

Volume 89, 2019

Editores

Alexandre Loureiro Madureira (Editor Chefe)

Laboratório Nacional de Computação Científica - LNCC
Petrópolis, RJ, Brasil

Amanda Liz Pacífico Manfrim Perticarrari

Universidade Estadual Paulista Júlio de Mesquita Filho - UNESP
Jaboticabal, SP, Brasil

Max Oliveira de Souza

Universidade Federal Fluminense - UFF
Niterói, RJ, Brasil

Eduardo V. O. Teixeira (Editor Executivo)

University of Central Florida - UCF
Orlando, FL, EUA

Lilian Markenzon

Universidade Federal do Rio de Janeiro - UFRJ
Rio de Janeiro, RJ, Brasil

Sandra Augusta Santos

Universidade Estadual de Campinas - UNICAMP
Campinas, SP, Brasil

A Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC publica, desde as primeiras edições do evento, monografias dos cursos que são ministrados nos CNMAC.

Para a comemoração dos 25 anos da SBMAC, que ocorreu durante o XXVI CNMAC em 2003, foi criada a série **Notas em Matemática Aplicada** para publicar as monografias dos minicursos ministrados nos CNMAC, o que permaneceu até o XXXIII CNMAC em 2010.

A partir de 2011, a série passa a publicar, também, livros nas áreas de interesse da SBMAC. Os autores que submeterem textos à série Notas em Matemática Aplicada devem estar cientes de que poderão ser convidados a ministrarem minicursos nos eventos patrocinados pela SBMAC, em especial nos CNMAC, sobre assunto a que se refere o texto.

O livro deve ser preparado em **Latex (compatível com o MikTeX versão 2.9)**, as **figuras em eps** e deve ter entre **80 e 150 páginas**. O texto deve ser redigido de forma clara, acompanhado de uma excelente revisão bibliográfica e de **exercícios de verificação de aprendizagem** ao final de cada capítulo.

Veja todos os títulos publicados nesta série na página
http://www.sbmac.org.br/p_notas.php

PROBLEMAS INVERSOS EM PESQUISA ESPACIAL

Haroldo Fraga de Campos Velho
haroldo.camposvelho@inpe.br

Laboratório Associado de Computação e Matemática Aplicada – LABAC
Coordenação de Laboratórios Associados – COCTE
Instituto Nacional de Pesquisas Espaciais – INPE



Sociedade Brasileira de Matemática Aplicada e Computacional

São Carlos - SP, Brasil
2019

Coordenação Editorial: Luiz Mariano Carvalho

Coordenação Editorial da Série: Alexandre L. Madureira

Editora: SBMAC

Capa: Matheus Botossi Trindade

Patrocínio: SBMAC

Copyright ©2008 by Haroldo Fraga de Campos Velho

Catálogo elaborado pela Biblioteca do IBILCE/UNESP
Bibliotecária: Maria Luiza Fernandes Jardim Froner

Campos Velho, Haroldo Fraga de
Problemas Inversos em Pesquisa Espacial – São Carlos, SP :
SBMAC, 2019, 100 p., 21.5 cm - (Notas em Matemática
Aplicada; v. 89)

ISBN 978-85-8215-092-4 / e-ISBN 978-85-8215-093-1

1. Problemas inversos 2. Soluções regularizadas 3. Métodos
implícitos 4. Redes neurais artificiais 5. Ciência Espacial
6. Engenharia Espacial 7. Aplicações Espaciais
I. Campos Velho, Haroldo Fraga. IV. Título. V. Série

CDD - 51

Agradecimentos

Durante os anos de pesquisa na área, muitos colegas do INPE e de outras instituições estão associados ao trabalho desenvolvido: Alexandre Martinez (USP, São José do Rio Preto), Atair Rios Neto (INPE e Embraer), Andrea S. Charão (Computação, UFSM), Antônio José da Silva Neto (IPRJ-UERJ), Antônio de Pádua Braga (Engenharia Eletrônica, UFMG), Cynthia F. Segatto (Instituto de Matemática, UFRGS), Domenico Anfossi (ISAC-CNR), Ezzat S. Chalhoub (LABAC-INPE), Fernando Manuel Ramos (LABAC-INPE), Gervásio A. Degrazia (Física, UFSM), João P. Braga (Instituto de Química, UFMG), João Batista C. Silva (Geofísica, UFPA), José Demísio S. da Silva (LAC-INPE), Juliana K. Vizzotto (CRS-INPE), Julio Cezar Ruiz Claeysen (Instituto de Matemática, UFRGS), Liliane B. Barichello (Instituto de Matemática, UFRGS), Mariângela Amendola (FEAGRI-UNICAMP), Marco Tulio de Vilhena (Instituto de Matemática, UFRGS), Milton Kampel (OBT-INPE), Nandamudi L. Vijaykumar (LABAC-INPE), Nelson J. Ferreira (CPTEC-INPE), Paolo Gasbarri (Depto. de Engenharia Aero-espacial, Universidade de Roma “La Sapienza”), Pedro P. B. de Oliveira (Universidade MacKenzie), Stephan Stephany (LABAC-INPE), Valéria Cristina F. Barbosa (ON-MCT); bem como estudantes de pós-graduação (atuais ou já formados) que trabalham (ou trabalharam) sob minha supervisão ou junto ao grupo de problemas inversos do INPE: Ana Paula C. Cuco, Alexandre G. Nowosad, Cláudio Faria, Cristiano Strieder, Débora R. Roberti, Eduardo F. P. da Luz, Elcio H. Shiguemori, Fabiana F. Paes, Fabiano L. de Sousa, Fabrício P. Härter, Helaine Cristina M. Furtado, João C. Carvalho, Leonardo D. Chiwiacowsky, Marcelo R. de Moraes, Mario R. Retamoso, Renata S. da Rocha Ruiz, Ricardo Varela Correa, Roberto Affonso da Costa Júnior, Roberto L. Galski, Rodrigo Barcelos Platte, Rosângela S. C. Cintra, Sabrina B. M. Sambatti, Wagner B. Muniz. A todos, meu especial agradecimento. Finalmente quero agradecer ao prof. Paolo Santini (*in memoriam*, emérito da Universidade de Roma “La Sapienza”).

Um agradecimento importante à Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC) e a Comissão de Programa do Congresso Nacional de Matemática Aplicada e Computacional (CNMAC) por apoiar o tema para apresentação como um Mini-Curso do CNMAC-2008.

Aos órgãos de suporte à pesquisa: CNPq, CAPES, FINEP e FAPESP.

Conteúdo

Prefácio	xiii
1 Introdução	1
1.1 Problemas Inversos: Conceito e Classificação	2
1.2 Condição Inicial em Condução de Calor	5
2 Técnicas de Inversão	9
2.1 Técnica de Regularização	9
2.2 Regularização de Tikhonov	11
2.3 Regularização Entrópica	11
2.4 Regularização por Entropia Não-Extensiva	12
2.4.1 Teorema da Unificação da Regularização	13
2.5 Parâmetro de Regularização	14
2.5.1 Princípio da discrepância generalizado	15
2.6 Técnicas de Otimização	16
2.6.1 Método de Newton e quase-Newton (MNe, MQNe)	16
2.6.2 Método de Levenberg-Marquardt (MLM)	18
2.6.3 Método do Gradiente Conjugado (MGC)	18
2.6.4 Recozimento Simulado (SA)	19
2.6.5 Algoritmo Genético Epidêmico (EGA)	20
2.6.6 Otimização por Colônia de Formigas (ACO)	22
2.7 Método Variacional	23
2.8 Redes Neurais Artificiais (RNA)	26
2.8.1 Perceptron de Múltiplas Camadas (PMC)	26
2.8.2 Função de Base Radial (FBR)	27
3 Problemas Inversos em Ciência Espacial	29
3.1 Inversão Magneto-telúrica	29
3.1.1 Formulação do Modelo Direto	29
3.1.2 Formulação do Problema Inverso	32
3.1.3 Resultados Numéricos	34
3.2 Outras Aplicações de Entropia de Alta Ordem	36
4 Problemas Inversos em Engenharia Espacial	37
4.1 Detecção de Danos na Estação Espacial Internacional (EEI)	38
4.2 Modelo Direto	39
4.3 Análise Inversa	40
4.3.1 O Problema de Sensitividade	40
4.3.2 O Problema Adjunto e a Equação do Gradiente	41

4.3.3	O Método do Gradiente Conjugado	42
4.3.4	O Critério de Parada: Princípio da Discrepância	42
4.3.5	Algoritmo Genético Epidêmico	43
4.4	Deteção de Danos: Resultados Numéricos	43
5	Problemas Inversos em Aplicações Espaciais	51
5.1	Meteorologia: Estimção de Temperatura Atmosférica com dados de satélite	52
5.1.1	Inversão com Regularização	54
5.1.2	Inversão por Redes Neurais Artificiais	54
5.2	Oceanografia: Propriedades Óticas	56
5.2.1	O Método LTS_N	58
5.2.2	Identificação Multi-espectral de Propriedades Óticas	60
6	Considerações Finais	67
A	Conceitos e Propriedades Matemáticas	69
A.1	Conceitos Preliminares	69
A.2	A Alternativa de Fredholm	71
A.3	Operadores Compactos	72
A.4	Espectro de Operadores Compactos	72
B	Termo-estatística Não Extensiva: Propriedades	75
B.1	Definições	75
B.2	Propriedades	75
	Bibliografia	77

Prefácio

Estas notas foram elaboradas através de versões mais reduzidas de mini-cursos ministrados em encontros científicos e por 2 anos consecutivos foram temas na Escola de Verão em Computação Aplicada (ELAC), promovida e patrocinada pelo Laboratório Associado de Computação e Matemática Aplicada (LABAC) do Instituto Nacional de Pesquisas Espaciais (INPE) — parte dos *slides* do mini-curso *Problemas Inversos: Conceitos Básicos e Aplicações*, apresentado durante o IV Encontro de Modelagem Computacional do IPRJ/EURJ em 2001, está disponível na internet: www.lac.inpe.br/~haroldo/scip.html (ver também: referência [14]).

Este minicurso é dividido em 2 partes. Primeiramente, descrição de técnicas gerais para a solução de problemas inversos, como o método de regularização, onde vários operadores são introduzidos: regularizadores de Tikhonov e de entropia – e variantes recentes: entropia de alta ordem e entropia não-extensiva, apresentando o teorema da unificação da regularização baseado no princípio da máxima entropia não extensiva. Métodos para o cálculo do parâmetro de regularização são apresentados, bem como uma generalização do princípio da discrepância de Morozov. Outras técnicas de inversão são também descritas: método variacional e redes neurais artificiais.

Quando o problema inverso é formulado como um problema de otimização, há várias técnicas para solução do problema são apresentadas no texto, baseadas em técnicas determinísticas, estocásticas e híbridas.

A segunda parte do minicurso estará voltada a aplicações.

As metodologias de inversão são aplicadas em problemas da pesquisa espacial. Há quadro pilares que compõem a área espacial: ciência espacial, engenharia ou tecnologia espacial, aplicações espaciais e medicina espacial. Os exemplos trabalhados no minicurso se situam nos 3 primeiros ramos da área espacial, que se constituem os campos de atuação do INPE.

Há vários anos se estabeleceu no LABAC-INPE um grupo de problemas inversos. Tivemos o apoio da FAPESP, via projeto temático: *Iterative Approaches to Inverse Problems with an Emphasis on Evolutionary Algorithms* (Proc. 1996/07200-8). A área espacial tem grande demanda em problemas inversos, que foram e são grandes desafios científicos e tecnológicos.

O objetivo do minicurso é apresentar um conjunto de metodologias e mostrar, através de problemas práticos, como estas técnicas podem ser empregadas em problemas reais que respondem a desafios científicos. As aplicações são todas da área espacial, mas as técnicas de inversão podem ser empregadas em aplicações de outras áreas.

São José dos Campos, 07 de maio de 2008.

Haroldo F. de Campos Velho

Capítulo 1

Introdução

A ciência tem um duplo papel: responder a necessidades do ser humano e da sociedade; desenvolver e organizar o conhecimento. Entender os fenômenos naturais e culturais é o desafio da ciência. É hoje um consenso que as aplicações da ciência é aspecto essencial para o desenvolvimento da sociedade. Osvaldo Cruz, um dos nomes mais importantes da ciência brasileira, se pronunciava da seguinte maneira: *Meditai se só as nações fortes podem fazer ciência ou se é a ciência que as torna fortes.*

Parte do conhecimento científico é o estudo de problemas diretos e de problemas inversos. Como a teoria geral de problemas inversos se estabeleceu muito recentemente, é natural termos uma idéia mais precisa do seja o problema direto.

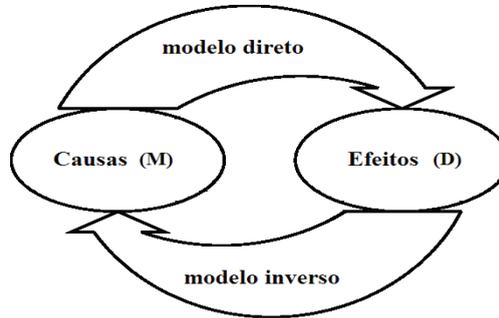
Observamos a luz e nos perguntamos: o que é a luz? Como se propaga? Ou ainda, como se dá o deslocamento de uma onda no mar, ou como se processa a dinâmica das órbitas dos corpos celestes? Tentar dar respostas a questões como estas é o desafio da ciência. Um marco histórico na ciência foi a edição do livro de Isaac Newton: *Princípios Matemáticos de Filosofia Natural* (o texto foi escrito em latim: *Philosophiae Naturalis Principia Mathematica*, e a 1a. edição data de 1687 — na verdade, são 3 livros: Livro I: *Do Movimento dos Corpos*, Livro II: *Do Movimento dos Corpos aos quais é oferecida Resistência na Razão da Velocidade*, Livro III: *Do Sistema do Mundo*). Há uma tradução para o português do *Principia* [62]. Newton muda, ou consolida, para sempre muitos aspectos do pensamento científico. Uma delas tem consequência direta para a física: diferentemente do que propunha Aristóteles, a física para o planeta Terra, não difere da física para os corpos celestes. Mais ainda, as ciências naturais devem ser quantitativas, isto é, as quantidades envolvidas se relacionam de acordo com relações matemáticas. Através do desenvolvimento do cálculo diferencial e integral (em que Newton teve papel fundamental), as relações quantitativas se expressam em equações diferenciais e integrais.

Assim, houve um longo desenvolvimento até que pudéssemos descrever, com relativa precisão, as equações (diferenciais) da dinâmica dos fluidos, por exemplo, para finalmente entendermos qualitativamente e quantitativamente o deslocamento de uma onda no mar. Deste modo, a descrição das equações matemáticas, propriedades do material (equações constitutivas) e as condições iniciais e de contorno, constitui o que chamamos de *Problema Direto*.

Mas, o que é um *Problema Inverso*?

É atribuído a Oleg Mikailivitch Alifanov, importante pesquisador russo na área de problemas inversos (PI), a afirmação: *a solução de um problema inverso consiste em determinar causas baseado na observação dos seus efeitos* (citação disponível internet: www.me.ua.edu/inverse/whatis.html).

A distinção entre o que seja um problema direto ou inverso para um dado fenômeno,



M: Espaço de parâmetros (modelos) **D:** Espaço de dados (observações)

Figura 1.1: Representação de problema direto (**M**) \times problema inverso (**D**).

está ligada a nossa cultura, isto é, trata-se do que se interpreta como causa e efeito!

Do ponto de vista prático, convencionou-se chamar problema direto àquele em que o estudo antecedeu-se historicamente. Tal ambigüidade (em relação ao par: *direto/inverso*), pode ser exemplificada do seguinte modo, se o modelo matemático é expresso por $A(u) = f$, o modelo inverso pode ser representado por: $A^{-1}(f) = u$. Por outro lado, definindo-se $B \equiv A^{-1}$, o par problema direto-inverso torna-se: $B(f) = u \implies B^{-1}(u) = f$! Mais adiante no texto, vamos notar que nem sempre a operação inversa como descrita aqui é possível.

Como se percebe, a própria definição de PI pode apresentar controvérsias. Entretanto, neste texto, a observação de Oleg M. Alifanov é a base para a conceituação de PI. Uma definição, bem como classificações de PI, é apresentada na próxima seção.

1.1 Problemas Inversos: Conceito e Classificação

Considera-se o astrofísico georgiano Viktor Amazaspovich Ambartsumian (para mais informações sobre este cientista consulte o sítio na internet: <http://www.physastro.sonoma.edu/BruceMedalists/Ambartsumian/>) como aquele que cunhou a expressão problema inverso (PI). Uma definição bastante abrangente, porém, é apresentada no livro de Engl et al. [44]: “*Resolver um problema inverso é determinar causas desconhecidas a partir de efeitos desejados ou observados*”. Note-se que a área de projeto ótimo ou projeto inverso (*inverse design*) também está incluída nesta definição. Em geral, as observações são imprecisas (dados contaminados com ruídos ou erros experimentais) e incompletas. Diferentemente, problemas diretos requerem um conhecimento completo e preciso das causas para a determinação dos efeitos.

A Figura 1.1 mostra de maneira pictórica a relação entre problema direto e inverso. Causas, num modelo matemático, são as condições iniciais e de contorno, termo de fontes/sumidouro e propriedades do sistema (material). Efeitos são as propriedades calculadas a partir de um modelo matemático (direto), como o campo de temperatura, concentração de partículas, corrente elétrica e propriedades vetoriais, como campo de velocidade.

Para a classificação de problemas inversos, o tipo de *causa* a ser determinada pode ser usado para classificar um PI. Contudo, outras classificações são possíveis:

1. Quanto a *natureza matemática* do método: Explícito (inversão direta), Implícito;
2. Quanto a *natureza estatística* do método: Determinista, Estocástico;
3. Quanto a *natureza da propriedade* estimada: Condição inicial, Condição de contorno, Termo de fonte/sumidouro, Propriedades do sistema;
4. Quanto a *natureza da solução* (Beck): Estimação de parâmetros, Estimação de função;
5. Quanto a *natureza da dimensão do problema* (Silva Neto e Moura Neto):
 - Tipo-1 (PDF e PIf),
 - Tipo-2 (PD ∞ e PIf),
 - Tipo-3 (PD ∞ e PI ∞),
 - Tipo-4 (PDF e PI ∞),

Nota-se que, a classificação dos itens 1 e 2 estão ligada aos métodos de solução do PI. No item 3, como já mencionando, tipifica-se o problema inverso pela *causa* a ser determinada. A classificação do item 4 foi sugerida pelo prof. James V. Beck. Ao que parece, o autor desta classificação tinha em mente como estimativa de função a noção de função contínua, desta forma a determinação de coeficientes c_k da expansão de uma função $f(t) = \sum_k c_k \phi_k(t)$ caracterizar-se-ia como estimação de parâmetros e não como estimação de função. Por outro lado, estimar propriedades características e constantes de um sistema (como a condutividade térmica k de um material, ou a intensidade S de uma fonte pontual de calor estacionária), é um problema de estimação de parâmetros. Num caso, o vetor de parâmetros desconhecidos é $X = [c_1 \ c_2 \ \dots \ c_N]^T$, noutro caso $X = [k, \ S]^T$. Um caso distinto é estimar uma função amostrada: $f(t_i) = f_i \ (i = 1, 2, \dots, N)$, nesse caso: $X = [f_1 \ f_2 \ \dots \ f_N]^T$.

A classificação indicada no item 5 foi proposta recentemente [135] e está baseada na dimensão do modelo do fenômeno físico (problema direto: PD) e na dimensão da quantidade a ser estimada (problema inverso: PI) se finita (f) ou infinita (∞): a expressão dimensão infinita está ligada ao conceito de função contínua (ou contínua por partes): esta classificação, entretanto, não é uma idéia de exclusão. Problemas inversos do tipo reconstrução de imagens são exemplos de PI do Tipo-1, estimação de parâmetros pode ser classificada como do Tipo-2 ou Tipo-3, enquanto estimação de função contínua é sempre um problema do Tipo-3.

É pertinente se fazer um comentário sobre o caso de estimação de função. Pode-se querer estimar a função, ela mesma, ou quando a função é obtida por uma expansão. Os dois casos de uma função real são listados a seguir:

$$\text{caso-1:} \quad f(x) = \sum_{k=0}^{\infty} c_k \phi_k(x);$$

$$\text{caso-2:} \quad f(x) \text{ com } x \in [0, 1].$$

Tanto no caso-1 como no caso-2, ter-se-á infinitos valores para se identificar. Entretando, no caso-1, os valores $\{c_k\}_{k=0}^{\infty}$ formam um conjunto enumerável, enquanto que no caso-2, a quantidade de valores a serem identificados formam um conjunto não enumerável. Em

ambos os casos, são conjuntos com uma quantidade infinita de elementos, mas a cardinalidade destes dois conjuntos são diferentes. Conjuntos enumeráveis (onde é possível de alguma forma uma associação com os números naturais), com uma quantidade infinita de elementos, tem cardinalidade \aleph_0 . O caso-2 é um exemplo de conjunto não enumerável, que tem cardinalidade $c = 2^{\aleph_0}$. George Cantor provou que $\aleph_0 \neq 2^{\aleph_0}$ [81]. Desta forma, um problema de identificação de parâmetros nunca é similar a um problema de identificação de função contínua.

Matematicamente, problemas inversos pertencem à classe de problemas malpostos. No início do século XX, o matemático francês Jacques Hadamard definiu um problema matematicamente bem-posto como sendo aquele que cumpre as três condições abaixo:

- (i) a solução existe;
- (ii) a solução é única;
- (iii) a solução tem uma dependência contínua (suave) com os dados de entrada.

Assim, o problema é dito mal-posto se alguma das condições acima não é satisfeita. Problemas discretos e finitos são chamados mal condicionados, se a condição (iii) não se cumpre. Em geral, nenhuma das condições de Hadamard é satisfeita num problema inverso!

Exemplos simples podem ser usados para ilustrar os conceitos acima. Por exemplo, considere a solução da equação do 1o. grau:

$$2x - 4 = 0. \quad (1.1.1)$$

O problema (direto) algébrico acima tem solução única: $x = 2$. O problema algébrico inverso

$$ax + b = 0 \quad (1.1.2)$$

com $x = 2$, não apresenta solução única. O problema de estabilidade é exemplificado por uma equação algébrica do 2o. grau:

$$ax^2 - 2x + 1 = 0 \quad (1.1.3)$$

o qual para $a = 1$, possui as seguintes soluções: $x_1 = x_2 = 1$. Introduzindo um erro de 1% no coeficiente a , isto é, $a = 1, 01$ - a solução da Eq. (1.1.3) torna-se: $x_{1,2} = 1 \pm 0, 1i$, sendo i a unidade dos números imaginários. Ou seja, uma variação de 1% (ruído) no coeficiente a , faz com que a Eq. (1.1.3) não tenha mais solução no campo dos números reais.

Os exemplos acima são simples, mas servem para ilustrar que problemas malpostos são muito frequentes. Contudo, Hadamard acreditava que fenômenos naturais (problemas da física) sempre se configuram como problemas bem-postos. Nos seus estudos sobre equações diferenciais, esta convicção foi desenvolvida (por vezes problemas malpostos eram identificados com formulações que violavam as leis da termo-dinâmica). Embora tenha apresentado exemplos de problemas malpostos, Hadamard afirmou que um modelo matemático para um fenômeno físico **deve** ser bem posto [55, 56].

A convicção de Hadamard impregnou por décadas a comunidade matemática. Desta forma, os matemáticos para *ajudar* as outras comunidades científicas, passaram a investigar tenazmente as duas primeiras propriedades de um problema bem-posto

Seja como for, na metade do século XX ficou evidente que problemas malpostos são problemas comumente encontrados em aplicações da ciência e da tecnologia.

Mesmo sendo uma área em franco desenvolvimento, PI é um capítulo relativamente recente na ciência. Há legítimos PI que não eram reconhecidos como tais. Contudo, existem várias outras áreas da ciência que estão correlacionadas com esta nova área, seja pela natureza do objetivo de estudo, seja do ponto de vista metodológico. A lista a seguir apresenta áreas correlatas aos PI (esta lista não pretende ser completa):

- Identificação de sistemas;
- Controle ótimo em sistemas estocásticos;
- Álgebra linear computacional em problemas de posto incompleto;
- Processamento de imagens;
- Teoria de filtragem;
- Assimilação/iniciação de dados;
- Teoria da estimação.

