir os limites que separam o sistema estudado do resto do universo. Esses limites servem para isolar o sistema da sua vizinhança, pois, para fins de análise, todas as interações entre sistema e vizinhança são consideradas inativas em determinados níveis. Como essas interações sempre existem, definir os limites do sistema é um passo na aproximação de um sistema real.

Uma vez que o sistema tenha sido identificado e limitado, é necessário estabelecer um critério no qual um ponto candidato pode ser selecionado de modo a se obter o melhor desempenho possível; por exemplo, em muitos problemas de engenharia (este problema inclusive), o critério econômico é utilizado. O problema é que definir múltiplos critérios para um problema pode resultar no fato de que alguns entram em conflito; normalmente, na engenharia, custo e desempenho caminham em direções opostas, por exemplo. Dessa maneira, normalmente toma-se um critério como primário e todos os outros se tornam critérios secundários.

O terceiro elemento na concepção de um problema de otimização é a escolha de variáveis independentes que possam caracterizar possíveis soluções para o sistema. Primeiramente, é necessário distinguir variáveis que podem mudar daquelas cujos valores são fixos devido a fatores externos, estando além dos limites dados do sistema em questão. Além disso, é importante diferenciar parâmetros fixos do sistema daqueles que são sujeitos a flutuações influenciadas por fatores externos não-controláveis. Segundamente,

é importante incluir todas as variáveis importantes que influenciam o desempenho do sistema ou afetam a definição do modelo. Finalmente, é necessário se considerar o nível de detalhamento no qual o sistema se encontra; embora seja importante tratar todas as variáveis independentes importantes, também é necessário que o problema não seja dificultado devido à inclusão de um número muito grande de detalhes de importância menor. Uma boa regra nesse último quesito é selecionar apenas variáveis que tenham impacto significativo no critério de desempenho do sistema estudado.

Por fim, o próximo passo na formulação de um problema de otimização para engenharia é construir o modelo que descreve como as variáveis do problema se relacionam e de que modo o critério de desempenho é afetado pelas variáveis independentes. Modelos são utilizados por que é caro, demorado ou arriscado usar o sistema real no estudo; logo, modelos são usados por oferecerem a maneira mais rápida e barata de se estudar os efeitos de mudanças das variáveis essenciais no desempenho geral do sistema. No geral, o modelo é composto de equações básicas de conservação de matéria e energia, relações de engenharia e equações de propriedades físicas que descrevem fenômenos físicos presentes no sistema estudado; essas equações são suplementadas por inequações que definem regiões de operação, especificam restrições de desempenho máximas ou mínimas, ou estabelecem limites de disponibilidade de recursos. Portanto, o modelo consiste de todos os elementos que devem ser considerados ao se prever o desempenho de um sistema de engenharia.

Capítulo 3

Proposta de Metodologia para o *Retrofitting* de Robôs Industriais

3.1 Introdução

Capítulo 4

Retrofitting do Robô ASEA IRB6-S2

4.1 Introdução

Capítulo 5

Conclusões

5.1 Conclusões do Trabalho

5.2 Sugestões para Trabalhos Futuros

REFERÊNCIAS BIBLIOGRÁFICAS

AGUIRRE, L. A. *Introdução à identificação de sistemas: técnicas lineares e não lineares: teoria e aplicação*. Belo Horizonte, MG: Editora UFMG, 2015.

GULLER, O. *Foundations of Optimization*. New York: Springer Science & Business Media, 2010. ISBN 978-0-387-68407-9.

IZMAILOV, A.; SOLODOV, M. *Otimização volume 1. Condições de otimalidade, elementos de análise convexa e de dualidade*. Rio de Janeiro, RJ: IMPA, 2005. ISBN 85-244-0238-5.

REKLAITIS, G. V. *Engineering Optimization: Methods and Applications*. New York: John Wiley, 1983. ISBN 0-471-05579-4.

ROSA, A. J.; CARVALHO, R. de S.; XAVIER, J. A. D. *Engenharia de Reservatórios de Petróleo*. Rio de Janeiro: Interciência: PETROBRAS, 2006. ISBN 85-7193-135-6.

THOMAS, G. *Cálculo, volume 1*. 11th. ed. São Paulo, SP: Addison Wesley, 2009. ISBN 978-85-88639-31-7.

THOMAS, G. *Cálculo, volume 2*. 11th. ed. São Paulo, SP: Addison Wesley, 2009. ISBN 978-85-88639-36-2.

THOMAS, J. E. et al. *Fundamentos de Engenharia de Petróleo*. 2. ed. Rio de Janeiro: Interciência: PETROBRAS, 2004. ISBN 85-7193-099-6.

YANG, X.-S. *Engineering optimization: an introduction with metaheuristic applications*. New York: Wiley, 2010. ISBN 978-0-470-58246-6.

APÊNDICES