Porque PI emergiram como uma nova área da ciência? Devido a sua importância científica, econômica, social e mesmo política.

Por exemplo, ondas de som constituem um dos resultados mais significativos a partir de mapas da radiação cósmica de fundo em microondas [64], ou RCFM, sendo uma previsão de modelos de instabilidade gravitacional da teoria da relatividade geral. Mapa de RCFM é um modelo matemático, onde o padrão de flutuações de temperatura é dado por uma expansão em harmônicos esféricos. Outro exemplo, o telescópio espacial Hubble levou cerca de 10 anos para ser construído e custou 4,7 bilhões de dólares¹: depois de seu lançamento, notou-se que as imagens produzidas não tinham a nitidez desejada (projetada). O problema apresentado na fabricação das lentes do telescópio foi solucionado por software (métodos matemáticos de PI [57]). Centenas de exemplos podem ser citados, mas não se pode deixar de citar a revolução na medicina (e na sociedade) desde a introdução da tomografia computadorizada [71, 100, 155], um clássico problema inverso. Para finalizar, lembramos a discussão sobre o aquecimento global: a metodologia de problemas inversos é até agora o único procedimento para identificar o ciclo bio-geo-químico dos gases do efeito estufa [72] – este tema tem despertado uma discussão mundial sobre o *aquecimento global* de efeito antrópico.

1.2 Condição Inicial em Condução de Calor

Aqui será tratado o problema de determinar a condição inicial de um problema de linear de condução de calor numa barra isotrópica e homogênea, sem fontes internas e considerando-se condições de contorno do tipo Neumann homogênea. O problema é formulado matematicamente como:

$$\frac{\partial \Theta}{\partial \tau} = k \frac{\partial^2 \Theta}{\partial x_L^2}, \quad \tau > 0, \quad x_L \in (0, L) \quad (1.2.4)$$

com as condições iniciais e de contorno:

$$\Theta(\tau, x_L)|_{\tau=0} = f(x_L); \quad \frac{\partial \Theta}{\partial x_L} = 0, \quad \text{em } x_L = 0, \quad x_L = L. \quad (1.2.5)$$

Na equação acima, num meio isotrópico e homogêneo, a condutividade térmica k é constante e definindo: $x \equiv x_L/L$, $T \equiv \Theta/\Theta_0$ e $t \equiv k\tau/L^2$ a Eq. (1.2.4) torna-se adimensional [97].

¹Custo original estimado em US\$ 400 milhões, mas o custo total foi estimado em US\$ 4,7 bilhões até o tempo do lançamento. O custo cumulativo foi estimado em cerca de US\$ 10 bilhões em 2010, em 20 anos de operação – fonte: Wikipedia (https://en.wikipedia.org/wiki/Hubble_Space_Telescope)

Desta forma, o *problema teste* considerado neste texto é o seguinte problema evolutivo de condução do calor numa barra com condições de contorno adiabáticas:

$$\begin{aligned}\frac{\partial T}{\partial t} &= \frac{\partial^2 T}{\partial x^2}, & t > 0, \quad x \in \Omega \equiv (0, 1) \\ \frac{\partial T}{\partial t} &= 0, & t > 0, \quad x = 0, \quad x = 1, \\ T(x, 0) &= f(x), & x \in [0, 1].\end{aligned}\tag{1.2.6}$$

A solução do problema direto para uma dada condição inicial $f(x)$ é obtida explicitamente usando separação de variáveis [106], para $(x, t) \in [0, 1] \times \mathbb{R}^+$:

$$T(x, t) = \sum_{m=0}^{+\infty} e^{-\beta_m^2 t} \frac{1}{N(\beta_m)} X(\beta_m, x) \int_{\Omega} X(\beta_m, x') f(x') dx', \tag{1.2.7}$$

onde $X(\beta_m, x)$ são as auto-funções associadas ao problema, β_m os auto-valores e $N(\beta_m)$ representa a *normalização da integral* (ou a *norma* [106]). Para as condições de contorno do problema, tem-se:

$$\begin{aligned}X(\beta_m, x) &= \cos(\beta_m x), & \text{para } m \in \mathbb{N} \cup \{0\} \\ \beta_m &= m\pi \\ N(\beta_m) &= \int_0^1 X^2(\beta_m, x) dx.\end{aligned}$$

A Eq. (1.2.7) pode ser escrita numa formulação integral:

$$T(x, t) = \int_0^1 K(x, x', t) f(x') dx', \tag{1.2.8}$$

onde o núcleo $K(x, x', t)$ é definido por:

$$K(x, x', t) = \sum_{m=0}^{+\infty} e^{-\beta_m^2 t} \frac{1}{N(\beta_m)} X(\beta_m, x) X(\beta_m, x'). \tag{1.2.9}$$

A representação acima requer que a temperatura inicial $f(x)$ seja uma função limitada satisfazendo as condições de contorno no intervalo $[0, 1]$ [20].

Se a temperatura $T(x, t)$ é conhecida para um tempo $t = \tau$ e usando a ortogonalidade das funções $X(\beta_m, x)$ a condição inicial pode ser obtida *analiticamente*.

Teorema 1.1. *Se a temperatura $T(x, t)$ é conhecida para um tempo $\tau > 0$ sobre todo o domínio espacial $x \in (0, 1)$, então a temperatura inicial $f(x)$ é dada por*

$$f(x) = \sum_{m=0}^{+\infty} \frac{e^{\beta_m^2 \tau}}{N(\beta_m)} X(\beta_m, x) \int_0^1 X(\beta_m, x') T(x', \tau) dx'. \tag{1.2.10}$$

Prova: ($\tau > 0$) A solução do problema direto é expressa como

$$T(x, \tau) = \sum_{m=0}^{\infty} c_m X(\beta_m, x) e^{-\beta_m^2 \tau}. \tag{1.2.11}$$

Multiplcando a Eq. (1.2.11) por $X(\beta_k, x)$ e integrando sobre o domíno espacial, obtém-se

$$\begin{aligned} \int_0^1 T(x', \tau) X(\beta_k, x') dx' &= \sum_{m=0}^{\infty} c_m e^{-\beta_m^2 \tau} \int_0^1 X(\beta_k, x') X(\beta_m, x') dx' \\ &= c_k e^{-\beta_k^2 \tau} N(\beta_k). \end{aligned}$$

Consequentemente, os coeficientes são calculados como

$$c_k = \frac{e^{\beta_k^2 \tau}}{N(\beta_k)} \int_0^1 T(x', \tau) X(\beta_k, x') dx'. \quad (1.2.12)$$

Aplicando a condição inicial ($t = 0$) e a expressão (1.2.12), obtém-se

$$\begin{aligned} f(x) &= \sum_{m=0}^{+\infty} c_m X(\beta_m, x) \\ &= \sum_{m=0}^{+\infty} e^{\beta_m^2 \tau} \frac{1}{N(\beta_m)} X(\beta_m, x) \int_0^1 X(\beta_m, x') T(x', \tau) dx'. \quad \square \end{aligned}$$

Para *testar* o resultado do teorema 1.1, podemos ir a um laboratório, instrumentar uma barra metálica (isto é, inserir sensores ao longo da barra para medir a temperatura em vários pontos) isolada nas extremidades e usar a fórmula:

$$f(x) = \sum_{m=0}^N e^{\beta_m^2 \tau} \frac{1}{N(\beta_m)} X(\beta_m, x) \int_0^1 X(\beta_m, x') \theta^\delta(x', \tau) dx' \quad (1.2.13)$$

onde a temperatura medida $\theta(x, \tau)$ difere da temperatura verdadeira $T(x, \tau)$ somente por um pequeno erro experimental

$$\theta^\delta(x, \tau) = T(x, \tau) + \delta. \quad (1.2.14)$$

Pode-se considerar que o desvio aleatório δ seja de no máximo 1%. O resultado da inversão sob estas condições é mostrada na figura 1.2 (ver também [124]). Fica claro que o resultado da inversão *analítica* não funcionou. A razão é que este problema inverso viola a 3a. condição de Hadamard.

Empregando o resultado do teorema 1.1, pode-se mostrar que o problema de determinar a condição inicial é *mal-posto* [96, 97].

Teorema 1.2. *O problema de determinar a condição inicial f , onde $(f, T) \in L_2([0, 1]) \times L_2([0, 1])$, não depende continuamente (na métrica- L_2) sobre os dados T (temperatura transiente), isto é, o problema é mal-posto no sentido de Hadamard, visto que a exigência de estabilidade não é satisfeita.*

Prova: Considere duas condições iniciais $f_1, f_2 \in L_2([0, 1])$ tal que $f_2(x) = f_1(x) + CX(\beta_n, x)$, com $C \in \mathbb{R} - \{0\}$ e $n \in \mathbb{N}$. Assuma que a correspondente solução transiente (para um $\tau > 0$ fixo) são respectivamente os perfis $T_1(x, \tau), T_2(x, \tau)$. Pela linearidade, ver Eq. (1.2.7), tem-se que

$$\begin{aligned} T_2(x, \tau) &= T_1(x, \tau) + \sum_0^{\infty} \frac{e^{-\beta_m^2 \tau} X(\beta_m, x)}{N(\beta_m)} \int_0^1 CX(\beta_n, x') X(\beta_m, x') dx' \\ &= T_1(x, \tau) + Ce^{-\beta_n^2 \tau} X(\beta_n, x) \end{aligned}$$

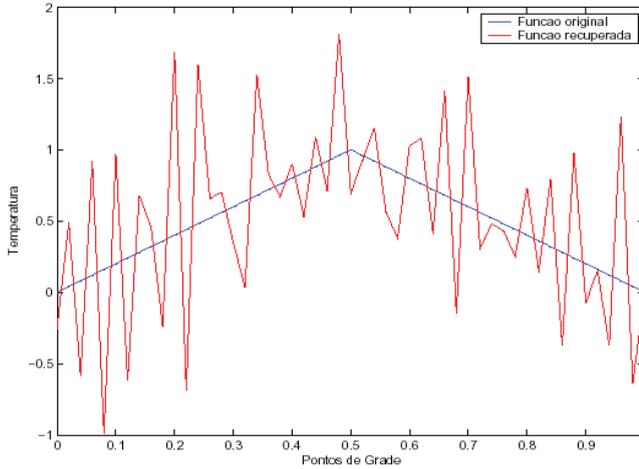


Figura 1.2: Solução calculada com a expressão (1.2.13).

Consequentemente, a diferença entre $T_1(x, \tau)$ e $T_2(x, \tau)$ é dada por

$$\|T_2 - T_1\|_2^2 = \int_0^1 [T_2(x, \tau) - T_1(x, \tau)]^2 dx = C^2 e^{-2\beta_n^2 \tau} N(\beta_n)$$

Então, para qualquer número C , a quantidade $\|T_2 - T_1\|_2$ pode ser feita arbitrariamente pequena, pela escolha de n suficientemente grande.

Analogamente, avaliando-se a diferença entre f_1 e f_2 na métrica- L_2 com $C \neq 0$, obtém-se

$$\|f_2 - f_1\|_2^2 = \int_0^1 C^2 X^2(\beta_n, x) dx = C^2 N(\beta_n) = \text{constante} > 0.$$

Consequentemente, com discrepâncias arbitrariamente pequenas entre T_1 e T_2 , pode-se escolher n e C de tal modo que a discrepância entre as soluções inversas correspondentes (f_1 e f_2) pode ser arbitrária, isto é:

$$\|T_1 - T_2\|_2 \rightarrow 0, \text{ mas } : \|f_1 - f_2\|_2 \not\rightarrow 0. \quad \square$$

O resultado não surpreende. Uma propriedade conhecida como a *alternativa de Fredholm* [45, 76] garante a solução para equações integrais do tipo (1.2.8). De forma mais precisa, a alternativa de Fredholm somente aplica-se quando o operador integral é um *operador compacto*. Dito de outra maneira, da teoria de Fredholm, núcleos suaves de equações integrais são operadores compactos. Tais operadores têm um espectro que é um sub-conjunto finito (inclui o número zero) ou infinito (mas *contável*), de números complexos (no último caso, tem o zero como seu *ponto limite*). Deste modo, da teoria espectral dos operadores compactos, é possível reconhecer que o operador inverso de Eq. (1.2.8) irá apresentar singularidades (uma descrição mais detalhada é feita no apêndice A).

No próximo capítulo, veremos algumas técnicas para calcular *boas inversões*, evitando as singularidades na operação de inversão.

Capítulo 2

Técnicas de Inversão

Como já mencionado, problemas inversos formam um tópico muito importante na ciência e tecnologia. Desta forma, há um esforço continuado no desenvolvimento teórico e de métodos para encontrar a solução desta classe de problemas.

Neste capítulo, técnicas de inversão usadas nas aplicações expostas neste texto são aqui explicitadas. Inicialmente a teoria de regularização é apresentada, formulada como um problema de otimização com restrições. Alguns operadores de regularização são descritos, bem como o *teorema da unificação* dos operadores de Tikhonov e de entropia. A teoria da regularização está completa com o cálculo do parâmetro de regularização e com métodos de solução do problema de otimização – quatro técnicas determinísticas e três estocásticas são detalhadas.

Por último, duas técnicas de solução de problemas inversos são descritas: método variacional e de redes neurais artificiais. Dois tipos de redes neurais são apresentados: perceptron de múltiplas camadas (PMC) e função de base radial (RBF) – está também incluída a rede RBF não-extensiva, desenvolvida no LABAC-INPE.

2.1 Técnica de Regularização

Como visto no capítulo anterior, mesmo tendo sucesso na busca de uma solução analítica para um problema inverso (outro problema resolvido explicitamente é a solução de Burgraff para identificação de fluxo superficial em problemas de condução de calor – ver Beck et al. [5], página 67), a solução inversa pode não ser estável em relação a presença de ruído nos dados (efeito de inserção de informação não ligada ao fenômeno observado ou efeito numéricos, como erro de truncamento de uma série ou mesmo o erro de arredondamento). Assim, mesmo que o problema direto seja bem-posto, com operadores suaves e contínuos (se for um operador de Fredholm com núcleo de quadrado integrável, por exemplo, implica num operador compacto), a operação inversa não é contínua [53]:

$$A(u) = f \not\Rightarrow u = A^{-1}(f).$$

No caso particular de alguns sistemas lineares indeterminados (com número de incógnitas inferior ao número de equações), uma solução estudada foi a abordagem de mínimos quadrados de norma mínima, isto é, existe um termo adicional (a norma do vetor de incógnitas) associado ao termo de discrepância quadrática. Isso pode ser generalizado. A idéia é a seguinte: para resolver problemas malpostos, é necessário fornecer informação adicional. Na década de 60, vários pesquisadores também notaram este fato. Nomes como V. K. Ivanov [67], D. L. Phillips [109] e S. Twomey [153] merecem destaque, mas foi com

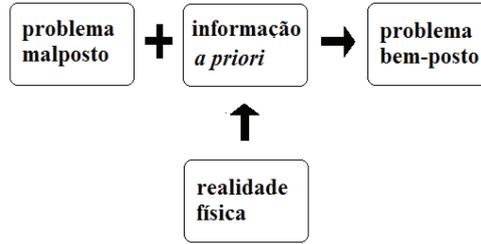


Figura 2.1: Idéia básica para solução de problemas mal-postos: informação adicional (*a priori*) é incorporada ao problema e o problema modificado (muito similar ao problema original) é um problema bem-posto!

o trabalho de Andrei Nikolaevich Tikhonov em 1963, o início de uma formulação geral para problemas malpostos, chamada teoria ou método de regularização. O prof. Tikhonov foi um matemático proeminente e trabalhou no prestigioso Instituto de Matemática Steklov da Academia Russa de Ciências (matemáticos como A. N. Krylov, D. K. Faddeev, L. S. Pontryagin, S. L. Sobolev, A. N. Kolmogorov e A. A. Markov, formam uma pequena lista de importantes cientistas do mesmo Instituto), tendo trabalhos importantes em topologia, análise funcional, matemática computacional e física-matemática.

O método da regularização consiste na determinação da solução aproximada e suave, compatível com os dados de observação, para certo nível de ruído. A busca da solução mais suave (regular) é uma informação adicional, que transforma o problema mal-posto num problema bem-posto (ver Figura 2.1).

Na implementação matemática do método, o PI é formulado como um problema de otimização com restrições:

$$\min_{u \in U} \|A(u) - f^\delta\|^2 \text{ sujeito a } \Omega[u] \leq \rho \quad (2.1.1)$$

onde $A(u) = f^\delta$ representa o modelo direto e $\Omega[u]$ é o operador de regularização [150]. A técnica dos multiplicadores de Lagrange permite colocar na mesma função custo os objetivos de fidelidade dos parâmetros com o modelo direto – primeiro termo da expressão abaixo – e de regularidade (ou, suavidade) – segundo termo da expressão abaixo - exigida da quantidade desconhecida:

$$\min_{u \in U} \left\{ \|A(u) - f^\delta\|_2^2 + \alpha \Omega[u] \right\} \quad (2.1.2)$$

onde α é o parâmetro de regularização. Note que, para $\alpha \rightarrow 0$, o termo de fidelidade dos dados na função objetivo é superestimado, enquanto que para $\alpha \rightarrow \infty$ toda a informação contida no modelo matemático é perdida.

Uma definição mais formal considera-se a aplicação $A(u) = f$ entre espaços normados U e F , onde um operador inverso modificado aproxima-se da operação de inversão, mas resolve a inversão de modo estável [53, 97].

Definição 2.1. Família de operadores de regularização: Uma família de operadores contínuos $R_\alpha : F \rightarrow U$ é chamada um esquema de regularização para a operação inversa de: $A(u) = f$, quando

$$\lim_{\alpha \rightarrow 0} R_\alpha \{A(u)\} = u \quad (2.1.3)$$

para $u \in U$ e α é o parâmetro de regularização.

A expressão (2.1.2) mostra uma implementação prática para a definição 2.1. Porém para completar a teoria é necessário indicar o operador de regularização e como calcular o parâmetro de regularização (α). Vários tipos de operadores de regularização têm sido investigados desde o trabalho dos pioneiros na área. Aqui somente serão comentadas três classes destes operadores de regularização.

Técnicas estatísticas para resolver problemas inversos em geral se baseiam na regra de Bayes, onde a função de distribuição de probabilidade *a posteriori* é obtida a partir do produto de uma função de máxima verossimilhança (é a distribuição da discrepância entre os dados de medida e os dados computados pelo modelo matemático) e de uma distribuição probabilidades *a priori* do vetor de parâmetros desconhecidos [149]. Sob certas circunstâncias (hipótese de distribuições gaussianas e discrepância avaliada pelo quadrado da norma-2) o resultado da estimação por máxima verossimilhança produz resultado equivalente ao resultado obtido com regularização de Tikhonov de ordem-zero – ver Seção 1.10 de [8] e [102].

2.2 Regularização de Tikhonov

Aqui o operador de regularização é expresso por:

$$\Omega[u] = \sum_{k=0}^P \mu_k \|u^{(k)}\|_2^2 \quad (2.2.4)$$

onde $u^{(k)}$ denota a k -ésima derivada (diferença) e $\mu_k > 0$. Em geral, $k = \delta_{kj}$ (delta de Kronecker), e o operador torna-se

$$\Omega[u] = \|u^{(k)}\|_2^2 \quad (2.2.5)$$

e a técnica é chamada de regularização de Tikhonov de ordem- k . O efeito da regularização de Tikhonov-0 é reduzir oscilações na função u (busca por funções suaves, neste caso $u^{(0)} \equiv u \approx 0$). Já na regularização de 1a. ordem, toma-se $u^{(1)} \approx 0$, ou seja, u é aproximadamente constante.

O problema inverso da Seção 1.2 foi resolvido usando este operador de regularização por Muniz et al. [96].

2.3 Regularização Entrópica

De maneira similar é técnica de Tikhonov, o método da máxima entropia busca regularidade global, produzindo as reconstruções mais suaves de acordo com os dados disponíveis (medidos).

O princípio da máxima entropia foi proposto por Jaynes [68] como um critério geral de inferência, baseado na teoria matemática da informação de Shannon [127]. A Figura 2.2 mostra diferentes probabilidades num fenômeno em que 8 estados são possíveis. A condição de máxima entropia ocorre quando todos os estados são igualmente prováveis (Figura 2.2a), enquanto o estado de mínima entropia é ilustrado na Figura 2.2b: todos os estados têm probabilidade nula de ocorrer, com exceção de um único estado.

Da mesma forma que a regularização de Tikhonov, podem ser estabelecidos operadores de entropia de várias ordens. Uma expressão genérica para regularização entrópica é:

$$S(u) = \sum_{q=1}^{N_q} s_q \log s_q, \quad \text{com: } s_q = \frac{r^{(k)}}{\sum_{q=1}^{N_q} r^{(k)}} \quad (2.3.6)$$

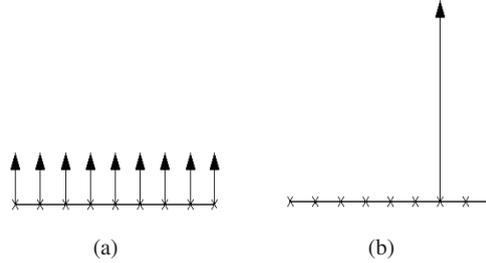


Figura 2.2: (a) Máxima entropia: estados igualmente prováveis, (b) mínima entropia: todos estados tem probabilidade nula de ocorrência exceto um único estado com probabilidade de 100% de ocorrência.

válida para o caso discreto, onde $u = [u_1 \ u_2 \ \dots \ u_q \ \dots \ u_{N_q}]^T$. Neste contexto, diferentes ordens do operador podem ser consideradas [12, 13, 114, 115]:

$$r_q^k = \begin{cases} u_q & \text{para } k = 0 \\ u_{q+1} - u_q + (u_{\max} - u_{\min}) + \varsigma & \text{para } k = 1 \\ u_{q+1} - 2u_q + u_{q-1} + 2(u_{\max} - u_{\min}) + \varsigma & \text{para } k = 2 \end{cases} \quad (2.3.7)$$

onde $r_q^{(k)}$ representa a k -ésima diferença (derivada) da quantidade a ser estimada. A função de entropia atinge seu valor máximo, $S_{\max} = \log(N_q)$, quando a função de densidade de probabilidade é uniforme, e tem seu valor mínimo, $S_{\min} = 0$, se os valores de r_q estiverem distribuídos por uma delta de Dirac (ou delta de Kronecker, no caso discreto) [98].

Em termodinâmica macroscópica, a entropia é uma maneira de se quantificar a irreversibilidade nos sistemas e a entropia tem um papel central na física estatística. A entropia de Shannon é uma medida da informação perdida. No contexto de problemas inversos, ela é uma *entropia de forma*.

O problema inverso da Seção 1.2 foi resolvido usando regularização entrópica de ordem zero e de alta ordem [98].

2.4 Regularização por Entropia Não-Extensiva

Uma formulation não extensiva para a entropia foi proposta por Constatino Tsallis [151, 152]. Recentemente, a forma não-extensiva da entropia (S_q) foi usada como um novo operador regularização [116], usando somente $q = 0.5$. O parâmetro q tem um papel central na termo-estatística de Tsallis, em que para $q = 1$ a entropia Boltzmann-Gibbs-Shannon é recuperada. A entropia de Tsallis inclui como um caso particular a entropia extensiva ($q = 1$). Entretanto outro importante case particular da entropia não-extensiva ocorre quando $q = 2$. Neste caso, o *princípio da máxima entropia não-extensiva* para o operador de regularização S_2 é equivalente a regularização de Tikhonov [150] (ou regularização de Tikhonov de ordem-zero). Na próxima seção este fato será demonstrado: o *Teorema da Unificação*.

Este novo operador regularização foi testado para estimar a condição inicial do problema de condução do calor descrito na seção 1.2 [17]. Uma investigação para resolver este problema foi avaliada para vários valores de q [17, 128].

A forma não-extensiva da entropia proposta por Tsallis [151] é dada pela expressão¹:

$$S_q(\mathbf{p}) = \frac{k}{q-1} \left[1 - \sum_{i=1}^{N_p} p_i^q \right] \quad (2.4.8)$$

onde p_i é uma probabilidade e q é um parâmetro livre (o *parâmetro de não-extensividade*). Em termodinâmica o parâmetro k é conhecido com a constante de Boltzmann. Analogamente como na teoria matemática da informação, $k = 1$ é considerado na teoria de regularização. A entropia de Tsallis se reduz a forma padrão de Boltzmann-Gibbs-Shanon no limite $q \rightarrow 1$:

$$S_1(\mathbf{p}) = -k \sum_{i=1}^{N_p} p_i \ln p_i. \quad (2.4.9)$$

Para $q < 5/3$, o teorema central do limite se aplica, implicando que se p_i é escrito como uma soma de N_p variáveis aleatórias independentes, no caso limite $N_p \rightarrow \infty$, a função de densidade de probabilidade para p_i no espaço de distribuições é uma *distribuições normal* (gaussiana) [152]. Entretanto, para $5/3 < q < 3$ o teorema central do limite de Levy-Gnedenko's aplica-se, resultando que para $M \rightarrow \infty$ a distribuição de Levy é função de distribuição de probabilidade para a variável aleatória p_i . O índice de tal distribuição de Levy é $\gamma = (3 - q)/(q - 1)$ [152].

Tal como na forma extensiva da entropia, a condição de equiprobabilidade produz o máximo para a função da entropia não-extensiva. Esta condição é expressa como

$$(S_q)_{\max} = \frac{1}{1-q} (N_p^{1-q} - 1) \quad (N_p \geq 1) \quad (2.4.10)$$

onde, no limite $N_p \rightarrow \infty$, S_q diverge se $q \leq 1$ e satura em $1/(q - 1)$ se $q > 1$ [152].

A condição de equiprobabilidade leva a um operador de regularização definido por $S_q(\mathbf{p})$, dado pela Eq. (2.4.8), que irá identificar a solução mais suave para o vetor de parâmetros desconhecidos \mathbf{p} [17, 116].

O problema inverso da Seção 1.2 também foi resolvido usando entropia não-extensiva [17].

2.4.1 Teorema da Unificação da Regularização

Propriedades de regularização para a entropia emergem do critério de inferência de Jaynes: o princípio da máxima entropia, onde todos os estados são igualmente prováveis: $p_i = 1/N_p$. O seguinte lema estende este resultado para a entropia não-extensiva.

Lema 2.1. *A função de entropia não-extensiva S_q tem um máximo quando $p_i = 1/N_p$ para todo i .*

Prova: O problema é determinar o máximo da função (2.4.8) sob a seguinte restrição

$$\sum_{i=1}^{N_p} p_i = 1 \quad (2.4.11)$$

visto que p_i representa uma probabilidade. Consequentemente, é possível definir uma função objetivo em que a restrição acima é adicionada a função não-extensiva (2.4.8):

$$J(\mathbf{p}) = S_q(\mathbf{p}) + \lambda \left(\sum_{i=1}^{N_p} p_i - 1 \right) \quad (2.4.12)$$

¹Ver também Eqs. (B.1.1) e (B.1.2) no Apêndice B.

onde λ é o multiplicador de Langrange, que, neste caso, pode ser determinado quando um mínimo para função objetivo é calculado como segue:

$$\frac{\partial J}{\partial p_i} = -\frac{1}{q-1} \left(q p_i^{q-1} \right) + \lambda = 0 \Rightarrow p_i = \left[\frac{\lambda(q-1)}{q} \right]^{\frac{1}{q-1}} \quad (2.4.13)$$

Este resultado pode ser usado para obter os valores de p_i 's que maximizam a função $J(\mathbf{p})$:

$$\begin{aligned} \sum_{i=1}^{N_p} p_i &= \sum_{i=1}^{N_p} \left[\frac{\lambda(q-1)}{q} \right]^{\frac{1}{q-1}} = \left[\frac{\lambda(q-1)}{q} \right]^{\frac{1}{q-1}} N_p = 1 \\ \Rightarrow p_i &= \frac{1}{N_p} \end{aligned} \quad (2.4.14)$$

ou seja, se $p_i = 1/N_p$ para todo $i = 1, \dots, N_p$ a entropia não-extensiva é máxima.

O teorema a seguir unifica os operadores de regularização de Tikhonov e de entropia extensiva [17].

Teorema 2.1. (Teorema da Unificação) Para os valores de $q = 1$ e $q = 2$ da entropia não extensiva S_q , o efeito de regularização deste operador é equivalente ao da regularização entrópica extensiva e de Tikhonov, respectivamente.

Prova: (i) $q = 1$: Tomando o limite,

$$\begin{aligned} \lim_{q \rightarrow 1} S_q(\mathbf{p}) &= \lim_{q \rightarrow 1} \frac{1 - \sum_{i=1}^{N_p} p_i^q}{q-1} = \lim_{q \rightarrow 1} \frac{1 - \sum_{i=1}^{N_p} e^{q \log p_i}}{q-1} \\ &= \lim_{q \rightarrow 1} \frac{-\sum_{i=1}^{N_p} \log p_i e^{q \log p_i}}{1} = \sum_{i=1}^{N_p} p_i \log p_i \end{aligned} \quad (2.4.15)$$

(ii) $q = 2$: Lembrando que $\max S_2$ é equivalente a $\min(-S_2)$, segue

$$\max S_2(\mathbf{p}) = \max \left(1 - \sum_{i=1}^{N_p} p_i^2 \right) \Leftrightarrow \min \{-S_2(\mathbf{p})\} = \min \left\{ \sum_{i=1}^{N_p} p_i^2 - 1 \right\}$$

agora, para os valores de máximo (mínimo) tem-se $\nabla_{\mathbf{p}} S_2 = 0$, conseqüentemente

$$\nabla_{\mathbf{p}} S_2(\mathbf{p}) = \nabla_{\mathbf{p}} \left(\sum_{i=1}^{N_p} p_i^2 - 1 \right) = \nabla_{\mathbf{p}} \left(\sum_{i=1}^{N_p} p_i^2 \right) = \nabla_{\mathbf{p}} \|\mathbf{p}\|_2^2 \quad (2.4.16)$$

Ou seja, $\max S_2(\mathbf{p}) = \min \|\mathbf{p}\|_2^2$: regularização de ordem-zero de Tikhonov. \square

2.5 Parâmetro de Regularização

Para completar a teoria de regularização, é necessário um método para calcular o parâmetro α na Eq. (2.1.2). O parâmetro de regularização realiza o balanço entre o termo da diferença quadrática entre os dados de observação e o modelo matemático (termo de *fidelidade dos dados*) e o termo de regularização (termo de *suavidade dos dados*). A literatura registra vários métodos para a determinação do multiplicador de Langrange [7]: critério da discrepância de Morosov, o método da curva L , validação cruzada.

O critério de Morosov baseia-se no fato de que a diferença, ou discrepância, entre os dados de observação e os dados do modelo deve ter a mesma magnitude do erro de medida. Dessa maneira, se δ é o erro no processo de medida, α^* é a raiz da equação

$$\{\|A(u) - f^\delta\|_{\alpha^*}^2\} = \delta \quad (2.5.17)$$

O critério da discrepância foi validado para o caso do operador de entropia de alta ordem [98] e para a entropia não-extensiva de Tsallis [17, 128].

O critério da curva L é um critério *geométrico* introduzido por Hansen [58]. A idéia aqui é encontrar o ponto de máxima curvatura da curva construída para o gráfico: $\Omega[u_\alpha] \times \|A(u_\alpha) - f^\delta\|_2^2$, que em geral apresenta uma forma tipo- L . Este critério também foi examinado para a regularização com a entropia não-extensiva [17, 128].

2.5.1 Princípio da discrepância generalizado

Se as medidas são tomadas por um mesmo instrumento e são independentes e, ainda, os erros de medidas podem ser expressos por uma estatística gaussiana, o princípio da discrepância pode ser expresso como:

$$\|A(u) - f^\delta\|^2 \approx N \sigma^2 \quad (2.5.18)$$

onde N é o número de parâmetros a serem estimados e σ^2 é a variância da distribuição gaussiana dos erros de medida (f^δ).

A hipótese implicitamente assumida é que a distribuição associada é uma distribuição normal, com média zero e desvio padrão σ . Esta é uma hipótese geralmente aceita e está embasada num dos teoremas mais fortes da estatística: o teorema central do limite².

Como fica o critério se a distribuição dos erros aleatórios da medida f^δ forem uma distribuição com o segundo momento estatístico não definido?

Há muitas distribuições estatísticas importantes que não tem o segundo momento (a variância) definido. Exemplos são as distribuições de Cauchy e de Lewi. Para casos deste tipo, uma generalização do critério de Morosov pode ser proposta.

Seja $\rho_f(x)$ a distribuição da variável aleatória δ , que está associada ao erro de medida da variável f :

$$f^\delta = f + r * \delta \quad (2.5.19)$$

onde r é o erro máximo. O segundo momento estatístico é definido por:

$$\sigma_p^2 \equiv \int_{-\infty}^{+\infty} x^2 \rho(x) dx . \quad (2.5.20)$$

É sempre possível normalizar a distribuição para fixar $\sigma^2 = 1$, desde que a integral (2.5.20) convergir. No caso da integral acima divergir ($\sigma^2 \rightarrow \infty$), pode-se redefinir o domínio da variável aleatória δ , denotada por $\hat{\delta} \in \hat{\rho}_{[-d,d]}$, tal que

$$\int_{-d}^d x^2 \rho(x) dx = 1 . \quad (2.5.21)$$

Assim, para uma distribuição normalizada ($\sigma^2 = 1$), o princípio da discrepância é re-definido como:

$$\|A(u) - f^\delta\|^2 \approx N , \quad (2.5.22)$$

²O resultado é o seguinte: Se uma variável aleatória δ resulta de uma soma infinita de variáveis aleatórias independentes e com momentos estatísticos finitos, a distribuição da variável aleatória δ é uma distribuição gaussiana, *mesmo* que as distribuições das demais variáveis não o sejam.

que será válido para distribuições com segundo momento finito e unitário, ou para distribuições com segundo momento não definidos, mas com *domínios* re-definidos pela Eq. (2.5.21).

Esta *forma generalizada do princípio da discrepância* para distribuições em que a variância não é definida, foi testada para a determinação de condição inicial (problema da Seção 1.2), onde os erros de medida eram variáveis aleatórias com distribuições de Cauchy, *t*-Student e de Tsallis com $q = 1, 5$ [130]. A generalização proposta mostrou-se efetiva em todos os casos.

2.6 Técnicas de Otimização

A solução inversa no método de regularização é obtida resolvendo-se o problema de otimização expresso pela Eq. (2.1.2). Há uma grande variedade de métodos na literatura, divididos em 2 grandes grupos: métodos determinísticos e métodos estocásticos:

- Determinísticos: máxima descida, método de Newton, quase-Newton, Gradiente Conjugado, Método de Levenberg-Marquadt, Método Simplex.
- Estocásticos: Recozimento simulado (SA: *Simulated Annealing*), Algoritmos genéticos (GA: *Genetic Algorithm*), busca Tabu, otimização extrema, otimização por colônia de formigas (ACO: *Ant Colony Optimization*), otimização por enxame de partículas (SPO: *Swarm Particle Optimization*).
- Métodos híbridos: Combinam a estratégia de busca global dos métodos estocásticos com busca local dos métodos deterministas (GAPlex: GA + Simplex, SAPlex: SA + Simplex, GA-CG: GA + Gradiente Conjugado).

2.6.1 Método de Newton e quase-Newton (MNe, MQNe)

O objetivo de um método de otimização é encontrar o mínimo/máximo de uma função escalar $J(X)$ sendo X um vetor de um espaço de dimensão m ($X \in \mathfrak{R}^m$). Porém, determinar o mínimo da função $J(X)$ é equivalente a encontrar o ponto X tal que $\nabla J(X) = 0$. O gradiente de uma função escalar é um vetor. Denotando este gradiente por $Q(X) \equiv \nabla J(X)$, a série de Taylor da função vetorial $Q(X)$ é dada por:

$$Q(X + \Delta X) = Q(X) + \nabla Q(X) \cdot \Delta X + \frac{1}{2!} \nabla[\nabla Q(X) \cdot \Delta X] \cdot \Delta X + O(\Delta X^3) \quad (2.6.23)$$

Realizando truncamento de 1a. ordem na expressão acima e adotando-se a aproximação: $(X + \Delta X) \approx 0$, chega-se a relação de recorrência do método de Newton ($\Delta X = X_{k+1} - X_k$):

$$X_{k+1} = X_k + [\nabla \nabla J(X_k)]^{-1} \nabla J(X_k) \quad (2.6.24)$$

Do ponto de vista computacional, a inversa da matriz hessiana da função $F(X)$ não é calculada explicitamente, ao invés disso resolve-se o sistema associado de equações lineares e avalia-se sua convergência – a precisão é definida previamente por uma quantidade ϵ . De forma algorítmica, o método de Newton pode ser descrito da seguinte forma:

1. Resolver o sistema linear de equações algébricas:

$$[\nabla \nabla J(X_k)] \Delta X = -\nabla J(X_k) ;$$

2. Atualizar o vetor de parâmetros: $X_{k+1} = \Delta X + X_k$;

3. Avaliação de convergência:

$$\frac{\|X_{k+1} - X_k\|}{\|X_{k+1}\|} \leq \epsilon \quad \begin{cases} \text{Sim: processo finaliza ,} \\ \text{Não: retorna ao passo-1 .} \end{cases}$$

Em notação indicial, a matriz hessiana ($N \times N$) é calculada como segue, onde o vetor de parâmetros desconhecidos de dimensão N é denotado por $X = [X_1, X_2, \dots, X_N]^T$:

$$[\nabla \nabla J(X)]_{ij} = \frac{\partial^2 J}{\partial X_i \partial X_j} .$$

O cálculo da matriz hessiana, em geral, é o processo computacionalmente mais custoso (pode-se usar técnicas numéricas para o cálculo da matriz). Várias estratégias são propostas para minimizar o custo computacional, evitando o cálculo da matriz hessiana. Estes são os métodos quase-Newton, ou ainda, métodos de métrica variável, com muitas vezes é denominado na literatura [82]. O processo implementado na rotina E04UCF da biblioteca numérica NAG [43] é o seguinte:

1. Resolver o problema direto para X_k e calcula o valor da função objetivo $J(X_k)$.
2. Calcular o gradiente por diferenças finitas: $\nabla J(X_k)$.
3. Calcular a aproximação quase-Newton positiva-definida H_k para a matriz hessiana ($\nabla \nabla J$):

$$H_{k+1} = H_k + \frac{b_k(b_k)^T}{b_k u_k} - \frac{H_{k-1} u_k (u_k)^T H_{k-1}}{(u_k)^T H_{k-1} u_k}$$

onde

$$\begin{aligned} b_k &\equiv X_k - X_{k-1} = \Delta X_k \\ u_k &\equiv \Delta J(X_k) - \Delta J(X_{k-1}) . \end{aligned}$$

4. Calcular a direção de busca d^k como a solução do seguinte sub-problema de programação quadrática:

$$\text{Minimize } [\Delta J(X_k)]^T d^k + \frac{1}{2} (d^k)^T H_k d^k ,$$

$$\text{sujeito a: } l_q - p_q^k \leq d_q \leq u_q^k - p_q^k .$$

5. Calcular: $X_{k+1} = X_k + \beta^k d^k$, onde o passo de comprimento β^k minimiza $J(X_k + \beta^k d^k)$.
6. Testar a convergência: finalizar ou retornar ao Passo 1.

É possível demonstrar que a solução ótima do processo acima, converge para solução do método de Newton [82]. Antes de finalizar esta seção, é pertinente informar que o método de Newton tem uma convergência quadrática, porém pode não convergir, dependendo se a estimativa inicial X_0 estiver fora de uma região de convergência (região próxima ao máximo/mínimo). Um método mais robusto em relação é escolha da estimativa inicial é o método da máxima descida, que tem convergência linear [82].

O método quase-Newton da rotina E04UCF (NAG) foi aplicado ao problema de otimização empregando regularização de Tikhonov, entrópia extensiva e não extensiva de várias ordens para o problema inverso da Seção 1.2 [17, 27, 96, 98].

2.6.2 Método de Levenberg-Marquardt (MLM)

Conforme mencionado anteriormente, os métodos de Newton e quase-Newton podem apresentar problemas de convergência: além disso, a matriz associada no algoritmo de solução do método de Newton pode ser singular ou quase singular – que pode apresentar características de uma matriz singular, devido às aproximações numéricas. O método da máxima descida, por sua vez, além de apresentar uma convergência lenta, não tem uma estratégia definida para a escolha da primeira estimativa do processo iterativo. Estes problemas podem ser minorados pelo método de Levenberg [79] e Marquardt [86], que combina as estratégias de máxima descida com o método de Newton.

Denotando $J(X) \equiv A(X) - f^\delta$, onde $A(\cdot)$ representa o modelo direto, f^δ representa os dados experimentais e $Q \equiv \Delta J$ (isto é: $Q_{ij} = \partial J_i / \partial X_j$) e assumindo a seguinte aproximação:

$$J(X + \Delta X) = J(X) + \nabla J \cdot \Delta X + O(\Delta X^2) \approx J(X) + Q\Delta X$$

as equações normais podem ser escritas como:

$$Q^T J(X + \Delta X) \approx Q^T [J(X) + Q\Delta X] = 0,$$

ou seja:

$$(Q^T Q) \Delta X = -Q^T J.$$

A idéia do MLM é modificar a matriz do sistema $M \equiv Q^T Q$, da seguinte maneira:

$$M \rightarrow \hat{M}; \text{ onde: } \hat{M} = M_{ij} + \lambda M_{ij} \delta_{ij}. \quad (2.6.25)$$

Se λ é grande ($\lambda \rightarrow \infty$), os termos da diagonal irão dominar a expressão, que é justamente o método da máxima descida. Quando $\lambda \approx 0$, o método torna-se similar ao método de Newton. Assim, a abordagem do MLM atualiza o cálculo para os parâmetros desconhecidos através da fórmula:

$$X_{k+1} = X_k - [Q^T Q + \lambda_k I]^{-1} Q^T J. \quad (2.6.26)$$

onde λ_k decai quando a iteração avança: $\lim_{k \rightarrow \infty} \lambda_k \rightarrow 0$.

A rotina DBCLSF da biblioteca numérica IMSL [42] implementa uma versão do método de Levenberg-Marquardt.

2.6.3 Método do Gradiente Conjugado (MGC)

Há uma extensa literatura sobre o Método do Gradiente Conjugado, devido à grande pesquisa dedicada a esta técnica – é considerada um dos 10 algoritmos mais importantes da computação científica! O emprego de direções conjugadas para a busca da solução ótima é utilizado com o objetivo de acelerar a convergência do método da máxima descida. Diferentemente do método da máxima descida, que na versão clássica assume passos em direções de busca idênticas a direções de busca anteriores, o MGC assume outra estratégia, empregando direções ortogonais conjugadas. Todavia, assim como o método da máxima descida, o MGC também emprega apenas informação referente ao gradiente (jacobiano) da função objetivo (ao invés de informações da derivada segunda – a hessiana da função objetivo).

Aqui vamos dar somente a versão algorítmica do método. A dedução dos passos do MGC é bem descrita na literatura especializada [82, 85]. Outra referência a ser mencionada é a tese de Chiwiacowsky [31]: além da dedução do MGC, o método variacional é descrito com detalhes na sua aplicação a problemas inversos – o método variacional será tratado mais adiante neste texto. Um algoritmo para o MGC pode ser descrito como:

1. Escolher uma estimativa inicial para X_0 ;
2. Cálculo do modelo direto $A(X_k)$ e da função objetivo $J(X_k)$;
3. Calcular o gradiente da função objetivo: $\nabla J(X_k)$;
4. Cálculo do coeficiente conjugado (versão Fletcher-Reeves):

$$\gamma^k = \frac{\|\nabla J(X_k)\|_2^2}{\|\nabla J(X_{k-1})\|_2^2}; \quad (2.6.27)$$

5. Cálculo da direção de busca:

$$p^k = \begin{cases} -J(X_k) & \text{para } k = 0 \\ -J(X_k) + \gamma^k p^{k-1} & \text{para } k > 0 \end{cases}; \quad (2.6.28)$$

6. Realizar uma busca em linha na direção p^k para calcular:

$$\beta^k = \min_{\beta} \{J(X_k + \beta p^k)\}; \quad (2.6.29)$$

7. Atualize o vetor de parâmetros: $X_{k+1} = X_k + \beta^k p^k$;
8. Teste de convergência: finaliza ou retorna ao passo-2.

A expressão *gradiente conjugado* deriva do fato de que as direções consecutivas de busca são A -ortogonais ou A -conjugadas: $(p^{k-1})^T A p^k = 0$.

2.6.4 Recozimento Simulado (SA)

Esta técnica foi o primeiro método de otimização estocástica inspirado em emular um processo natural de otimização de funções, no caso, minimizar a energia interna de ligação entre átomos de uma rede cristalina. A técnica é muito conhecida pela expressão em inglês: *Simulated Annealing* (SA): é uma meta-heurística para o problema de otimização global. O nome da técnica foi cunhado a partir do processo em metalurgia de aquecimento e resfriamento controlado de um material para a obtenção de uma rede cristalina com o mínimo de defeitos – o recozimento (annealing). O aquecimento altera a energia interna de ligação entre os átomos da rede cristalina, permitindo que o sistema se desloque aleatoriamente através de estados de energia mais alta (novas configurações da rede cristalina); o resfriamento lento torna o sistema capaz de se acomodar numa configuração com energia interna mais baixa do que a energia inicial, ou seja, numa configuração mais estável (menor energia) e a rede cristalina será mais bem estruturada, ou ainda, com menos imperfeições.

Em analogia com o processo físico, no algoritmo SA a energia interna é identificada como a função objetivo do problema de otimização. Em cada passo a solução atual é atualizada, sendo substituída por uma outra solução grosseiramente aleatória, escolhida com uma probabilidade que depende dos valores da função objetivo entre a solução anterior e a solução atual e de um parâmetro global T (“temperatura”), que vai gradualmente reduzindo seu valor durante o processo iterativo. Quando o parâmetro T é grande, praticamente qualquer solução é permitida, mas, a medida que T se aproxima de zero, somente soluções que reduzam o valor da função objetivo são consideradas. O estágio do processo iterativo que permite soluções com valor da função objetivo acima do valor obtido na iteração anterior torna o método efetivo em evitar estacionar em mínimos locais. Outra vantagem do SA é que neste método não é necessário o cálculo do gradiente da função objetivo, nem a exigência de continuidade da função objetivo e de suas derivadas.

O método foi descrito por Kirkpatrick et al. [73] e é uma adaptação do algoritmo de Metropolis, um método de Monte Carlo empregado em sistemas termodinâmicos [89]. O algoritmo a seguir foi implementado por Press et al. [110]:

1. Definir a temperatura inicial e um conjunto inicial de M soluções: X_0^m ($m = 1, 2, \dots, M$);
2. Cálculo da função objetivo de cada solução candidata da iteração: $J(X_k^m)$;
3. Escolher novas soluções candidatas aleatoriamente e recalcular a função objetivo associado (se o valor da função objetivo coincidir em algumas das soluções geradas ou com soluções candidatas no passo da iteração anterior, escolhem-se outras);
4. Avaliação das soluções candidatas:
 - (i) Se $J(X_k^m) < J(X_{k-1}^m)$ a solução é aceita - caso contrário,
 - (ii) Se um valor aleatório *rand* satisfizer: $rand > e^{-\Delta E/T_m}$, onde $\Delta E \equiv |J(X_k^m) - J(X_{k-1}^m)|$, a solução é aceita;
5. Redução da temperatura T : $T_{k+1} = r * T_k$ (r é o fator de redução: $|r| < 1$);
6. Teste de convergência: finaliza ou retorna ao passo-1.

Há muitas variantes do algoritmo SA: *fast SA*, *very fast SA*, *simulated re-annealing* (ou SA com ciclos térmicos).

2.6.5 Algoritmo Genético Epidêmico (EGA)

Algoritmos genéticos são métodos de otimização baseados na teoria da evolução proposta por Charles Darwin e Alfred Russel Wallace. Darwin e Wallace coletaram evidências para sua teoria evolucionária a partir de viagens que ambos fizeram na América do Sul. Darwin foi o naturalista convidado durante a viagem do navio *Beagle*, em expedição para mapear a costa da América do Sul, e Wallace foi um naturalista que viveu na região Amazônica por 4 anos. Ambos conceberam o mecanismo da seleção natural após terem lido o livro de Thomas Malthus: *An Essay on the Principle of Population*, onde Malthus afirma que a população humana cresce mais rápido do que a produção de alimentos.

A teoria de Darwin-Wallace é a idéia principal da moderna teoria da evolução, mas a proposta original tinha problemas. Por exemplo, não estava muito claro com a especiação funcionava (mais ainda, como ocorriam explosões de bio-diversidade, como identificado no período Cambriano): também não estava claro como se davam as pequenas mudanças entre indivíduos de uma mesma espécie. Um nome central para a síntese da moderna teoria evolucionária é Ernst Mayr. Ele é identificado com o principal responsável pela grande síntese, que integra a teoria da hereditariedade de Mendel com a teoria da evolução de Darwin-Wallace (e o processo de seleção natural). A teoria da evolução ainda está em desenvolvimento. As propostas do equilíbrio pontuado e da endo-simbiose são dois exemplos de idéias novas no campo.

Na formulação dos algoritmos genéticos (GA: *Genetic Algorithm*) os procedimentos são concebidos a partir da versão moderna da teoria da evolução – a grande síntese [61]. No contexto da otimização, a função objetivo (chamada de *função de aptidão*) funciona como o ambiente (o mecanismo da pressão evolutiva). Seguindo a seleção natural, somente os indivíduos mais aptos na população serão selecionados para serem os pais dos indivíduos da nova geração. A nova geração será levemente distinta da geração anterior por uma variação aleatória em algum ou mais itens do genótipo.

Na versão clássica de GA, há 3 operadores fundamentais: (i) seleção, (ii) cruzamento, (iii) mutação. Na versão empregada neste minicurso, um novo operador genético, denominado *epidêmico*, é implementado [27, 88]. O operador epidêmico é ativado quando um número específico de gerações é atingido sem alteração do melhor indivíduo (mais apto) durante algumas iterações. Assim, a população é afetada por uma *praga* e somente os indivíduos com melhor aptidão (um percentual pré-definido de 2% a 10%) irão sobreviver. Os indivíduos extintos serão substituídos por novos indivíduos (como imigrantes recém chegados) gerados aleatoriamente. Assim, dois novos parâmetros são necessários para a versão do EGA (*Epidemic GA*): o número de gerações sem melhoria do melhor indivíduo para ativar o operador epidêmico e a quantidade de indivíduos que sobreviverão à praga.

Outro ponto importante na codificação do GA clássico é a descrição dos indivíduos (números ou funções) em codificação binária. Esta descrição permite uma similaridade aparentemente mais nítida com o mapa gênico de um indivíduo para as operações de cruzamento e de mutação. Assim, o problema era codificado para estrutura binária e o resultado final, decodificado para números reais decimais. Na verdade, esta codificação/decodificação binária não é necessária. Há muitas formas de se aplicarem os operadores genéticos com números reais. Aqui só será descrito o algoritmo do EGA usado nas aplicações espaciais descritas na próxima seção.

1. Geração aleatória da população de M indivíduos (possíveis soluções): X_0^m ($m = 1, 2, \dots, M$);
2. Cálculo da aptidão (função objetivo) de cada indivíduo: $J(x_k^m)$;
3. Ativa operador de seleção (estratégia torneio [27]);
4. Ativa o operador de cruzamento: $z_k = x_i^\mu y_i^\mu$ (x_i e y_i são 2 pais selecionados ($x_i, y_i \in X^m$) e z_i um indivíduo da nova geração $k + 1$, $\mu = 0,25$ é o parâmetro usado para o cruzamento: equivale igual peso para ambos os pais);
5. Ativa operador de mutação:

$$z'_i = \begin{cases} z_i + \Delta(k, l_{\text{sup}} - z_i) & \text{se: dígito aleatório binário é 0} \\ z_i - \Delta(k, z_i - l_{\text{inf}}) & \text{se: dígito aleatório binário é 1} \end{cases};$$

onde: $\Delta(k, y) = y [1 - \text{rand}^{(1-k/N_k)^b}]$ (rand =número aleatório entre $[0,1]$, k é o passo da iteração atual, N_k é o número máximo de iterações, $b = 5$);

6. Teste de convergência: finaliza ou avalia aplicação do operador epidêmico – se a epidemia não acontece, retorna ao passo-2.

A estratégia da epidemia tem uma similaridade com a estratégia dos ciclos térmicos (ou *re-annealing*) do SA. O EGA é muito similar a implementação do micro-GA [77]. Cuco et al. [38] discutem a similaridade entre as duas abordagens (EPA e μ -GA) e comparam os resultados na solução de um problema inverso de identificação de parâmetros (constante de dissociação e da concentração máxima adsorvida) em transferência de massa com adsorção de bio-moléculas em leito de resina. Este trabalho também faz uma avaliação do desempenho da solução inversa obtida com o EGA e GEO (*Generalized Extreme Optimization* [144]).

O método EPA foi aplicado ao problema de otimização empregando regularização de Tikhonov para o problema inverso da Seção 1.2 [17].

2.6.6 Otimização por Colônia de Formigas (ACO)

A otimização por colônia de formigas (ACO: *Ant Colony Optimization*) é um método que emprega uma meta-heurística baseada no comportamento coletivo de formigas escolhendo um caminho entre o ninho e uma fonte de alimento [41]. Nota-se que o trajeto estabelecido pela maioria das formigas é o menor caminho entre o ninho e o alimento. Assim, ao se emular artificialmente este comportamento das formigas, ter-se-á um método para otimização de funções.

Cada formiga marca seu caminho depositando uma quantidade de ferormônio em tal caminho, estabelecendo uma *comunicação química* precebida por outras formigas. A Figura 2.3 ilustra o comportamento mencionado, onde a maior parte das formigas (Figura 2.3c) *escolhe* o caminho mais curto. No método ACO, várias gerações são produzidas. Para cada geração, uma quantidade fixa de formigas (n_a) é avaliada. Cada formiga está associada a um caminho possível e este caminho representa uma solução candidata. Cada formiga é gerada a partir destes caminhos.

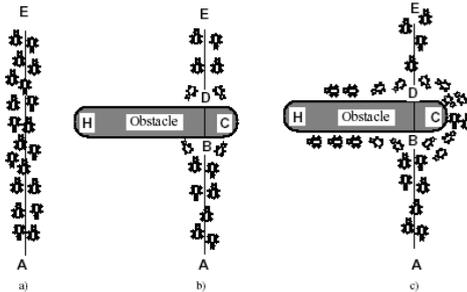


Figura 2.3: Formigas ultrapassando um obstáculo (adaptada de [41]).

Quando a estratégia ACO é empregada para identificar uma função contínua $f(x)$, inicialmente adota-se a seguinte partição: $f_r = f_{\min} + r * \Delta_f$ ($r = 0, 1, \dots, n_s$), onde $\Delta_f \equiv |f_{\max} - f_{\min}|/n_s$, $f_{\max/\min} = \max/\min |f(x)|$. A variável $x \in [a, b]$ é então discretizada: $x_s = a + s * \Delta_x$ ($s = 0, 1, \dots, n_p$), com $\Delta_x = |b - a|/n_p$. Os caminhos serão constituídos pela poligonal que une os pontos escolhidos aleatoriamente entre os nodos r da partição f_r no nível x_s (um grafo unindo pontos na matriz $n_s \times n_p$). Cada solução é então expressa por uma n_s -upla $[i, j]$.

Uma matriz de ferormônio \mathbf{T} é definida com dimensão $n_s \times n_p$. Para cada geração, todas as formigas (soluções candidatas) são avaliadas por meio da função objetivo e a melhor formiga é marca com ferormônio o seu *caminho* - na verdade o valor dos elementos correspondentes a esta n_s -upla na matriz T_{ij} é magnificado. Denotando por ρ a taxa de decaimento do ferormônio, a quantidade de ferormônio T_{ij} para a geração k é dada por

$$T_{ij}(k) = (1 - \rho) T_{ij}(k - 1), \quad \text{para: } k = 1, 2, \dots, m_{it} \quad (2.6.30)$$

onde m_{it} é o número máximo de iterações e a quantidade inicial de ferormônio em $k = 0$ é calculada por [10]

$$T_{ij}(0) \equiv T_0 = \frac{1}{n_s Q}. \quad (2.6.31)$$

A probabilidade de um dado caminho $[i, j]$ ser escolhido para cada geração de formigas é

dada por

$$P_{ij}(k) = \frac{[T_{ij}(k)]^\alpha [\eta_{ij}(k)]^\beta}{\sum_l \{ [T_{il}(k)]^\alpha [\eta_{il}(k)]^\beta \}} \quad (2.6.32)$$

onde $l \in [1, n_p]$ e η_{ij} é a visibilidade/custo para cada caminho (este conceito associa-se ao problema do caixeiro viajante). Os parâmetros α e β são escolhidos de forma a estabelecer um limiar entre feromônio e visibilidade. Nos problemas trabalhados usou-se $\alpha = 1$ e $\beta = 0$. Fixando-se um valor q_0 , é gerado um número aleatório $rand$ e comparado com q_0 . Se $rand > q_0$ então o caminho escolhido é o de maior probabilidade P_{ij} . Caso contrário, será escolhido o caminho (i, j^*) cuja a probabilidade cumulativa $\sum_{j=1}^{j^*} P_{ij} > r_0$.

Após este processamento, a formigas são avaliadas e no caminho de menor custo $(i, j)_{\min}$ é depositada uma quantidade adicional de feromônio:

$$T_{\{ij\}_{\min}}(k) = T_{\{ij\}_{\min}}(k) + T_0. \quad (2.6.33)$$

ACO com Regularização Intrínseca

Uma informação *a priori* assumida em problemas inversos é impor uma certa suavidade na solução. Deste modo, após se produzir uma geração de formigas, antes de se fazer uma avaliação destas soluções em relação a função objetivo, faz-se uma *pré-seleção*, onde somente os caminhos mais suaves são selecionados [112]. A suavidade θ pode ser quantificada por uma norma de Tikhonov-2:

$$\theta(f_r) = \left\| f_r^{(2)} \right\|_2^2. \quad (2.6.34)$$

Esta estratégia permite não só reduzir o esforço computacional, como também obter melhores soluções inversas. Outras penalidades podem ser incorporadas, como o critério da 2a. derivada [141].

2.7 Método Variacional

A técnica variacional foi primeiramente desenvolvida pelo pesquisador russo Oleg M. Alifanov, especialista de problemas inversos com aplicações em transferência de calor da área espacial. Basicamente o esquema consiste em associar o método do gradiente conjugado à equação adjunta (algumas vezes é chamado de *método da equação adjunta*). É devido ao uso do MGC que esta técnica de resolução de problemas inversos é apresentada após a seção de otimização.

O método variacional foi desenvolvido para se determinarem extremos de funcionais. Uma função é uma transformação (ou mapeamento) que relaciona número (e também vetores e/ou tensores) com números (ou vetores/tensores). Funcionais são mapeamentos que relacionam funções com números. No caso de funções, usamos propriedades das derivadas para o cálculo de extremos de funções. No caso de funcionais, usa-se a variação (ao invés da diferenciação) para se determinar extremos de funcionais (ou ainda, pontos estacionários do funcional). Numa descrição pouco técnica, se duas funções diferem uma da outra pela soma de outra função $\hat{u}(x) = u(x) + \epsilon \eta(x)$, a variação da função $u(x)$ é dada por:

$$\delta u(x) = \hat{u}(x) - u(x) = \epsilon \eta(x), \quad (2.7.35)$$

onde $\eta(x)$ é uma função arbitrária (mas que se anula na fronteira) e ϵ é um parâmetro pequeno. A diferença entre o operador variacional δ e o operador diferencial d é clara: δu e du são ambas mudanças diferenciais da função $u(x)$; porém: du se refere a uma mudança infinitesimal de $u(x)$ causada pela variação infinitesimal do argumento dx , isto é,

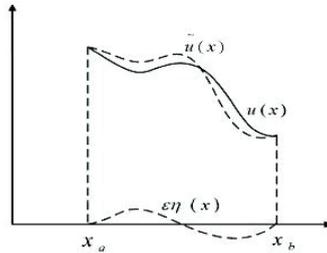


Figura 2.4: Funções $u(x)$, $\hat{u}(x)$ e a variação $\delta u = \epsilon\eta(x)$ – adaptado de [39].

$du = u(x+dx) - u(x)$; enquanto que δu é uma variação infinitesimal de $u(x)$ que produz uma nova função: $\delta u = \eta(x)$ (ver Fig. 2.4).

O cerne da teoria do método variacional é uma metodologia para se determinar o extremo de um funcional J tomando-se a primeira variação deste funcional e igualando-se a zero: $\delta J = 0$, similar ao estudo de funções do cálculo diferencial. Não será apresentada a teoria do cálculo variacional. Porém, somente para ilustrar o procedimento, vamos considerar a variação da integral I abaixo, onde as funções $u(x)$, sua derivada e a variável x estão relacionadas por uma função $F(\cdot)$:

$$\begin{aligned} \delta I &= \delta \int_a^b F[u(x), u'(x), x] dx = \int_a^b \delta F dx \\ &= \epsilon \int_a^b \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right) dx, \quad x \in (a, b); \end{aligned} \quad (2.7.36)$$

onde $u'(x)$ é a derivada da função $u(x)$. Integrando por partes a expressão acima:

$$\delta I = \epsilon \int_a^b \eta \left(\frac{\partial F}{\partial u} - \frac{d}{dx} \frac{\partial F}{\partial u'} \right) dx + \epsilon \eta(x) \frac{\partial F}{\partial u'} \Big|_a^b. \quad (2.7.37)$$

Para determinar o ponto estacionário do funcional $\delta I/\epsilon$, lembramos que a função $\eta(x)$ é arbitrária e anula-se nos extremos: $\eta(x) = 0$, para $x = a$ e $x = b$. Assim, a condição para tornar nula a expressão acima é:

$$\frac{\partial F}{\partial u} - \frac{d}{dx} \frac{\partial F}{\partial u'} = 0. \quad (2.7.38)$$

A equação (2.7.38) é a condição necessária e suficiente para anular δJ e é conhecida como a equação de Euler-Lagrange, em homenagem aos dois grandes matemáticos que trabalharam no desenvolvimento do cálculo variacional. Entretanto, diferentemente da técnica de regularização, é difícil gerar uma formulação variacional para qualquer problema. A dedução da equação de Euler-Lagrange é calculada caso a caso.

O caso de minimização de um funcional com certos vínculos pode ser holomônico (quando o vínculo é dado a partir de um outro funcional da relação entre as funções do funcional original a ser minimizado), ou não holomônicos (equações que relacionam funções do funcional a ser minimizado). Vínculos são equações adicionais que impõem certo comportamento ou propriedade da função que minimiza o funcional. Neste caso, o problema pode ser formulado pela minimização de um novo funcional sem restrições, onde as restrições são incorporadas ao novo funcional com a ajuda de multiplicadores de Lagrange,

que são funções desconhecidas. As equações que definem os multiplicadores de Lagrange são as equações adjuntas.

Para o caso de condições auxiliares (vínculos):

$$\begin{aligned} \text{restrição holonômica : } & \int_a^b g[u, u', x] dx = 0, \\ \text{restrição não holonômica : } & g[u, u', x] = 0. \end{aligned}$$

A equação de Euler-Lagrange para o primeiro caso é

$$\frac{\partial}{\partial u}(F + \lambda g) - \frac{d}{dx} \frac{\partial}{\partial u'}(F + \lambda g) = 0. \quad (2.7.39)$$

no segundo caso se a função g depender de $u'(x)$, o vínculo é uma equação diferencial e a Euler-Lagrange associada é dada por

$$\frac{\partial}{\partial u}[F + \lambda(x)g] - \frac{d}{dx} \frac{\partial}{\partial u'}[F + \lambda(x)g] = 0 \quad (2.7.40)$$

onde λ é um multiplicador de Lagrange [78]. Para o vínculo holonômico (2.7.39) o multiplicador de Lagrange é constante, no caso não holonômico λ depende da variável x . Veremos nas aplicações como determinar os multiplicadores de Lagrange.

O método iterativo de Alifanov[1] pode ser sumarizado como:

1. Escolher uma estimativa inicial X_0 ,
2. Resolver o problema direto $A(u) = f$, onde: $X = [u_1, \dots, u_N]^T$,
3. Resolver o problema adjunto: multiplicador de Lagrange (λ),
4. Conhecendo λ , calcular o gradiente: $\partial J/\partial X$,
5. Calcular o coeficiente conjugado (γ^k): Eq. (5.1.9),
6. Calcular a direção descendente p^k : Eq. (2.6.28),
7. Fixar: $\Delta X = p^k$ para o problema de sensibilidade,
8. Calcular o tamanho do passo (β^k): Eq. (2.6.29),
9. Atualizar a estimativa X_{k+1} ,
10. Verificar a convergência.

O método variacional de Alifanov, como descrito no algoritmo acima, foi aplicado ao problema inverso da Seção 1.2 [17].

Embora o método variacional de Alifanov seja uma técnica elegante e poderosa, quando aplicada em problemas inversos de vibração (vibração em placas são problemas hiperbólicos), o procedimento não apresentou o mesmo desempenho como em problemas inversos de condução do calor (aqui o problema direto é parabólico). Diferentemente do que afirma Huang [65, 66], a técnica falha quando são considerados muito graus de liberdade [28], ou no caso de danos severos, ou ainda porque a rigidez na estrutura não tem um alto valor [32]. Basicamente a convergência não ocorre se a estimativa inicial não estiver próxima a bacia de convergência do MGC (Seção 2.6.3).

Deste modo, para o método variacional ser aplicado em problemas inversos de vibração com vários graus de liberdade, sob danos severos ou uma estrutura de baixa rigidez, é necessário uma técnica híbrida [28]. Em nossas aplicações, combinamos o método EGA com o método variacional. O EGA será executado para suprir a estimativa inicial para o passo-1 do método variacional [32]. Mais detalhes serão apresentados no Capítulo 4.

2.8 Redes Neurais Artificiais (RNA)

O uso de redes neurais para solução de problemas inversos é muito recente. No Brasil, o grupo do prof. João P. Braga do Instituto de Química da UFMG foi também um dos pioneiros na aplicação de redes neurais para problemas inversos, empregando, por exemplo, a rede de Hopfield [101]. O prof. Braga também desenvolveu uma formulação não linear da rede de Hopfield para ser aplicada em problemas inversos. As redes neurais examinadas nesta Seção, são aquelas usadas de maneira mais sistemática no grupo do LABAC-INPE.

2.8.1 Perceptron de Múltiplas Camadas (PMC)

Redes neurais artificiais (RNAs) são arranjos de elementos de processamento chamados neurônios (Figura 2.5b). O modelo de neurônio artificial consiste de uma combinação linear seguida de uma função de ativação (ver Figura 2.5a):

$$y_m = \varphi(\nu_m) = \varphi \left(\sum_{n=1}^{N_r} w_{mn} x_n + b_m \right) \quad (2.8.41)$$

onde w_{mn} são os pesos das conexões e b_m é o viés [59]. A função de ativação pode assumir várias formas funcionais. As mais usadas são as funções sigmóide e tangente hiperbólica (respectivamente):

$$\varphi(x) = (1 + e^{-ax})^{-1} \quad (2.8.42)$$

$$\varphi(x) = \tanh(x) = (1 - e^{-ax}) (1 + e^{-ax})^{-1} \quad (2.8.43)$$

As redes podem ser supervisionadas ou não supervisionadas. Nas redes supervisionadas, os pesos das conexões são determinados de tal maneira que a saída da rede neural produz um resultado similar ao do objetivo alvo. Assim, para as redes supervisionadas, tem-se 2 conjuntos y^{alvo} e $y^{\text{rede}}(\mathbf{W})$, onde \mathbf{W} é a matriz dos pesos das conexões. Para determinar esta matriz, o objetivo é minimizar a diferença quadrática:

$$J(\mathbf{W}) = \|y^{\text{alvo}} - y^{\text{rede}}(\mathbf{W})\|_2^2 \quad (2.8.44)$$

Um método não muito eficiente, mas muito usado, é o algoritmo de retro-alimentação (*back-propagation*), onde a atualização da matriz \mathbf{W} é dada pela regra δ :

$$W_{mn}^k = W_{mn}^{k-1} + \Delta W_{mn} \quad (2.8.45)$$

onde o termo ΔW_{mn} é calculado como:

$$\Delta W_{mn} = \eta \frac{\partial E}{\partial w_{mn}} = \eta \delta_n x_{nm}, \quad (2.8.46)$$

$$E = \sum_{n_e \text{ saída RN}} [y^{\text{alvo}} - y^{\text{rede}}(\mathbf{W})]^2, \quad (2.8.47)$$

$$\frac{\partial E}{\partial w_{mn}} = \frac{\partial E}{\partial y_n^{RN}} \frac{\partial y_n^{RN}}{\partial w_{mn}} = \frac{\partial E}{\partial y_n^{RN}} x_n \Rightarrow \delta_n \equiv \frac{\partial E}{\partial y_n^{RN}}. \quad (2.8.48)$$

entre o valor de saída da rede (y^{RN}) e o valor de referência (y^{alvo}). Qualquer técnica de otimização estudada até aqui pode ser usada para encontrar a matriz \mathbf{W} , de pesos das conexões entre os neurônios artificiais.

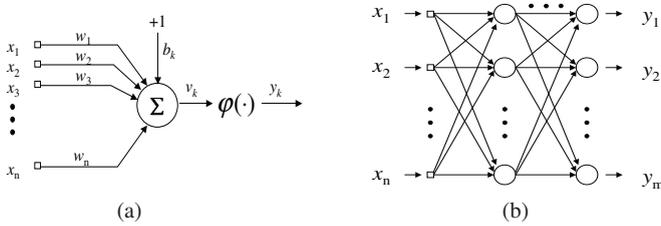


Figura 2.5: (a) Neurônio artificial simples, (b) rede neural de multi-camada.

Quando se considera além das entradas x_n , um limiar b_m , a regra δ de atualização dos pesos sinápticos na camada- l torna-se

$$W_{mn}^k = W_{mn}^{k-1} + \alpha [W_{mn}^k - W_{mn}^{k-1}] + \delta_m^{(l)}(k) y_n^{l-1}(k). \quad (2.8.49)$$

Porque a regra δ funciona? A resposta é simples. Se a direção de busca (ΔW) na equação (2.8.45) fosse dada em termos do método de Newton, na expressão ter-se-ia o produto da inversa da matriz hessiana \mathbf{H} pelo o gradiente do erro. Como vimos no método quase-Newton, a matriz hessiana não é calculada explicitamente: a inversa da matriz hessiana é aproximada por outra expressão. Assim, num extremo de simplificação, a inversa da matriz hessiana poderia ser dada pela matriz diagonal homogênea: $\mathbf{H}^{-1} \approx \eta \mathbf{I}$, sendo \mathbf{I} a matriz identidade³.

Aqui foi feita uma descrição muito breve da rede perceptron de múltiplas camadas. Todavia, redes neurais é um campo vasto e fascinante e há uma verdadeira explosão de aplicações e de novas redes neurais sendo desenvolvidas.

As redes neurais foram efetivas para resolver o problema inverso da Seção 1.2 [129].

2.8.2 Função de Base Radial (FBR)

Este tipo de rede tem somente uma única camada escondida [59]. Vários tipos de funções de base radial podem ser usadas, mas geralmente é empregada a função gaussiana – aqui chamamos de FBR *padrão*. Os pesos da camada de entrada para a camada escondida representam *centros de classes* e podem ser determinados como nos métodos de agrupamento. Os pesos entre a camada escondida e a camada de saída podem ser determinada por qualquer regra de treinamento (otimização), como a regra delta, revisada na Seção anterior.

O processo de ativação da rede é dado pela equação:

$$y(x) = \sum_{i=1}^N w_i \phi(\|x - t_i\|) \quad (2.8.50)$$

onde x é o vetor de entrada, w_i são pesos da conexão entre os neurônios, ϕ é a função de base radial e t_i são os centros da i -ésima FBR.

A estatística não extensiva de Tsallis pode ser usada para desenvolver uma nova rede neural de base radial não extensiva. Como apresentado por Tsallis [151], as distribuições $p_q(r)$ em (2.4.8) assumem expressões distintas para diferentes valores de q (parâmetro de não extensividade).

Formas funcionais para a distribuição de Tsallis de acordo com o parâmetro q :

³Uma matriz C é uma inversa aproximada [36] da matriz A se: $\|I - CA\| < 1$.

Se $q > 1$:

$$p_q(r) = \phi(r) = \frac{1}{\sigma} \left[\frac{q-1}{\pi(3-q)} \right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{1}{q-1}\right)}{\Gamma\left(\frac{3-q}{2(q-1)}\right)} \frac{1}{\left(1 + \frac{q-1}{3-q} \frac{r^2}{\sigma^2}\right)^{\frac{1}{q-1}}}, \quad (2.8.51)$$

Se $q = 1$ (FBR padrão):

$$p_q(r) = \phi(r) = \frac{1}{\sigma\sqrt{2\pi}} e^{-r^2/2\sigma^2}; \quad (2.8.52)$$

Se $q < 1$:

$$p_q(r) = \phi(r) = \frac{1}{\sigma} \left[\frac{1-q}{\pi(3-q)} \right]^{\frac{1}{2}} \frac{\Gamma\left(\frac{5-3q}{2(1-q)}\right)}{\Gamma\left(\frac{2-q}{1-q}\right)} \left[1 - \frac{(1-q)}{(3-q)} \frac{r^2}{\sigma^2} \right]^{\frac{1}{(1-q)}}. \quad (2.8.53)$$

Capítulo 3

Problemas Inversos em Ciência Espacial

3.1 Inversão Magneto-telúrica

O objetivo da prospecção geofísica é inferir o tipo de estrutura geológica sob a superfície da Terra a partir de medidas de alguma propriedade de interesse.

Na inversão magneto-telúrica, o campo geo-eleto-magnético é medido por sensores na superfície. A crosta terrestre é compreendida como um condutor estratificado que, com a chegada de uma onda eletro-magnética proveniente de fontes naturais (raios cósmicos, por exemplo), vai se propagar para o interior da Terra e induzir uma corrente neste condutor. A anomalia no campo geomagnético é então registrada. Porém, este registro depende das propriedades do material. Aqui vamos apresentar uma técnica de inversão, chamada de método da mínima entropia de primeira ordem (MINENT-1), para reconstrução de distribuições bidimensionais de condutividade geolétrica, a partir de dados magnetotelúricos (MT). O método combina uma busca iterativa com uma técnica de regularização baseada na minimização da medida de entropia do vetor de diferenças primeiras das condutividades a serem estimadas. Simulações numéricas, com a utilização de dados sintéticos contaminados com ruído gaussiano, mostram que o algoritmo MINENT-1 produz excelentes reconstruções de condutividade, com resultados melhores que os obtidos pelo método da máxima entropia.

Diferentemente dos outros esquemas clássicos de regularização, que maximizam suavidade para um dado conjunto de dados, o método proposto limita a classe de possíveis soluções a um conjunto restrito de modelos de baixa entropia, constituído por regiões localmente lisas separadas por descontinuidades abruptas. Esta abordagem pode ser bastante eficaz para incorporação de informação *a priori* sobre a natureza da suavidade local do modelo físico real.

A descrição aqui segue o artigo publicado na revista Brasileira de Geofísica. [12] (ver também [114]).

3.1.1 Formulação do Modelo Direto

Várias abordagens são possíveis para resolver o problema direto, podendo ser técnicas analíticas ou numéricas (diferenças finitas, elementos de contorno ou elementos finitos, volumes finitos, métodos espectrais e técnicas híbridas). A escolha do método esta ligada a muitos fatores, deste considerações como velocidade de processamento, precisão e simplicidade. Nesta aplicação, os cálculos para solução do problema direto, requeridos pelo

esquema de inversão, são realizados por um código de diferenças finitas, baseado num modelo de condutividade não homogênea bi-dimensional proposto por Jones e Price [70].

A formulação matemática do problema é dada pelas equações de Maxwell numa região bi-dimensional com condições de contorno adequadas. Esquematicamente, o domínio do problema é esboçado na Figura 3.1.1, onde Ω^+ e Ω^- correspondem à zona condutiva ($z < 0$) e a zona de espaço livre ($z > 0$), respectivamente. O campo oscilante tem período $2\pi/\omega$, suficientemente longo para permitir que correntes de deslocamento possam ser ignoradas. A permeabilidade magnética é tomada como unitária. As equações são consequentemente

$$\vec{\nabla} \times \vec{H} = 4\pi\sigma\vec{E} \quad (3.1.1)$$

$$\vec{\nabla} \times \vec{E} = -i\omega\vec{H} \quad , \quad (3.1.2)$$

onde o fator temporal $\exp(i\omega t)$ é assumido em todas as quantidades de campos e $\sigma = \sigma(y, z)$ é a condutividade elétrica.

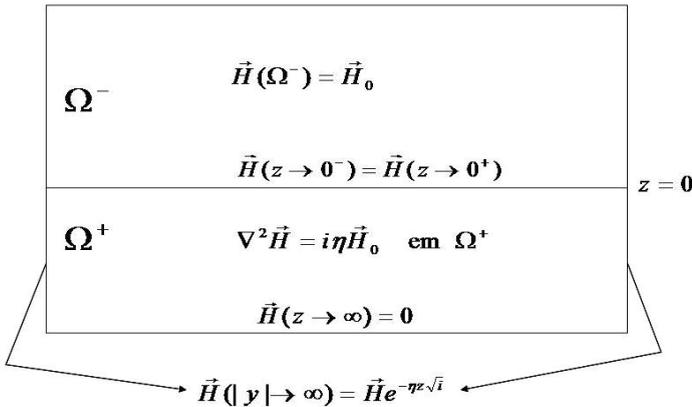


Figura 3.1: Geometria e condições de contorno do problema direto.

Visto que \vec{H} e \vec{E} são independentes da direção- x , Eqs. (3.1.1) e (3.1.2) tomam a forma de dois conjuntos de equações

$$\begin{cases} \partial H_z / \partial y - \partial H_y / \partial z & = 4\pi\sigma E_x \\ \partial H_x / \partial z & = 4\pi\sigma E_y \\ -\partial H_x / \partial y & = 4\pi\sigma E_z \end{cases} \quad (3.1.3)$$

e também,

$$\begin{cases} \partial E_z / \partial y - \partial E_y / \partial z & = -i\omega H_x \\ \partial E_x / \partial z & = -i\omega H_y \\ -\partial E_x / \partial y & = -i\omega H_z \end{cases} \quad (3.1.4)$$

que podem ser resolvidas separadamente.

Eliminando H_y e H_z das Eq. (3.1.3), as equações do modo elétrico transversal (TE) (problema de E -polarização) são obtidas:

$$\vec{\nabla}^2 E_x = \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} = i\eta^2 E_x \quad , \quad (3.1.5)$$

$$\partial E_x / \partial z = -i \omega H_y \quad (3.1.6)$$

$$- \partial E_x / \partial y = -i \omega H_z \quad , \quad (3.1.7)$$

onde

$$\eta^2 = 4 \pi \sigma \omega \quad . \quad (3.1.8)$$

Analogamente, eliminando E_y e E_z da Eq. (3.1.4), as equações do modo magnético (TM) (problema de *H-polarização*) são dadas por

$$\vec{\nabla}^2 H_x = \frac{\partial^2 H_x}{\partial y^2} + \frac{\partial^2 H_x}{\partial z^2} = i \eta^2 H_x \quad , \quad (3.1.9)$$

$$\partial H_x / \partial z = 4 \pi \sigma E_y \quad (3.1.10)$$

$$- \partial H_x / \partial y = 4 \pi \sigma E_z \quad . \quad (3.1.11)$$

Dados magneto-telúricos (MT) podem ser caracterizados pelas impedâncias dos modos TE e TM, dados respectivamente por

$$Z_{xy} = \frac{E_x}{H_y} \quad (3.1.12)$$

$$Z_{yx} = \frac{E_y}{H_x} \quad , \quad (3.1.13)$$

ou pelas resistividades e fase aparentes, da forma:

$$\rho_a = \frac{1}{\mu_0 \omega} |Z|^2 \quad (3.1.14)$$

$$\phi = \arg(Z) \quad , \quad (3.1.15)$$

onde Z refere-se a Z_{xy} e Z_{yx} para os modos TE e TM, respectivamente.

Embora a inversão 2D agrupe ambas as polarizações [8], por simplicidade somente o problema da *H-polarização* será considerada na análise.

Condições de Contorno

É assumido que as não homogeneidades condutivas inseridas em Ω^+ são suficientemente pequenas para permitir que o meio comporte-se como um condutor uniforme, para grandes distâncias das descontinuidades em σ . Consequentemente, quando $y \rightarrow \pm\infty$, a Eq. (3.1.9) torna-se

$$\frac{\partial^2 H_x}{\partial z^2} = i \eta^2 H_x \quad (3.1.16)$$

e o campo tende a zero para grandes valores positivos de z . Desta forma, a solução apropriada da Eq. (3.1.16) é

$$H_x = H_0 e^{-\eta z \sqrt{i}} \quad . \quad (3.1.17)$$

Nas interfaces entre Ω^+ e Ω^- ($z = 0$) e também entre medios diferentes dentro da região condutiva, H_x é contínua. Fora do condutor, onde $\sigma = 0$, H_x é independente de y e z .

Formulação Numérica

Separando o campos magnético em suas componentes real e imaginária, $\vec{H} = \vec{F} + i\vec{G}$, Eq. (3.1.9) pode ser escrita como

$$\nabla^2 H = \nabla^2 F + i \nabla^2 G = (-\eta^2 G) + i (\eta^2 F) \quad (3.1.18)$$

ou, escrita na forma matricial

$$\nabla^2 \Phi = i\eta^2 \mathbf{M} \Phi \quad (3.1.19)$$

com

$$\Phi = \begin{bmatrix} F \\ G \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Discretizando a Eq. (3.1.19) por diferenças finitas sobre uma malha não uniforme de primas retangulares, cada um tendo uma condutividade elétrica uniforme $\sigma_{j,k}$, o sistema resultante de equações algébricas pode ser iterativamente resolvido pelo método de Gauss-Seidel [60].

Um código de computador foi escrito baseado no procedimento de cálculo acima e validado contra resultados apresentado por [70].

3.1.2 Formulação do Problema Inverso

O vetor de condutividades a ser determinado pela análise inversa é denotado por

$$\mathbf{p} = \{p_1, p_2, \dots, p_q, \dots, p_Q\} = \{\sigma_{11}, \sigma_{21}, \dots, \sigma_{jk}, \dots, \sigma_{JK}\},$$

onde $q = J(k-1) + j$, com $j = 1, \dots, J$ and $k = 1, \dots, K$. A inversão MT pode ser formulada como um problema de minimização com restrição,

$$\min J(\mathbf{p}), \quad l_q \leq p_q \leq u_q, \quad q = 1, \dots, Q, \quad (3.1.20)$$

onde

$$J(\mathbf{p}) = R(\mathbf{p}) - \gamma_0 S_0(\mathbf{p})/S_{max} + \gamma_1 S_1(\mathbf{p})/S_{max}, \quad (3.1.21)$$

S_0 e S_1 são operadores de regularização, γ_0 e γ_1 são parâmetros positivos de regularização e S_{max} é uma constante de normalização. Os limites l_q e u_q são escolhidos para permitir que a inversão vá se restringir dentro de limites físicos conhecidos *a priori*.

A diferença quadrática entre dados do modelo e de observação é dada por

$$R(\mathbf{p}) = \sum_{j=1}^{N_y} \sum_{m=1}^M [\Phi_{j,m}^E - \Phi_{j,m}^C(\mathbf{p})]^2, \quad (3.1.22)$$

onde os superescritos E e C denotam os dados experimentais e dados calculados, respectivamente. É assumido que as medidas $\Phi_{j,m}^E$ estão disponíveis para $j = 1, 2, \dots, N_y$ posições horizontais e para ω_m , $m = 1, 2, \dots, M$ frequências diferentes. Considerando que o campo magnético para o modo TM é constante em $z = 0$, o algoritmo de inversão é suprido com valores H_x preditos pelo modelo direto sobre a grade horizontal *abaixo* da superfície da Terra. Isto é equivalente a uma aproximação de diferença finita de primeira ordem da Eq.(3.1.10) para o cálculo de Z_{yx} em $z = 0$.

A escolha do operador de regularização e dos parâmetros de regularização serão discutidos na próxima Seção.

Regularização de Mínima Entropia de Primeira Ordem

É bem conhecido que dados observacionais são geralmente insuficientes para prover uma solução única e estável para lidar com um problema inverso. Uma abordagem recomendada neste caso é o uso de alguma técnica de regularização, para assegurar que a variação dos parâmetros estejam limitadas de tal modo que a solução final seja fisicamente razoável. Geralmente, esta noção de vaga significa de fato uma condição de *suavidade* imposta a solução inversa. Em outras palavras, técnicas clássicas de regularização, tais como a regularização de Tikhonov e o formalismo da máxima entropia, buscam por regularidade *global* e permitem as reconstruções mais suaves que são consistentes com os dados disponíveis.

O princípio da máxima entropia foi primeiramente proposto por Jaynes como um procedimento geral de inferência [68], com base na caracterização axiomática de Shannon da quantidade de informação [127]. O princípio da máxima entropia tem sido sucessivamente usado a uma variedade de aplicações, incluindo radioastronomia [54], tomografia [137], teste não destrutivo [113], reconhecimento de padrões [47] e cristalografia [9].

Este novo operador de regularização foi introduzido baseado na minimização da medida de entropia S_1 do vetor das *diferenças-primeiras* de \mathbf{p} . Adotando a terminologia padrão [150], esta técnica de regularização é chamada de *mínima entropia de primeira-ordem* (MINENT-1). Analogamente, o método da máxima entropia, que usa a medida de entropia de ordem zero $S_0(\mathbf{p})$ como operador de regularização, é denotada por MAXENT-0. Deste modo, os operadores de regularização na Eq. (3.1.21) são dados por

$$S_\alpha(\mathbf{p}) = - \sum_{q=1}^Q s_q \log(s_q) \quad , \quad \alpha = 0, 1 \quad , \quad (3.1.23)$$

onde

$$s_q = r_q / \sum_{q=1}^Q r_q \quad (3.1.24)$$

e r_q é dado de acordo com o método de regularização

$$r_q = \begin{cases} p_q & \text{se } \alpha = 0 \\ |p_q - p_{q-1}| + \varsigma & \text{se } \alpha = 1 \quad , \end{cases} \quad (3.1.25)$$

ς é uma pequena constante positiva ($\varsigma = 10^{-15}$) que assegura a entropia de primeira-ordem ter sempre um valor definido. A função S_α tem seu máximo global quando todos os valores r_q são os mesmos, que corresponde a uma distribuição uniforme com valor de $S_{max} = \log Q$. Por outro lado, o menor valor de entropia, $S_{min} = 0$, é atingido quando todos os elementos, exceto um, são nulos.

Aspectos essenciais do operador de regularização MINENT-1 são discutidos por Campos Velho e Ramos [12]. Naquele trabalho, três modelos bem distintos: distribuição uniforme (modelo-a), distribuição gaussiana (modelo-b) e uma função de onda quadrada (modelo-c), são avaliados. Estes 3 modelos apresentam entropia de ordem-zero muito similares: $S_0^a = 1,0000$, $S_0^b = 0,9969$, $S_0^c = 0,9955$. Entretanto, a medida de entropia de primeira-ordem para estes modelos é: $S_1^a = 1,0000$, $S_1^b = 0,5958$, $S_1^c = 0,1502$. Assim, a descontinuidade no modelo acarreta um valor mais baixo para sua entropia de primeira-ordem, permitindo quantificar as diferenças entre os 3 modelos de maneira muito mais nítida.

Desta forma, enquanto os esquemas de regularização clássicos, tais como máxima entropia e inversão de Occam, buscam pelo *modelo mais suave que concorda com os dados dentro de uma tolerância esperada* [35], o método MINENT-1 busca por regiões localmente suaves separadas por descontinuidades de forma.

O teorema da concentração de entropia [69] é uma justificativa quantitativa para o método MINENT-1. De acordo com este teorema, a grande maioria dos resultados possíveis num experimento aleatório tem distribuição próximas a uniforme. Em outras palavras, distribuições com baixa entropia são completamente atípicas. Consequentemente, se há uma evidência *a priori* de um conteúdo de baixa entropia de primeira-ordem de uma estrutura geológica, o método MINENT-1 leva a uma redução drástica no número de soluções candidatas (isto é, aquelas que são consistentes com os dados disponíveis) para serem iterativamente avaliadas pelo algoritmo de inversão.

O(s) parâmetro(s) de regularização pode(m) ser determinado(s) por algum método mencionado no Capítulo 2. No exemplo descrito neste Capítulo, os valores destes parâmetros foram determinados por experimentação numérica.

3.1.3 Resultados Numéricos

O método descrito nas seções anteriores deste Capítulo foi testado para diferentes modelos de crosta terrestre, usando dados sintéticos. Em todas as simulações, o semi-espaco condutivo (Ω^+ , ver Fig. 3.1.1) foi dividido em células computacionais de 8×11 blocos, com $\Delta y = 10 \text{ km}$ e Δz variando de 1 a 10 km. Dados MT (parte real e imaginária de H_x) foram geradas pelo modelo direto, usando a mesma malha do esquema de inversão, para 11 estações em $z = 0$, e para 20 frequências logaritmicamente espaçadas variando de 0,0001 a 0,01 Hz. Para simular os erros experimentais, um ruído gaussiano foi adicionado aos dados exatos. As computações foram realizadas até atingir convergência, usando uma conditividade uniforme para o semi-espaco σ^+ como modelo de partida.

Os resultados são apresentados na forma de mapas de condutividade 2D em escala logarítmica. Os valores desconhecidos de condutividade foram agrupados no vetor \mathbf{p} , partindo do topo do canto esquerdo. A coluna mais a esquerda (condição de contorno) e a linha mais acima (superfície da Terra) são assumidas conhecidas e, portanto, estão fora do processo de inversão.

O método de inversão MINENT-1 foi aplicado a uma estrutura consistindo de um prisma condutivo Ω^c e um prima resistivo Ω^r , ambos inseridos no semi-espaco Ω^+ , com a razão de condutividade $\sigma^c/\sigma^+ = 10$ e $\sigma^r/\sigma^+ = 0.1$. Resultados numéricos foram calculados considerando os seguintes casos testes: (1) sem regularização; (2) MINENT-1 regularização; (3) e (5) MAXENT-0 regularização, e (4) MAXENT-0 e MINENT-1 regularização híbrida.

Mapas de condutividade, em escala logarítmica, são mostrados na Fig. 3.1.3a até 3.1.3f, mostrando o modelo verdadeiro (referência) usado para gerar os dados sintéticos dos casos analisados. Para cada caso teste, a Tabela 3.1.3 apresenta o número de iterações até a convergência final I_f , o resíduo normalizado rms , $\rho = R(\mathbf{p})/R(\mathbf{p}^0)$, as entropias normalizadas \mathcal{S}_0 e \mathcal{S}_1 , bem como o erro da raiz média quadrática rmq definido por

$$\epsilon = \left[\sum_{q=1}^Q (p_q - p_q^{exact})^2 / (p_q^0 - p_q^{exact})^2 \right]^{1/2}. \quad (3.1.26)$$

A figura 3.1.3c mostra que o algoritmo de inversão MINENT-1 recupera apropriadamente a distribuição condutiva. A combinação de de ambas as técnicas entrópicas num esquema híbrido, mostrado na Figura 3.1.3e, ocorre uma pequena degradação, se comparado com a Fig. 3.1.3c. Por outro lado, as reconstruções 3.1.3b e 3.1.3d foram contaminadas com valores incorretos. Em particular, Fig. 3.1.3d indicaria que γ_0 deveria aumentar para melhorar a regularização MAXENT-0. Entretanto, uma valor mais alto para o parâmetro de regularização neste caso somente aumenta a filtragem e estruturas com baixa entropia

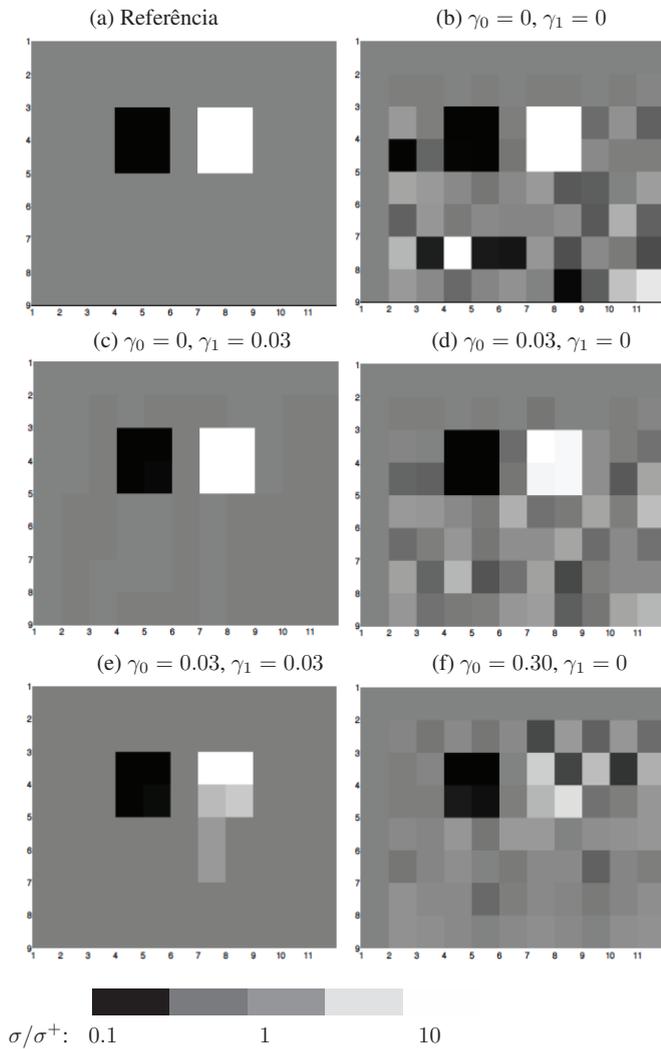


Figura 3.2: Mapas de condutividade em escala logarítmica.

na solução inversa, removendo artefatos e também informação útil, como mostrado na Fig. 3.1.3f.

Tabela 3.1: Resultados numéricos para os casos teste da Figura 3.1.3.

Test case	Figure	γ_0	γ_1	I_f	ϵ	ρ	S_0	S_1
1	4b	0.00	0.00	88	0.6355	0.0210	0.8536	0.7953
2	4c	0.00	0.03	210	0.0078	0.0233	0.8791	0.4054
3	4d	0.03	0.00	70	0.3134	0.0222	0.9003	0.8123
4	4e	0.03	0.03	176	0.5095	0.0272	0.9160	0.4726
5	4f	0.30	0.00	138	0.8000	0.0280	0.9575	0.8266

Uma comparação dos resultados de entropia na Tabela 3.1.3 indica que, enquanto os valores de entropia de ordem-zero se situam num intervalo estreito ($0,85 < S_0 < 0,96$), a entropia de primeira-ordem tem uma variabilidade muito maior ($0,41 < S_1 < 0,83$), os níveis mais baixos estão associados as melhores inversões. Este resultado sugere que a técnica de regularização MINENT-1 confina a classe de soluções possíveis a um conjunto restrito de modelos de baixa entropia, constituído por regiões suaves separadas por descontinuidades de forma.

3.2 Outras Aplicações de Entropia de Alta Ordem

Nesta Seção não serão descritas nenhuma aplicação adicional, mas serão mencionadas outros resultados de inversões em usando entropia de alta ordem.

Na linha de inversão geofísica, outra abordagem não sísmica é a inversão gravimétrica, onde a variação de densidade do(s) corpo(s) incrustado(s), o que acarreta numa anomalia da gravidade observada. Resultados com a técnica MINENT-1 obtiveram bons resultados em inversão gravimétrica [16, 104, 136].

Inversão com máxima entropia de segunda-ordem:

- Em meteorologia de camada limite: determinação de propriedades turbulentas: coeficiente de difusão turbulenta [13], termo de contra-gradiente [118].
- Identificação de fontes de poluição atmosférica (esta metodologia é também apropriada para identificar o ciclo bio-geo-químico de gases do efeito estufa). Primeiros resultados obtidos com otimização determinística quase-Newton [118, 119, 120, 121]. O problema de estimação de fonte de área [119, 120] obteve melhor estimativa quando um método estocástico de busca foi utilizado (PSO: *Particle Swarm Optimization*) [84].
- Identificação de perfis de temperatura atmosférica a partir de dados de satélites [115].

Capítulo 4

Problemas Inversos em Engenharia Espacial

O setor espacial enseja muitos desafios tecnológicos: tele-comunicação, análise térmica e de radiação e análise estrutural, de tal forma a garantir um bom funcionamento de todos os componentes e do sistema como um todo em um ambiente que se tinha poucas informações (na verdade, o espaço é um lugar inóspito: é um vácuo com temperaturas muitas baixas, de intenso fluxo de radiação e de partículas cósmicas), além do controle e atitude de naves espaciais.

Embora a área espacial tenha 4 grandes pilares, é a engenharia espacial que dá nome a área: a *era espacial* só tem início quando a humanidade desenvolveu tecnologia para levar um artefato para longe da superfície da Terra, com capacidade de controle e mesmo deixar sensores orbitando em torno do planeta. O estudo da resposta de estruturas a vibrações provocadas pela ascensão e deslocamento de um foguete, devido a propulsão e até mesmo o acoplamento acústico, foram alguns dos desafios dos pioneiros da tecnologia espacial.

A teoria matemática das vibrações constitui-se num tema permanente de investigação. Um área particularmente importante é o monitoramento da saúde estrutural, ou identificação de danos estruturais. Este é um problema de inverso de vibração. Pode-se configurar o problema de identificação de danos em vários níveis:

1. Existência do dano;
2. Localização do dano;
3. Avaliação da severidade do dano;
4. Previsão da vida útil da estrutura.

Aqui vamos descrever uma técnica desenvolvida recentemente para detecção de danos estruturais, que está baseada no método variacional [1], porém a estimativa inicial para o esquema do gradiente conjugado é obtida por um algoritmo genético epidêmico: desenvolvido na tese de mestrado de Medeiros [88] e aplicada à problemas inversos (estimação de condição inicial, ver Seção 1.2) por Chiwiacowsky e Campos Velho [27]. A técnica foi desenvolvida para o domínio do tempo [28, 29, 49] e para o domínio da frequência [30, 32, 49].

É pertinente destacar que, embora a abordagem variacional seja mencionada na literatura de problemas inversos em engenharia mecânica (por exemplo: [11]), só muito recentemente esta técnica foi proposta para problemas de identificação de danos [28, 65, 66].

Todavia, a elegante técnica variacional, diferentemente do que afirma Huang em seus artigos [65, 66], **não** funciona em identificação de danos para sistemas com altos graus de liberdade (maiores do que 5) [28] e/ou quando o sistema não é tão rígido (por exemplo, estruturas de alumínio [31]). Foi prevendo uma falha na convergência do método variacional *a per se*, que a técnica híbrida foi proposta, combinando o Algoritmo Genético Epidêmico (AGE) – ou seu acrônimo em inglês EPA: *Epidemic Genetic Algorithm* – com o método do gradiente conjugado com a equação adjunta, ou seja, a técnica híbrida: EGA + abordagem variacional.

4.1 Detecção de Danos na Estação Espacial Internacional (EEI)

Futuras missões espaciais envolverão o uso de estruturas grandes e leves. Porém, estruturas espaciais grandes e flexíveis apresentam uma grande incerteza no comportamento do seu comportamento dinâmico. Em geral, o estudo das características físicas de componentes de grandes estruturas espaciais é uma aproximação ao sistema real e somente pode ser verificado após a estrutura estar em órbita, para então se avaliar a resposta da estrutura aos distúrbios que estará submetida. De fato, muitos parâmetros estruturais, tais como rigidez de alguns dos seus membros ou coeficiente de amortecimento, entre outros, podem variar durante a missão, modificando a dinâmica e/ou comportamento estático da própria estrutura. Além disso, estruturas complexas tem um número tão grande de graus de liberdade, que para objetivos práticos é necessária realizar uma redução no modelo. Esta redução introduz um alto grau de incerteza no modelo matemático que descreve a dinâmica do sistema a ser usado e no projeto do sistema de controle para vibrações estrutural, bem como controle de atitude de toda a estação ou aeronave espacial. Desta forma, problema de identificação de parâmetros estruturais e de detecção de falhas (ou danos) estruturais são consideradas técnicas adequadas para obter um modelo de estrutura com um alto grau de fidelidade sobre o tempo de vida orbital de uma estação ou nave espacial.

Dano é definido como uma alteração introduzida num sistema, que intencional ou não, que afeta o desempenho atual ou futuro daquele sistema e métodos de identificação de danos são aqueles capazes de identificar e localizar um dano numa estrutura. Técnicas de detecção de danos são: inspeção visual ou técnicas experimentais (métodos acústicos e ultrassônicos, *eddy-current* e métodos de térmicos). Estas técnicas experimentais requerem que um conhecimento *a priori* da localização do dano (sem muita precisão) e/ou que a porção da estrutura para inspeção seja acessível.

A necessidade de detecção quantitativa de danos aplicada a estruturas complexas (como as estruturas aero-espaciais), levaram ao desenvolvimento de métodos que avaliem as alterações nas características da vibração da estrutura. Uma revisão de várias técnicas pode ser vista nas referências [40, 138]. A idéia de que alteração na resposta da vibração poderá prover informação relativa ao dano é muito intuitiva. Entretanto, o problema mostrou-se muito difícil.

Aqui será apresentada a aplicação de uma metodologia que usa diretamente a resposta dinâmica da estrutura para a identificação de um ou mais danos em uma grande estrutura espacial. A técnica pode ser classificada como um método particular de atualização (*updating*) que combina o método do gradiente conjugado e a equação adjunta (abordagem variacional) e o método do algoritmo genético, ou seja, um esquema híbrido.

A figura 4.1 mostra a EEI e sua representação simplificada por um aestrutura de treliça plana com 68 graus de liberdade.

A estação espacial internacional (EEI) é um dos principais e mais importantes projetos de engenharia espacial da atualidade. É um esforço que envolve o consórcio de 15 países,

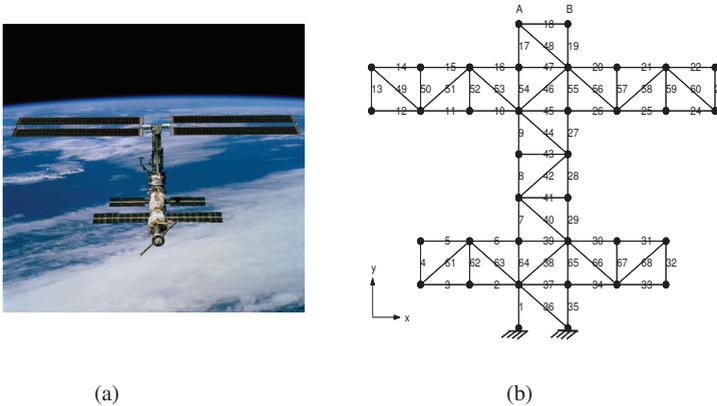


Figura 4.1: (a) Estação Espacial Internacional [99], (b) estrutura simplificada da EEI e numeração dos elementos.

sob a liderança dos Estados Unidos. A técnica será empregada na detecção de danos da estação espacial internacional¹.

A descrição aqui segue o trabalho publicado recentemente na revista *Acta Astronautica* [33].

4.2 Modelo Direto

Na análise da EEI, a versão simplificada da estação constitui-se de 72 barras de alumínio conectadas por nós. Cada elemento não danificado tem as seguintes características: comprimento $L = 6\text{ m}$, área da seção transversal $A = 80\text{ cm}^2$, densidade $\rho = 2700\text{ Kg} \cdot \text{m}^{-3}$, módulo de Young $E = 70\text{ GPa}$. A estratégia de discretização dos elementos finitos é aplicada e o comportamento mecânico da estrutura é representado por um conjunto de equações diferenciais:

$$\mathbf{M} \ddot{\mathbf{x}}(t) + \mathbf{C} \dot{\mathbf{x}}(t) + \mathbf{K} \mathbf{x}(t) = \mathbf{f}(t), \quad (4.2.1)$$

com as condições iniciais:

$$\mathbf{x}(0) = \mathbf{x}_0 \quad \text{and} \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0, \quad (4.2.2)$$

onde \mathbf{M} , \mathbf{C} e \mathbf{K} são as matrizes de massa, amortecimento e rigidez, respectivamente, $\mathbf{x}(t)$ é o vetor $(N \times 2)$ de deslocamento nodal (sendo N o número de nós) e o ponto superposto indica derivada em relação ao tempo. O vetor $\mathbf{f}(t)$ representa as forças externas aplicadas aos nós da estrutura. A matriz de amortecimento \mathbf{C} é assumida ser proporcional às matrizes de massa e de rigidez:

$$\mathbf{C} = \alpha \mathbf{M} + \beta \mathbf{K}. \quad (4.2.3)$$

Sob a hipótese acima, as equações da dinâmica podem ser desacopladas, usando a transformação modal

$$\mathbf{x} = \Phi \mathbf{y}, \quad (4.2.4)$$

¹ A sugestão desta aplicação partiu do prof. P. Gasbarri (Universidade de Roma “La Sapienza”) e foi realizada no período em que Leonardo D. Chiwiacowsky esteve no período de doutorado sanduiche

onde Φ é a matriz modal cujas colunas são auto-vetores. Substituindo as Eqs. (4.2.3) e (4.2.4) na Eq. (4.2.1) e pré-multiplicando por Φ^T , segue

$$\Phi^T \mathbf{M} \Phi \ddot{\mathbf{y}} + (\alpha \Phi^T \mathbf{M} \Phi + \beta \Phi^T \mathbf{K} \Phi) \dot{\mathbf{y}} + \Phi^T \mathbf{K} \Phi \mathbf{y} = \Phi^T \mathbf{f}, \quad (4.2.5)$$

sendo \mathbf{y} o vetor de amplitudes modais. Se os auto-vetores são orto-normalizados com respeito à matriz de massa tal que $\Phi^T \mathbf{M} \Phi = \mathbf{I}$ (onde \mathbf{I} é a matriz identidade) o seguinte resultado é obtido:

$$\ddot{\mathbf{y}} + \mathbf{D} \dot{\mathbf{y}} + \Omega^2 \mathbf{y} = \mathbf{f}^*, \quad (4.2.6)$$

onde Ω^2 é uma matriz diagonal de auto-valores ω_k^2 , que são o quadrado das frequências naturais, \mathbf{f}^* o vetor de forças generalizado (projeção sobre o espaço modal) e \mathbf{D} é a matriz modal de amortecimento:

$$\mathbf{D} = \alpha \mathbf{I} + \beta \Omega. \quad (4.2.7)$$

4.3 Análise Inversa

A técnica de problema inverso de estimação de dano localiza e e quantifica o dano a partir de medidas do deslocamento, que foi mencionada na Seção 2.7. Cada etapa do processo é detalhada a seguir, exceto o modelo direto, que já foi descrito pela Eq. (4.2.1).

4.3.1 O Problema de Sensitividade

O problema é formulado a partir de uma perturbação da matriz de rigidez, que altera o valor do deslocamento. São N_e parâmetros de rigidez desconhecidos (N_e é o número de elementos da estrutura discretizada). Para o caso sob consideração, p. ex.; o i -ésimo elemento da treliça, este coeficiente é representado pelo valor genérico $K_i = EA/L_e$; onde E é o módulo de Young's, A é a área da seção transversal e L_e é o comprimento do i -ésimo elemento finito. A matriz de rigidez da estrutura é uma função de K_i , i.e., $\mathbf{K} = f[\mathbf{K}]$, onde $\mathbf{K} = [K_1, \dots, K_{N_e}]$. Para deduzir o problema de sensitividade para cada função desconhecida K_i , cada parâmetro de rigidez deveria ser perturbado. Supondo que K_i é perturbado por uma quantidade pequena $\Delta K_i \delta_{ij}$, onde δ_{ij} é o delta de Kronecker e $j = 1, \dots, N_e$, como consequência uma pequena variação no deslocamento $\Delta x_{ij}(t)$ é obtida. A quantidade $\Delta x_{ij}(t)$ representa a variação do deslocamento em cada nó- i associado a variaçãp de rigidez do elemento j . O problema de sensitividade é obtido repassando no modelo direto, Eqs. (4.2.1)–(4.2.2), K_i por $K_i + \Delta K_i \delta_{ij}$, $x_i(t)$ by $x_i(t) + \Delta x_{ij}(t)$, e subtraindo da expressão resultante o problema direto original, negligenciando os termo de segunda ordem. Desta maneira, os N_e problemas de sensibilidade obtidos para cada parâmetro de rigidez perturbado, pode ser escrito como

$$\mathbf{M} \Delta \ddot{\mathbf{x}}_j(t) + \mathbf{C} \Delta \dot{\mathbf{x}}_j(t) + \mathbf{K} \Delta \mathbf{x}_j(t) = \Delta \mathbf{K}_j \mathbf{x}(t) \quad (j = 1, \dots, N_e) \quad (4.3.8)$$

onde $\Delta \mathbf{x}_j = [\Delta x_{1j}, \Delta x_{2j}, \dots, \Delta x_{N_j}]^T$, com as condições iniciais

$$\Delta \mathbf{x}_j(0) = 0 \quad \text{e} \quad \Delta \dot{\mathbf{x}}_j(0) = 0. \quad (4.3.9)$$

A matriz ΔK_j representa a contribuição da uma alteração associada a cada elemento K_j na matriz de rigidez completa. Consequentemente, estas matrizes tem elementos nulos, exceto aqueles com conexões com o elemento- j .

4.3.2 O Problema Adjunto e a Equação do Gradiente

O problema inverso é formulado como um problema de otimização, onde a função desconhecida \mathbf{K} deveria minimizar a it distância entre os deslocamentos calculados e os deslocamentos medidos, isto é minimizar o funcional $J[\mathbf{K}]$ definido por

$$J[\mathcal{K}] = \int_0^{t_f} [\mathbf{x}(\mathcal{K}, t) - \mathbf{x}^{exp}(t)]^T [\mathbf{x}(\mathcal{K}, t) - \mathbf{x}^{exp}(t)] dt, \quad (4.3.10)$$

onde t_f é o tempo final, $\mathbf{x}(t)$ and $\mathbf{x}^{exp}(t)$ são os deslocamentos calculados e medidos no tempo t , respectivamente. Para resolver o problema de otimização (4.3.10), o vetor multiplicador de Lagrange $\lambda(t)$ é usado para associar as restrições (4.2.1) a forma funcional

$$\begin{aligned} J[\mathcal{K}] &= \int_0^{t_f} [\mathbf{x}(\mathcal{K}, t) - \mathbf{x}^{exp}(t)]^T [\mathbf{x}(\mathcal{K}, t) - \mathbf{x}^{exp}(t)] dt + \\ &+ \int_0^{t_f} \lambda^T \{ \mathbf{M}\ddot{\mathbf{x}}(\mathcal{K}, t) + \mathbf{C}\dot{\mathbf{x}}(\mathcal{K}, t) + \mathcal{K}\mathbf{x}(\mathcal{K}, t) - \mathbf{f}(t) \} dt. \end{aligned} \quad (4.3.11)$$

A variação $\Delta J_j[\mathcal{K}]$ do funcional é obtida perturbando \mathcal{K} por $\Delta \mathcal{K}_j$ na Eq. (4.3.11) e isto implica (de acordo com a Eq. (4.3.8)) numa perturbação do deslocamento $\mathbf{x}(t)$ por $\Delta \mathbf{x}_j(t)$. Subtraindo a expressão resultante a Eq. original (4.3.11) e negligenciando os termos de segunda ordem, a variação do funcional segue

$$\begin{aligned} \Delta J_j[\mathcal{K}] &= \int_0^{t_f} 2 [\mathbf{x} - \mathbf{x}^{exp}]^T \Delta \mathbf{x}_j dt + \\ &+ \int_0^{t_f} \lambda^T \{ \mathbf{M}\Delta \ddot{\mathbf{x}}_j + \mathbf{C}\Delta \dot{\mathbf{x}}_j + \mathbf{K}\Delta \mathbf{x}_j + \Delta \mathbf{K}_j \mathbf{x} \} dt. \end{aligned} \quad (4.3.12)$$

Integrando por partes o lado direito da expressão acima e empregando as condições iniciais nulas do problema de sensibilidade, o seguinte problema adjunto é obtido (equação para os multiplicadores de Lagrange, o vetor $\lambda(t)$):

$$\mathbf{M} \ddot{\lambda}(t) - \mathbf{C} \dot{\lambda}(t) + \mathbf{K} \lambda(t) = 2 [\mathbf{x}^{exp}(t) - \mathbf{x}(t)], \quad (4.3.13)$$

com condição final

$$\lambda(t_f) = 0 \quad \text{and} \quad \dot{\lambda}(t_f) = 0. \quad (4.3.14)$$

Como o problema adjunto não tem dependência sobre a rigidez perturbada ($\Delta \mathbf{K}_j$), o subscrito j foi omitido.

O problema adjunto é um problema de valor final. Entretanto, o problema (4.3.13)-(4.3.14) pode ser transformado num problema de valor inicial, introduzindo uma nova variável: $\tau = t_f - t$. Então a técnica padrão do método de *Newmark* pode ser aplicada para a solução do problema transformado.

Após obter a equação do problema adjunto, a primeira variação do funcional pode ser deduzida das Eqs. (4.3.8) e (4.3.13):

$$\Delta J[\mathcal{K}] = \sum_{j=1}^N \int_0^{t_f} \lambda^T \Delta \mathbf{K}_j \mathbf{x} dt. \quad (4.3.15)$$

Por definição, a derivada direcional de $J[\mathcal{K}]$ na direção de um vetor $\Delta \mathbf{K}$ é dada por

$$\Delta J[\mathcal{K}] = \sum_{j=1}^N \mathbf{J}'_j \Delta \mathbf{K}_j, \quad (4.3.16)$$

onde \mathbf{J}' é a direção do gradiente do funcional $J[\mathbf{K}]$. Uma comparação das Eqs. (4.3.15) e (4.3.16) revela que a j -ésima componente da direção do gradient, \mathbf{J}'_j , é dada por

$$\mathbf{J}'_j[\mathcal{K}] = \int_0^{t_f} \lambda^\top \Delta \tilde{\mathbf{K}}_j \mathbf{x} dt, \quad (4.3.17)$$

onde $\Delta \tilde{\mathbf{K}}_j$ refere-se a j -ésima matriz de rigidez perturbada, i.e., $\Delta \tilde{\mathbf{K}}_j \equiv \partial[\Delta \mathbf{K}]/\partial K_j$.

4.3.3 O Método do Gradiente Conjugado

O procedimento iterativo do método do gradiente conjugado é usado para estimar a matriz de rigidez:

$$\mathcal{K}^{n+1} = \mathcal{K}^n - \beta^n \mathbf{P}^n, \quad n = 0, 1, 2, \dots, \quad (4.3.18)$$

onde β^n é vetor do tamanho do passo e \mathbf{P}^n é a direção do vetor descendente definido como

$$\mathbf{P}^n = \mathbf{J}'^n + \gamma^n \mathbf{P}^{n-1}, \quad \text{with } \gamma^0 = 0, \quad (4.3.19)$$

onde γ^n é o vetor coeficiente conjugado. O caso especial $\gamma^n = 0$, para qualquer n , corresponde ao método da máxima descida. Aqui, o vetor coeficiente conjugado é adotado como

$$\gamma^n = \frac{[\mathbf{J}'^n(\mathcal{K})]^2}{[\mathbf{J}'^{n-1}(\mathcal{K})]^2}, \quad n = 1, 2, \dots \quad (4.3.20)$$

O vetor do tamanho do passo β^n , Eq. (4.3.18), é determinado por minimizando o funcional $J[\mathcal{K}^{n+1}]$ dado pela Eq. (4.3.10) em relação β^n , i.e.

$$\min_{\beta^n} J[\mathcal{K}^{n+1}(t)] = \min_{\beta^n} \int_0^{t_f} [\mathbf{x}(\mathcal{K}^{n+1}, t) - \mathbf{x}^{exp}(t)]^2 dt. \quad (4.3.21)$$

Fazendo uma expansão em série de *Taylor* do integrando da Eq. (4.3.21) o valor de β^n para o mínimo pode ser avaliado analiticamente:

$$\beta^n = \left\{ \int_0^{t_f} [\Delta \mathbf{x}(t)]^\top [\Delta \mathbf{x}(t)] \right\}^{-1} \times \left\{ \int_0^{t_f} [\Delta \mathbf{x}(t)]^\top [\mathbf{x}(t) - \mathbf{x}^{exp}(t)] \right\} dt. \quad (4.3.22)$$

4.3.4 O Critério de Parada: Princípio da Discrepância

Na ausência dos erros experimentais (erros dos instrumentos de medida) um critério de parada seria:

$$J[\mathcal{K}] < \epsilon^*, \quad (4.3.23)$$

onde $J[\mathcal{K}]$ é definido pela Eq. (4.3.10) e ϵ^* é um valor pequeno especificado. Entretanto, nas aplicações, erros de medida estão sempre presentes. O *princípio da discrepância* [93, 130] pode ser aplicado para estabelecer um critério de parada. Seja o desvio padrão σ dos erros de medida independentes e comuns a todas os sensores e medidas, isto é,

$$\|\mathbf{x}(t) - \mathbf{x}^{exp}(t)\|_2 \cong \sigma. \quad (4.3.24)$$

Substituindo este resultado na Eq. (4.3.10), produz

$$\int_0^{t_f} \sigma^2 dt = \sigma^2 t_f \equiv \epsilon^2. \quad (4.3.25)$$

Então o princípio da discrepância para o critério de parada é tomado como:

$$J[\mathcal{K}] < \epsilon^2 = \sigma^2 t_f . \quad (4.3.26)$$

Com a aplicação deste critério de parada, o processo iterativo prossegue até $J[\mathcal{K}]$ atingir um valor abaixo dos erros de medida. A iteração é interrompida mesmo que o mínimo não tenha sido alcançado. De fato, se o processo iterativo continuar além desse limite, $J[\mathcal{K}]$ continuará a decrescer, porém o resultado da inversão irá se degradar.

4.3.5 Algoritmo Genético Epidêmico

As fases mais importantes nos algoritmos genéticos são a seleção (competição), reprodução, mutação e avaliação da função objetivo. A seleção é a operação para decidir que indivíduos da população serão usados para as ações de reprodução e mutação, para gerar uma nova população, na busca do ponto de ótimo. Reprodução (ou *cross over*) é o processo pelo qual o *material genético* dos pais é combinado para obter um ou mais descendente. Mutação é aplicada para introduzir uma alteração aleatória na genética do descendente. É testando os elementos da população na função objetivo é que se avalia a adequação dos indivíduos ao *ambiente* (solução ótima). Conduzindo a busca sobre todo o espaço de solução e associado ao operador de mutação, os algoritmos genéticos reduzem a chance de convergência em mínimos locais.

Tal como no método do gradiente conjugado, a aplicação de algoritmos genéticos para resolver o problema de identificação de danos é também um problema de minimização. O objetivo é minimizar o funcional $J[\mathcal{K}]$ da Eq. (4.3.10). Na implementação do AGE (ou *EPA*), o algoritmo opera sobre um tamanho fixo de população, gerada aleatoriamente. Os indivíduos desta população são variáveis reais que representam a função que o algoritmo está tentando otimizar (\mathcal{K}). Especificamente, os operadores evolucionários empregados são:

1. Seleção por torneio [90];
2. Cruzamento (*cross-over*) aritmético [90];
3. Mutação não-uniforme [90];
4. Epidemia [27, 88].

4.4 Detecção de Danos: Resultados Numéricos

Antes de estudar os efeitos dos danos sobre a estrutura e de como estimá-los, é interessante uma análise sobre as características dinâmicas da estrutura (e.g., as frequências naturais e modos normais de vibração) da configuração íntegra. Figura 4.1 mostra o esboço da treliça considerada (o número de cada elemento do modelo de elementos finitos está também representado na figura). A estrutura é considerada engastada na parte inferior. O cálculo dos auto-valores mostra que as frequências naturais da estrutura variam de um valor mínimo de 2,2 Hz a um valor máximo de 450 Hz. A tabela 4.1 registra os primeiros 15 valores de frequência.

Figura 4.2 mostra 2 modos de vibração: modo-1 e o modo-11. É possível perceber que para baixas frequências a estrutura comporta-se similar a uma viga (barras simples não estão envolvidas no movimento global da estrutura), composta por uma viga vertical e 2 vigas horizontais conectadas a viga vertical. Quando as frequências aumentam, os

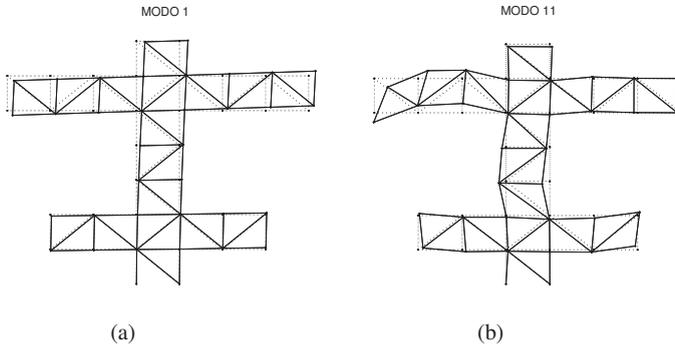


Figura 4.2: Esboço de modo-1 ($\omega_1 = 2.2$ Hz) e modo-11 ($\omega_{11} = 62.2$ Hz).

efeitos locais das barras tornam-se muito mais relevantes e os modos relevantes são aqueles associados a suas deformações (ver Figura 4.2b).

Agora o efeito de danos sobre o comportamento dinâmico da estrutura será considerado. Aqui, o efeito do dano é assumido com uma redução da rigidez original ($K_i, i = 1, \dots, N_e$, onde N_e é o número de elementos na treliça). Uma análise de freqüência foi efetuada e as Figuras 4.3-4.4 mostram a variação percentual das freqüências induzidas pelo dano. A intensidade e diferentes localizações dos danos são mostrados nas Tabelas 4.2 e 4.3.

É bem conhecido que o deslocamento da freqüência não assume o mesmo valor para todas as freqüências, como mostrado nas Figuras 4.3 a 4.4. Este efeito é relevante quando o algoritmo de detecção de danos; pois o método é aplicado no domínio do tempo, isto força considerar todos os graus de liberdade da estrutura, e.g., integrar diretamente as equações do movimento sem aplicar a técnica de redução modal, ver see Eq. (4.2.6), que, de um lado poderia nos ajudar a reduzir a ordem do sistema (reduzindo o tempo computacional), mas por outro lado reduz também informações associadas ao comportamento dinâmico da estrutura, necessárias para a caracterização da própria estrutura. Uma história temporal típica é apresentada na Figura 4.5, onde os deslocamentos \mathbf{u} ao longo da direção horizontal e \mathbf{v} ao longo da direção vertical. Em toda a análise realizada, somente o coeficiente α do amortecimento estrutural foi considerada. Em particular foi assumido $\alpha = 1.7$, que corresponde ao amortecimento crítico $\zeta_1 = 6\%$ sobre o modo-1, $\zeta_2 = 2\%$ sobre o modo-2, $\zeta_3 = 1\%$ sobre o modo-3, e os outros valores de amortecimento decrescem de acordo com

No. do modo	ω (Hz)	No. do modo	ω (Hz)
1	2.24	9	48.49
2	6.90	10	50.61
3	12.66	11	62.19
4	14.48	12	62.55
5	18.43	13	72.27
6	26.38	14	74.07
7	28.64	15	83.88
8	40.24		

Tabela 4.1: Primeiras 15 freqüências naturais da estrutura não danificada.

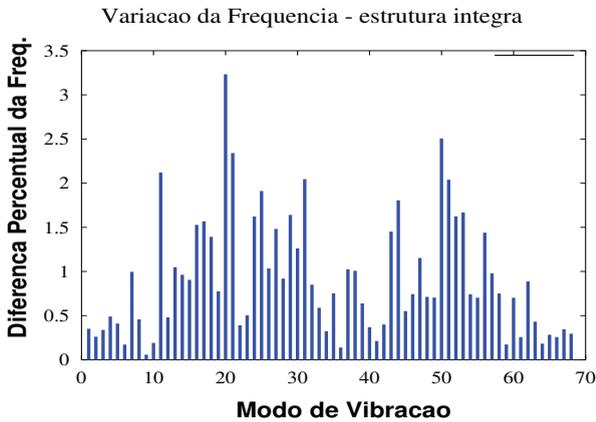


Figura 4.3: Seis elementos danificados: 15, 29, 42, 58, 62, 68.

Element	42	29	58	15	68	62
% Reduction	5%	10%	15%	20%	25%	35%

Tabela 4.2: Seis elementos danificados e redução de rigidez.

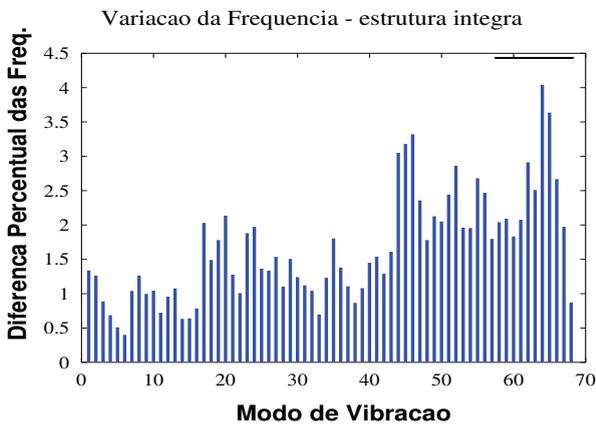


Figura 4.4: Vinte elementos danificados: 4, 8, 10, 15, 17, 24, 29, 30, 32, 35, 42, 43, 46, 50, 52, 55, 58, 60, 62, 68.

Element	30	8, 24, 42 46, 55, 68	4, 10, 29 35, 52, 60	17, 32 50, 58	15, 43 62
% Reduction	2%	5%	10%	15%	20%

Tabela 4.3: Vinte elementos danificados e redução de rigidez.

a Eq. (4.2.3) quando a frequência aumenta.

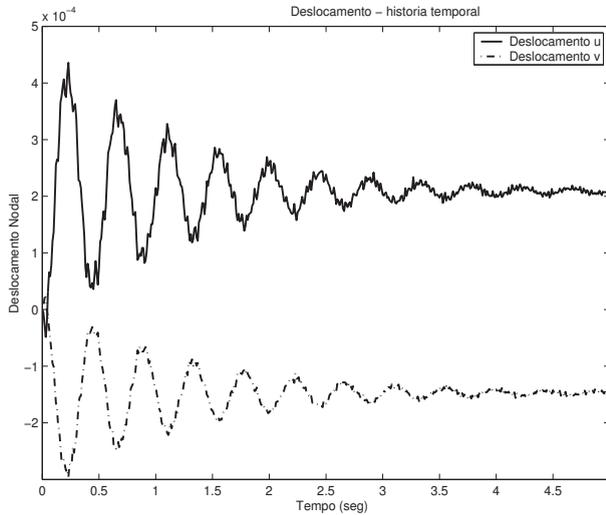


Figura 4.5: História temporal do nodo-2 para a estrutura com vinte elementos danificados.

A integração numérica da Eq. (4.2.1) foi realizada usando o método de *Newmark* [31, 33], assumindo o tempo final de $t_f = 5$ s e passo de tempo $\Delta t = 5.0 \times 10^{-2}$ s. Forças externas foram aplicadas aos nós *A* e *B* (Figura 4.1), com componentes $F_x^{(A)} = 10$ N, $F_y^{(A)} = 5$ N, $F_x^{(B)} = 10$ N e $F_y^{(B)} = 20$ N, e condições iniciais: $\mathbf{x}(0) = 0$ e $\dot{\mathbf{x}}(0) = 0$.

Dados experimentais sintéticos são obtidos da solução exata do problema direto adicionando-se uma perturbação aleatória, da seguinte forma:

$$\mathbf{x}^{exp}(t) = \mathbf{x}(t) [1 + \sigma \mathcal{R}], \quad (4.4.27)$$

onde σ é o desvio padrão dos erros de medida e \mathcal{R} é uma variável aleatória com distribuição gaussiana, tal que $\mathcal{R} \sim \mathcal{Normal}(0;1)$. O valor adotado para σ é para produzir um ruído com 1% de desvio médio nos dados (este nível de ruído é somente para teste do algoritmo, na prática o valor pode ser mais alto). O critério de parada é fixado usando a Eq. (4.3.23) ($\epsilon^* = 10^{-10}$, para o caso de dados sem ruído). A Eq. (4.3.26) é aplicada para o caso de dados com ruído.

A abordagem variacional foi avaliada através da inversão de diferentes problemas de detecção de danos, onde a localização, intensidade e quantidade dos elementos danificados foram calculados. Os cenários de danos adotados são apresentados nas Tabelas 4.2-4.3 e para avaliar a qualidade da estimacção, o fator de dano foi definido como

$$DF_i = \frac{K_i^u - K_i^d}{K_i^u} \quad i = 1, \dots, N_e \quad (4.4.28)$$

onde K_i^u e K_i^d são parâmetros para a estrutura não danificada e danificada, respectivamente.

Figuras 4.6 e 4.7 apresentam uma comparação entre valores de fator de danos estimado e exato, quando dados sem ruído são empregados. A abordagem variacional padrão foi bem sucedida com uma perfeita estimacção de danos.

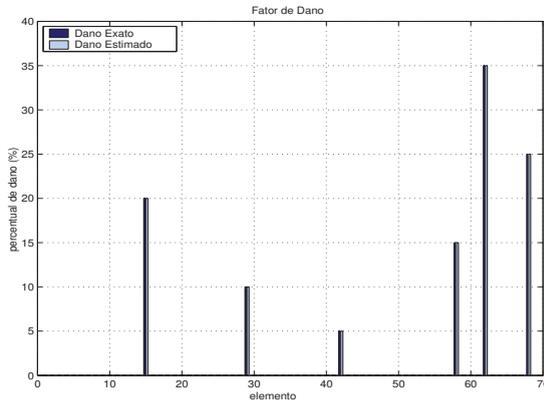


Figura 4.6: Fator de dano estimado com 6 elementos danificados – ver Tabela 4.2 (dados sem ruído).

Num caso mais realístico, o problema da estimação de danos é resolvido com dados experimentais corrompidos por ruído. Se alguma informação adicional está disponível, como a localização e/ou a dimensão do dano, deveria ser usada para garantir uma boa estimativa inicial para o método do gradiente conjugado e assim, assegurar a convergência. Entretanto, este tipo de informação pode não ser disponível. Assim, uma outra metodologia deve ser empregada para escolher uma melhor estimativa inicial para o processo iterativo.

O uso de dados contaminados com ruído na detecção de danos estruturais, tem-se observado que a estimativa inicial para o método do gradiente conjugado **não** pode ser arbitrária, pois não ocorre convergência.

Por exemplo, tomando o cenário de danos descrito pela Tabela 4.2, a hipótese de usar como estimativa inicial a configuração não danificada mostra-se inadequada. Figura 4.8 apresenta uma comparação entre os valores exatos e estimados do fator de danos, quando a configuração não danificada foi adotada como iniciativa inicial para o caso de 6 elementos danificados. Os resultados não são satisfatórios.

Para contornar esta dificuldade, o método estocástico de algoritmo genético (GA: *Genetic Algorithm*), foi adotado para calcular uma solução primária, que é assumida como estimativa inicial para o método do gradiente conjugado. Com esta abordagem híbrida, boas estimativas foram calculadas. As Figuras 4.9 e 4.10 apresentam uma comparação entre os valores do fator de dano estimados e exatos, no caso em que os valores de rigidez estimados foram maiores do que seus valores não danificados, esta estimativa foi negligenciada.

Em relação a qualidade da estimação dos danos, há uma certa dificuldade em se determinar os valores exatos. O critério de parada (princípio da discrepância) causa uma interrupção do processo de convergência, antes que o ruído nos dados experimentais possam corromper as estimativa de rigidez. devido a esta interrupção, alguns coeficientes de rigidez estimados não apresentaram convergência completa aos valores exatos, como fica evidente nas Figuras 4.9 e-4.10. Contudo, os resultados quanto a localização e quantificação danos podem ser considerados satisfatórios.

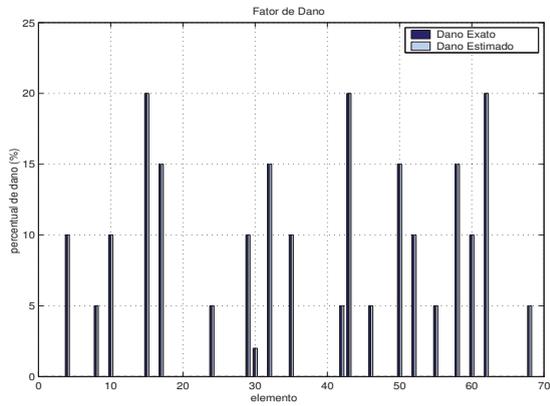


Figura 4.7: Fator de dano estimado com 6 elementos danificados – ver Tabela 4.3 (dados sem ruído).

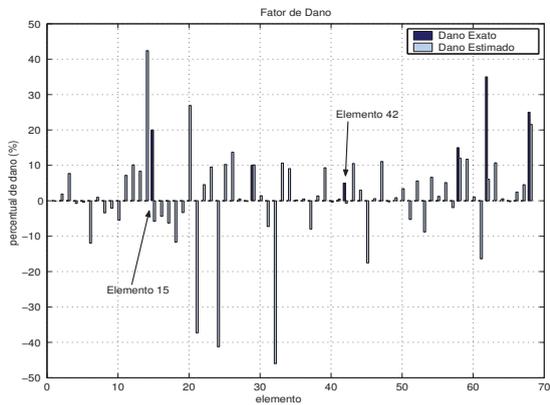


Figura 4.8: Estimação de danos de 6 elementos com uma estimativa inicial arbitrária.

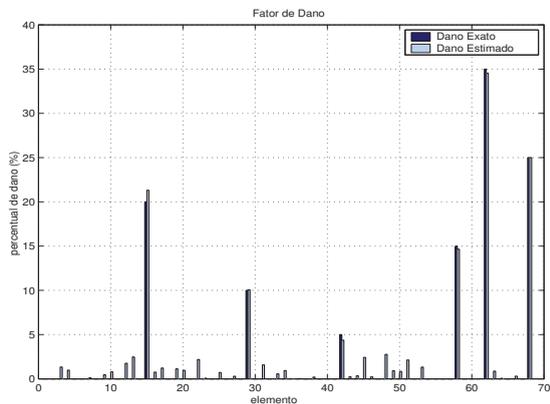


Figura 4.9: Fator de dano estimado com 6 elementos danificados – ver Tabela 4.2 (dados com ruído).

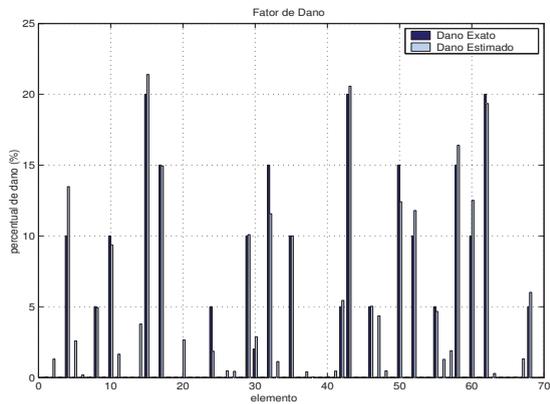


Figura 4.10: Estimação de danos com 20 elementos danificados – ver Tabela 4.3 (dado experimental com ruído).

Capítulo 5

Problemas Inversos em Aplicações Espaciais

Há um interesse permanente do ser humano em entender o cosmos onde estamos inseridos. O estudo da astronomia nos seus primórdios estava contaminado com questões sobrenaturais. A medida que o conhecimento humano foi se desenvolvendo e amadurecendo, a investigação da dinâmica dos corpos celestes foi se dando em bases mais científicas, desde das *Leis de Kepler* até culminar com a cosmologia de Isaac Newton, de um Universo estático e infinito. Nos tempos modernos, houve um grande avanço da astrofísica, com uma melhor compreensão da evolução estelar, do movimentos dos planetas, asteróides, cometas, galáxias e estruturas de grande escala. Foi graças ao desenvolvimento de modernos telescópios, que hoje temos um modelo padrão de um cosmos dinâmico: o Universo em expansão. No final do século XX, os cosmólogos descobriram que o Universo está em expansão *acelerada* (e não desacelerada, como se acreditava): para explicar este comportamento sugeriu-se a introdução do conceito de *energia escura* (uma forma de energia gravitacionalmente repulsiva).

A tecnologia espacial tem estado presente nos desenvolvimento da ciência espacial: astronomia, astrofísica, física solar e em muitas outras áreas, através de lançamentos de satélites científicos e o lançamento do telescópio espacial Hubble. Porém, estas são aplicações científicas da engenharia espacial, que é um item mandatório de qualquer programa espacial.

Porém, quando se menciona *aplicações espaciais*, estamos nos referindo a aplicações permanentes, como sistema de tele-comunicações, sistemas globais de navegação por satélites (GNSS: *Global Navigation Satellite Systems*, hoje só existe o sistema americano totalmente operacional: o GPS: *Global Positioning System*), satélites meteorológicos e satélites de observação da Terra (voltados para a avaliação e monitoramento dos recursos naturais, circulação oceânica, dinâmica da criosfera e, mais recentemente, satélites ambientais de monitoramento de gases minoritários da atmosfera).

Neste capítulo serão trabalhadas 2 aplicações: estimação de perfis verticais de temperatura e umidade atmosféricas a partir de dados de satélite; e, identificação de propriedades óticas do oceano.

Antes de encerrar este preâmbulo, é importante ressaltar que atividades até recentemente de cunho estritamente científico (como observação da atividade solar e ejeções de massa coronal, por exemplo), irão desempenhar um papel mais relevante no século XXI, com o desenvolvimento da atividade de Clima Espacial (*Space Weather*). De forma muito resumida, clima espacial é uma atividade complexa que pretende englobar desde o moni-

toramento, entendimento e previsão da atividade solar, sua propagação através do espaço, para finalmente interagir com o planeta Terra. Durante erupções solares e/ou ejeções de massa coronal, partículas energéticas, propagam-se pela heliosfera, atingindo a magnetosfera terrestre, para então alterar as propriedades da ionosfera. Entender, modelar e desenvolver uma ciência para fazer previsões de toda esta cadeia de processos é o objetivo final do clima espacial. O INPE está iniciando uma nova atividade que pretende estabelecer um Centro de Previsão para clima espacial: pode-se acessar o sítio do clima espacial diretamente da página principal do INPE na internet (<http://www.inpe.br>). O quadro e a importância que marcaram a previsão numérica do tempo no século XX, deverá haver um paralelo para o Clima Espacial no século XXI.

5.1 Meteorologia: Estimação de Temperatura Atmosférica com dados de satélite

No início do século XX, a ciência fez a humanidade mudar sua visão do mundo e do Universo. Em 1900 Max Planck propunha a hipótese quântica e em 1905, Albert Einstein apresentou a teoria da relatividade restrita. Um ano antes, 1904, a previsão do tempo era finalmente reconhecida como um problema de valor inicial (teorema do meteorologista norueguês Vilhelm Bjerknes). Desde então, o desenvolvimento da meteorologia foi meteórico: em 1922 Lewis Fry Richardson publica o livro *Weather Prediction by Numerical Process* (o procedimento fracassou por várias razões, uma delas é que a análise numérica não tinha se desenvolvido e não se sabia sobre as restrições que deveriam ser impostas às discretizações numéricas).

No período de 1941 a 1943, a Inglaterra desenvolveu o primeiro computador digital (projeto Colossus), quase ao mesmo tempo os Estados Unidos no período de 1943 a 1945 desenvolveu o computador digital ENIAC. A primeira aplicação do ENIAC foi realizar o sonho de Richardson: fazer previsão do tempo por processos numéricos. Isso mudou definitivamente a meteorologia e os benefícios que uma previsão mais precisa do tempo e do clima vem trazendo para a sociedade são incalculáveis. Porém, desde que a primeira previsão numérica do tempo foi realizada até os anos 80, um aspecto persistiu: a previsão para o hemisfério norte era significativa e sistematicamente melhor do que a previsão para o hemisfério sul. O desenvolvimento do computador, dos métodos numéricos e melhor entendimento do fenômeno (ou ainda, melhores *parametrizações*, como denominam os profissionais de modelos computacionais para previsão numérica do tempo, para turbulência, micro-física de nuvens e outros processos) não estavam resolvendo este hiato de previsão. No final da década de 70 apareceram os satélites ambientais, que permitiam obter informações sobre todo o planeta – isto foi crucial para garantir informações em partes do planeta desabitadas (florestas tropicais, os pólos do planeta e sobre os oceanos). Com a introdução destas novas informações e o contínuo desenvolvimento obtidas com a tecnologia espacial, em meados de 2002 finalmente as previsões para ambos os hemisférios tornaram-se equivalentes em qualidade.

A radiação do corpo negro é uma das muitas evidências que dão suporte a hipótese quântica. Neste caso, a energia radiante está relacionada com a temperatura do corpo (Lei de Planck):

$$B_{\lambda}(T) = \frac{2hc^2}{\lambda^5 [e^{(hc)/(\lambda k_B T)} - 1]} \quad (5.1.1)$$

sendo B_{λ} a energia irradiada, T a temperatura absoluta, λ o comprimento de onda, c a velocidade da luz no vácuo e h , k_B as constantes de Planck e Boltzmann. A Eq. (5.1.1) só foi deduzida graças a hipótese de Planck da quantização da luz, isto é, a energia luminosa

é propagada em *pacotes* (ou *partículas*): $E = nh\nu$ ($\lambda = 1/\nu$).

A equação de Planck para a radiação do corpo negro, mostra que a temperatura atmosférica pode ser inferida pela radiação medida por um satélite. Esta proposta foi formulada na metade da década de 1950 [74].

A transferência diferencial da intensidade de radiação monocromática I_λ pode ser matematicamente expressa como (aqui vamos usar a notação padrão dos textos de meteorologia, ver por exemplo [80]):

$$dI_\lambda(t, p, s) = -(\kappa_\lambda + \sigma_\lambda) I_\lambda \rho ds + j_\lambda \rho ds \quad (5.1.2)$$

onde t , p e s são o tempo, a pressão e a direção do feixe incidente e ρ é a massa específica do ar. Os coeficientes κ_λ , σ_λ e $j_\lambda = \kappa_\lambda B_\lambda(T)$ representam a absorção, espalhamento e termo de fonte, respectivamente. Uma simplificação importante para a Equação de Transferência Radiativa (ETR) é quando não ocorre espalhamento (equação de Schwarzschild), que na geometria plano-paralela é dada por:

$$\mu \frac{dI_\lambda(\tau, \mu)}{d\tau} = I_\lambda(\tau, \mu) - B_\lambda(T) \quad (5.1.3)$$

onde $\mu \equiv \cos \theta$, com θ o ângulo de espalhamento polar e $\tau = \int_z^\infty k_\lambda \rho dz'$ é profundidade ótica. A formulação integral da ETR para um sensor posicionado no zênite ($\mu = \cos \theta \approx 1$) é dada por

$$I_\lambda(\tau) = I_\lambda(\tau_1) e^{-(\tau_1 - \tau)} - \int_\tau^{\tau_1} B_\lambda[T(\tau')] d[e^{(\tau' - \tau)}] \quad (5.1.4)$$

sendo que $\tau = \tau_1$ para $z = 0$ e no topo da atmosfera: $z \rightarrow \infty$ e $\tau = 0$.

Definindo-se a *transmitância monocromática* como:

$$\mathfrak{S}(\lambda, z) \equiv e^{-\tau} \quad (5.1.5)$$

a radiação medida por um sensor a bordo de um satélite pode ser aproximada por:

$$I(\lambda_i) = B[\lambda_i, T(p_s)] \mathfrak{S}(\lambda_i, p_s) + \int_{p_s}^{p_0} B[\lambda_i, T(p)] \frac{\partial \mathfrak{S}(\lambda_i, p)}{\partial p} dp \quad (5.1.6)$$

escrita em coordenada de pressão, sendo p_s a pressão à superfície e p_0 a pressão no topo da atmosfera. A Eq. (5.1.6) é uma equação de Fredholm de primeira espécie, ou seja, a inversão da ETR é uma problema mal-posto, pois viola a 3a. condição de Hadamard. O modelo direto (5.1.6) é resolvido com uma aproximação de diferenças finitas centradas:

$$I_i = B_{i,s}(T_s) \mathfrak{S}_{i,s} + \sum_{j=1}^{N_p} \left(\frac{B_{i,j} + B_{i,j-1}}{2} \right) [\mathfrak{S}_{i,j} - \mathfrak{S}_{i,j-1}] \quad (5.1.7)$$

sendo $i = 1, 2, \dots, N_\lambda$, $j = 1, 2, \dots, N_p$, com $I_i \equiv I_{\lambda_i}(0)$, N_λ é o número de canais no radiômetro do satélite e N_p é o número de camadas atmosféricas consideradas.

A sondagem remota da atmosfera por satélites meteorológicos é uma questão crucial para previsão do tempo para o hemisfério sul. Os satélites da série TIROS-N (*Television and Infrared Observational Satellite*) da NOAA (*National Oceanic and Atmospheric Administration*) conta hoje com sensores do tipo AVHRR (*Advanced Very High Resolution Radiometer*) e MSU (*Microwave Sounding Unit*). Várias técnicas de inversão tem sido propostas para dados de satélites [122] da classe TOVS (*TIROS Operational Vertical Sounder*). Um dos pacotes de inversão disponíveis é o ITPP-5.0 (*International TOVS Processing Package*), cuja a versão incorpora algoritmos para o sensor AVHRR, constituído de um conjunto de câmeras que utilizam 2 bandas do CO₂, sete canais na banda de 15 μm e seis canais na banda 4,3 μm .

Pressão (hPa)	EMQ Max-Ent-2	NE-MaxEnt-2 $q = 0.5$	NE-MaxEnt-1 $q = 0.5$
50-0.1	13,248	13,646	9,598
250-50	7,442	8,723	5,818
500-250	5,216	5,508	5,532
700-500	1,283	1,817	1,478
1000-700	4,428	3,475	5,334

Tabela 5.1: Erro médio quadrático (EMQ).

5.1.1 Inversão com Regularização

Conforme visto na Seção 2.1, o problema inverso pode ser formulado como um problema de otimização, com o objetivo de minimizar o funcional de Tikhonov (2.1.2). O operador de regularização aqui será a entropia (extensiva) de 2a. ordem (técnica MaxEnt-2) [21, 22, 115] e a entropia não extensiva de Tsallis [17] (NE-MaxEnt-n).

Como resultado do teorema da unificação, a forma da entropia não extensiva de ordem- γ é definida como

$$S_q^\gamma \equiv \frac{k}{q-1} \left[1 - \sum_{i=1}^{N_p} r_i^q \right]; \quad (5.1.8)$$

onde o vetor $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_{N_p}]^T$ é dado por:

$$\mathbf{r} = \Delta^\gamma \mathbf{p} \quad (5.1.9)$$

onde $\gamma = 0, 1, 2, \dots$, \mathbf{p} é o vetor de termos desconhecidos (temperatura ou umidade) e Δ é o operador de diferença. O operador MaxEnt-2 pode ser deduzido de (5.1.8) e (5.1.9) impondo $\gamma = 2$ e $q = 1$. Um pequeno valor deve ser adicionado ao valor do operador de diferença ($\varsigma = 10^{-15}$) para assegurar uma quantidade bem definida para todos os valores de q .

A Figura 5.1a mostra a inversão de temperatura atmosférica a partir de dados de sensor HIRS-2 do satélite NOAA-14 com estimativa inicial a partir de médias climatológicas para o TOVS. Figura 5.1b é a mesma inversão, mas considerando uma estimativa inicial de um perfil uniforme e idêntico a da temperatura do solo. Esta figura mostra que o método implícito baseado em MaxEnt-2 é mais robusto em relação perfil inicial usado no processo iterativo de inversão do que o método instalado no ITPP-5.

Os resultados para MaxEnt-0 não-extensiva (simplesmente: NE-MaxEnt) não foram bons [17]. Entretanto, resultados com NE-MaxEnt-1 e NE-MaxEnt-2 são similares aos obtidos com MaxEnt-2. Porém, na camada mais próxima ao solo (1000 a 700 hPa), os resultados com $q = 0,5$ (outros valores foram testados, mas não melhoram a inversão) para NE-MaxEnt-2 foram os melhores obtidos considerando as técnicas de regularização baseadas em entropia [17] (ver Tabela 5.1).

5.1.2 Inversão por Redes Neurais Artificiais

O mesmo problema tratado na Seção 5.1.1 pode ser resolvido com a aplicação de redes neurais artificiais. O primeiro resultado foi obtido com a rede perceptron de múltiplas camadas. Um aspecto importante é o conjunto de dados usados no treinamento. Três tipos de dados foram usados para treinamento das redes: (i) dados sintéticos SDB1, com radiancias calculadas com o modelo direto (5.1.7) a partir de perfis de temperatura conhecidos;

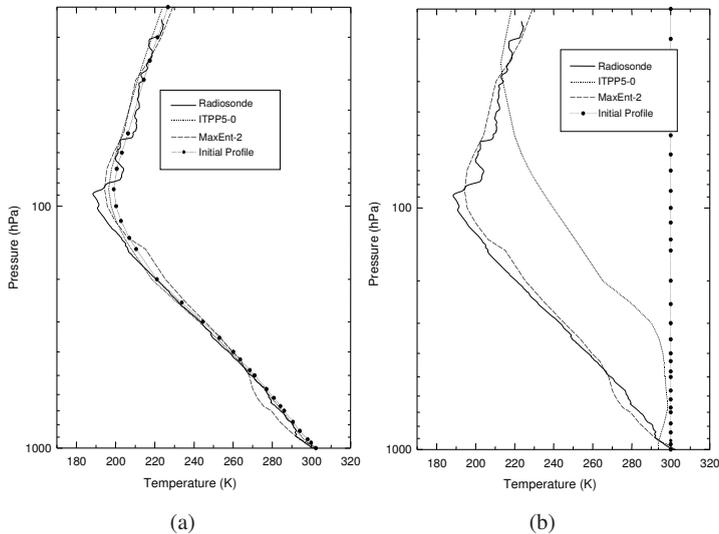


Figura 5.1: Inversão de temperatura atmosférica com MaxEnt-2 e ITTP-5 usando dados de satélite NOAA-14 com 2 tipos de estimativa inicial: (a) TOVS; (b) perfil uniforme.

(ii) banco de dados TIGR (*TOVS Initial Guess Retrieval*): uma biblioteca de situações atmosféricas disponível na internet; (iii) um banco de dados constituído pela união dos anteriores: SDB1+TIGR. A melhor inversão é obtida com o conjunto de treinamento (iii). A figura 5.2a mostra a inversão com RNA-PMC [131]. Resultados similares são obtidos com redes neurais do tipo FBR [133]. Um resultado muito significativo vem com a definição de uma nova rede neural do tipo FBR: a *FBR não extensiva*. A rede FBR não extensiva foi também aplicada para o problema da inversão de temperatura, melhorando os resultados obtidos com a rede FBR padrão [134].

A importância do uso de redes neurais se deve também por um aspecto de grande relevância: é possível implementar as redes neurais em dispositivos de hardware, os neurocomputadores. Este tipo de computação (uma máquina não von Neumann) pode ser implementada em sistemas de hardware configuráveis: FPGA (*Field Programmable Gate Array*), cuja configuração pode ser realizada com o uso de VHDL (*VHSIC Hardware Description Language*, VHSIC: *Very-High-Speed Integrated Circuit*). Algumas linhas de computadores de alto desempenho (computação paralela) estão desenvolvendo máquinas híbridas: parte do processamento realizado por software e parte por hardware. Em outubro de 2004, a Cray Research lançou a linha XD1: computadores multi-processados com dispositivos de FPGA incluídos no sistema, com capacidade de processamento em ponto flutuante (em muitas versões, as FPGAs operavam somente com ponto fixo, sendo esta uma das limitações apontadas para a tecnologia) – atualmente a linha XD1 foi substituída pela família XT5.

A tecnologia de FPGA irá permitir que o processamento seja feito a bordo dos satélites, com dispositivos acoplados diretamente aos radiômetros (uma nova geração de sensores inteligentes), ao invés do processamento computacionalmente dispendioso que é realizado hoje em dia com os procedimentos de inversão em grandes computadores de alto desem-

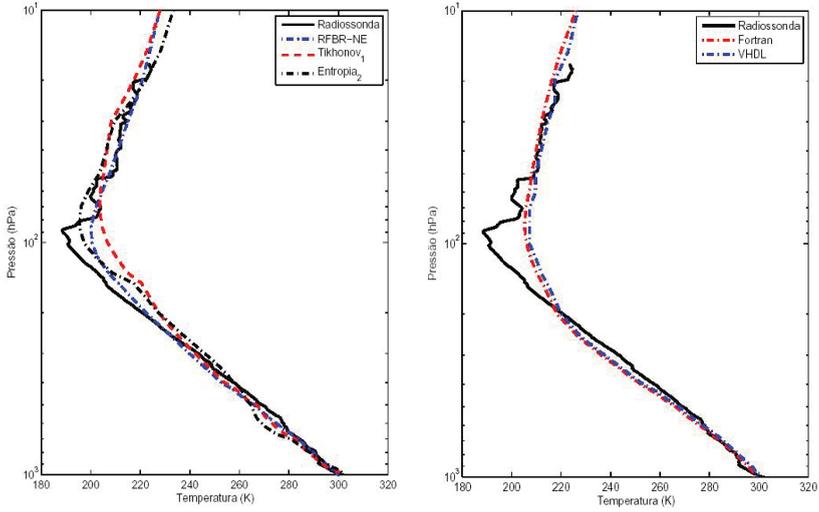


Figura 5.2: Inversão para a temperatura atmosférica usando dados do satélite NOAA-14 (treinamento com SDB1+TIGR): (a) RNA-PMC, (b) software \times hardware.

penho. A figura 5.2b mostra a inversão por redes neurais com software e com hardware (versão VHDL) [19].

5.2 Oceanografia: Propriedades Óticas

O programa de pesquisa em ótica hidrológica inversa junto ao grupo do LAC/INPE tem mais de uma década de atividade, identificando inicialmente fontes de bioluminescência [146, 147] a partir de dados de radiância *in situ*. McCormick [87] e Gordon [52] apresentam artigos de revisão nesta área importante com aplicações em ciência e engenharia.

O problema direto é dado pela equação linear de Boltzmann do transporte de fótons no interior do corpo d'água – a figura 5.3 ilustra o problema direto. Em qualquer ponto do meio, a luz pode ser absorvida, espalhada ou transmitida, de acordo com os coeficientes de absorção (a_λ), espalhamento (b_λ) e a função de fase de espalhamento. Um coeficiente de atenuação é c_λ é definido como $c = a + b$ e, como antes, a profundidade é substituída pela espessura ótica τ . Assumindo geometria plano-paralela, a Eq. de Boltzmann 1D (ETR) pode ser escrita para um dado comprimento de onda como (aqui será usada a notação típica da área, ver [91])¹:

$$\mu \frac{\partial}{\partial \tau} L(\tau, \mu, \varphi) + L(\tau, \mu, \varphi) = \frac{\varpi_0}{4\pi} \int_{-1}^1 \int_0^{2\pi} \beta(\tau, \mu, \varphi; \mu', \varphi') L(\tau, \mu', \varphi') d\varphi' d\mu' + S(\tau) \quad (5.2.10)$$

onde τ é a profundidade ótica, $\mu \equiv \cos \theta$, θ e φ são os ângulos incidentes polar e azimutal, respectivamente; $\varpi_0 = b/c$ é o albedo de espalhamento simples e $S(\tau)$ é termo de fonte. A

¹A dedução desta equação foi apresentada por Souto [139].

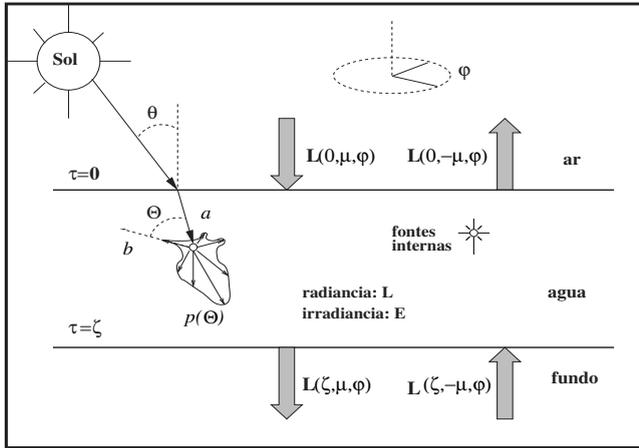


Figura 5.3: Diagrama esquemático do problema direto em ótica hidrológica.

função de fase de espalhamento $\beta(\tau, \mu, \varphi; \mu', \varphi')$ representa a distribuição angular do feixe luminoso. Para simplificar a notação, a dependência com o comprimento de onda λ foi omitida, quando necessário será explicitado. A equação está sujeita as seguintes condições de contorno

$$\begin{aligned} L(0, \mu, \varphi) &= f(\mu, \varphi) \\ L(\tau_{\max}, -\mu, \varphi) &= g(\mu, \varphi) \end{aligned}$$

Em nossos estudos, três técnicas foram aplicadas na solução numérica da Eq. (5.2.10):

Inserção Invariante: Sua descrição é feita por Mobley [91]. Nesta abordagem, a ETR para um corpo de água é transformado num problema de valor inicial não linear (equação matricial de Riccati, resolvida por um bem estabelecido código de solução de equações diferenciais ordinárias. O pacote Hydrolight [92] permite a escolha de vários tipos de fontes internas e diferentes condições do tempo.

Método S_N Analítico: Aqui, primeiramente a radiancia é separada em componentes espalhada e não-espalhada. A solução para a componente não espalhada é dada por uma expressão simples, enquanto que a componente espalhada é expandida em soluções elementares das equações de ordenadas discretas. Os coeficientes da expansão são obtidos resolvendo o sistema de equações algébricas lineares associado. Uma descrição completa do método é feita na referência [23].

Método LTS_N : Este esquema apareceu no início dos anos 90 no contexto de transporte de neutrons e foi mais tarde aplicado a problemas de transferência radiativa [126]. Convergência do método foi estabelecida usando teoria de semi grupo- C_0 [111].

Em todos os métodos citados, o tratamento da variável azimutal se dá através da decomposição espectral de Fourier:

$$L^\mp(\tau, \mu, \varphi) = \sum_{m=0}^{N_\varphi} [L_{1,m}^\mp(\tau, \mu, \varphi) \cos(m\varphi) + L_{2,m}^\mp(\tau, \mu, \varphi)] \quad (5.2.11)$$

$$\beta^{\pm}(\tau, \mu', \varphi'; \mu, \varphi) = \sum_{k=0}^{N_{\beta}} \beta_k^{\pm}(\tau, \mu; \mu') \cos[k(\varphi' - \varphi)] \quad (5.2.12)$$

onde o sinal “+” denota radiâncias direcionadas para baixo e o sinal “-” para radiâncias direcionadas para cima, os subscritos 1,2 denotam as amplitudes espectrais. A dependência em relação ao ângulo azimutal φ é expressa pelos m -modos. Ou seja, a decomposição espectral de Fourier, transformada a equação integro-diferencial (5.2.10) com dependência em μ e φ , num sistema de N_{ϕ} equações integro-diferenciais com dependência em μ somente.

A precisão e eficiência dos diferentes métodos para resolver a equação (5.2.10) foi investigada por Chalhoub et al. [25] e o desempenho da versão paralela dos códigos foi realizada por Souto et al. [141].

5.2.1 O Método LTS_N

O método das ordenadas discretas é uma esquema de colocação, onde a variável angular μ é discretizada em N_{μ} direções (ou ordenadas) discretas [26]. A função de fase de espalhamento é escrita como expansão em polinômios de Legendre:

$$\begin{aligned} \mu \frac{\partial}{\partial \tau} L^m(\tau, \mu) + L^m(\tau, \mu) = \\ \frac{\varpi_0}{2} \sum_{l=m}^{N_{\beta}} \omega_l^m P_l^m(\mu) \int_{-1}^1 P_l^m(\mu') L^m(\tau, \mu') d\mu' + S^m(\tau) . \end{aligned} \quad (5.2.13)$$

onde N_{β} é o grau de *anisotropia* do meio. Finalmente, a integral de espalhamento é discretizada por uma regra de quadratura gaussiana. Isto é, a Eq. (5.2.13) é escrita como:

$$\begin{aligned} \mu_j \frac{d}{d\tau} L^m(\tau, \mu_j) + L^m(\tau, \mu_j) = \\ \frac{\varpi_0}{2} \sum_{l=m}^{N_{\beta}} \omega_l^m P_l^m(\mu_j) \sum_{i=1}^{N_{\eta}} \eta_i P_l^m(\mu_i) L^m(\tau, \mu_i) + S^m(\tau) \end{aligned} \quad (5.2.14)$$

onde $j = 1, 2, \dots, N_{\mu}$ e η_i é o peso da quadratura gaussiana. Deste modo, a equação integro-diferencial (5.2.10) é transformada num sistema de $N_{\phi} \times N_{\mu}$ equações diferenciais.

No problema inverso deste capítulo, o método LTS_N será empregado para resolver o problema direto associado. Neste método, aplica-se a transformada de Laplace na equação de ordenada discreta (5.2.14), resultando num sistema de equações algébricas sobre a variável complexa s :

$$\begin{aligned} s \bar{L}^m(s) + \frac{1}{\mu_j} \bar{L}^m(s) - \frac{\varpi_0}{2\mu_j} \sum_{l=m}^{N_{\beta}} \beta_l^m P_l^m(\mu_j) \sum_{i=1}^{N_{\beta}} \eta_i P_l^m(\mu_i) \bar{L}^m(s) = \\ L_j^m(0) + \frac{1}{\mu_j} \bar{S}^m(s) \end{aligned} \quad (5.2.15)$$

sendo $\bar{L}^m(s) = \int_0^{\infty} L^m(\tau) e^{-s\tau} d\tau$. A forma matricial da Eq. (5.2.15) torna-se

$$\bar{M}_{N_{\mu}}^m(s) \bar{I}^m(s) = L^m(0) + \bar{Q}^m(s) \quad (5.2.16)$$

onde a matriz $\bar{M}_r^m(s)$ de ordem N_{μ} , chamada de matriz LTS_N, é dada por

$$\bar{M}_r^m(s) = s\mathbf{I} + A^m \quad (5.2.17)$$

sendo \mathbf{I} a matriz identidade de ordem N_μ , enquanto a matriz A^m é dada por

$$A^m(i, j) = \begin{cases} \frac{1}{\mu_j} - \frac{\varpi_0}{2\mu_j} \sum_{l=m}^{N_\phi} \beta_l^m P_l^m(\mu_j) \eta_j P_l^m(\mu_j), & \text{se } i = j, \\ -\frac{\varpi_0}{2\mu_j} \sum_{l=m}^{N_\phi} \beta_l^m P_l^m(\mu_j) \eta_i P_l^m(\mu_i), & \text{se } i \neq j \end{cases} \quad (5.2.18)$$

os vetores $\bar{I}^m(s)$,

$$\bar{I}^m(s) = [\bar{I}_1^m(s) \quad \bar{I}_2^m(s) \quad \dots \quad \bar{I}_{N_\mu}^m(s)],$$

$$\bar{Q}^m(s) = \begin{bmatrix} \bar{S}_1^m(s) & \bar{S}_2^m(s) & \dots & \bar{S}_{N_\mu}^m(s) \\ \mu_1 & \mu_2 & \dots & \mu_{N_\mu} \end{bmatrix}.$$

Para resolver a equação matricial (5.2.16), ela deve ser multiplicada pela matriz inversa $\bar{M}_{N_\mu}^m(s)$, como segue

$$\bar{I}^m(s) = [\bar{M}^m(s)]^{-1} I^m(0) + [\bar{M}^m(s)]^{-1} \bar{Q}^m(s) \quad (5.2.19a)$$

$$\bar{I}^m(s) = \bar{B}^m(s) I^m(0) + \bar{B}^m(s) \bar{Q}^m(s). \quad (5.2.19b)$$

Aplicando a transformada inversa de Laplace

$$I^m(\tau) = B^m(\tau) I^m(0) + H^m(\tau) \quad (5.2.20)$$

onde

$$B^m(\tau) = \mathcal{L}^{-1} [\bar{B}^m(s)], \quad H^m(\tau) = B^m(\tau) * Q^m(\tau) \quad (5.2.21)$$

no qual a convolução é denotada por “*”.

Inversão matricial é uma operação computacionalmente custosa. A diagonalização tira vantagem do fato de que a matriz LTS_N matrix, Eq. (5.2.17), é não degenerada, isto é todos os auto-valores são distintos e conseqüentemente pode ser diagonalizada

$$A^m = X^m D^m (X^m)^{-1} \quad (5.2.22)$$

onde D^m é uma matriz diagonal contendo os auto-valores de A^m e X^m é a correspondente matriz de auto-vetores.

$$\begin{aligned} B^m(\tau) &= \mathcal{L}^{-1} [(s\mathbf{I} + A^m)^{-1}] \\ &= X^m \mathcal{L}^{-1} [(s\mathbf{I} + D^m)^{-1}] (X^m)^{-1} \\ &= X^m e^{D^m \tau} (X^m)^{-1} \end{aligned} \quad (5.2.23)$$

substituindo a Eq. (5.2.23) na Eq. (5.2.20), chega-se

$$I^m(\tau) = X^m e^{D^m \tau} (X^m)^{-1} I^m(0) + H^m(\tau) \quad (5.2.24)$$

onde d^+ e d^- são auto-valores positivos e negativos, respectivamente.

O método, tal como descrito pelas Eqs. (5.2.20) ou (5.2.24), não funciona devido ao *overflow* numérico: valor da função exponencial quando avaliada em domínios com grande espessura e/ou grandes valores de N_μ . Um mudança de variáveis evita este efeito indesejado. A Eq. (5.2.20) pode ser escrita como segue

$$I^m(\tau) = \mathcal{B}^m(\tau)\xi^m(0) + H^m(\tau) \quad (5.2.25)$$

onde

$$\mathcal{B}^m(\tau) \equiv X^m \begin{bmatrix} e^{d^+(\tau-\zeta)} & 0 \\ 0 & e^{d^-\tau} \end{bmatrix}^m, \quad (5.2.26)$$

$$\xi^m(0) \equiv \begin{bmatrix} e^{d^+\zeta} & 0 \\ 0 & 1 \end{bmatrix}^m (X^m)^{-1} I^m(0). \quad (5.2.27)$$

Equação (5.2.25) pode ser representada por matrizes bloco:

$$\begin{bmatrix} I^1(\tau) \\ I^2(\tau) \end{bmatrix}^m = \begin{bmatrix} \mathcal{B}_{11}(\tau) & \mathcal{B}_{12}(\tau) \\ \mathcal{B}_{21}(\tau) & \mathcal{B}_{22}(\tau) \end{bmatrix}^m \begin{bmatrix} \xi^1(0) \\ \xi^2(0) \end{bmatrix}^m + \begin{bmatrix} H^1(\tau) \\ H^2(\tau) \end{bmatrix}^m \quad (5.2.28)$$

onde os índices 1 e 2 indexam as radiancias transmitidas nas direções para baixo e para cima, respectivamente. Na superfície da água, $N_\mu/2$ valores de radiancias incidentes são conhecidas. Deste modo, aplicando as condições de contorno (por conveniência, o índice do modo azimutal m foi omitido):

$$\mathcal{B}_{21}(\zeta)\xi^1(0) = I^2(\zeta) - \mathcal{B}_{22}(\zeta)\xi^2(0) - H^2(\zeta) \quad (5.2.29)$$

permitindo calcular os valores desconhecidos de $N_\mu/2$ componentes do vetor $\xi(0)$, para cada $m = 0, 1, 2, \dots, N_\phi$.

5.2.2 Identificação Multi-espectral de Propriedades Óticas

O grupo do LAC-INPE iniciou o estudo de problemas inversos em ótica hidrológica, com uma metodologia que fazia uso de dados experimentais *in situ*, ou seja, medidas de radiancia eram tomadas em várias profundidades do corpo d'água. Dados deste tipo tem relevância, porém este tipo de observação experimental tem aplicação local. É difícil inferir um comportamento do estado de uma propriedade física em posição distante do local da medida.

Desta forma é preciso tecnologia espacial (satélites) para se fazer observações em escala planetária. Porém, são poucos os métodos de inversão apresentando estimação de propriedades óticas a partir de imagens de satélites. Frequentemente o problema é simplificado, por exemplo fazendo a hipótese de oceano homogêneo ou a lei de empregando a lei de Gershun [91] na estimação de propriedades óticas aparentes.

A recuperação de propriedades óticas de sensores multi-espectrais (como aqueles usados em satélites) é um tema em aberto. O primeiro resultado do grupo de problemas inversos do LABAC-INPE em estimação multi-espectral em ótica hidrológica foi a estimativa de fontes de bioluminescência [24], onde o termo de fonte foi parametrizado por uma forma funcional parabólica e o método de Levenberg-Marquardt foi usado para identificar os coeficientes da parábola. Neste problema, modelos bio-óticos [91] foram empregados para representar os coeficientes de absorção e espalhamento. Além dessas considerações, nenhuma outra hipótese simplificadora do problema foi adotada.

Uma hipótese chave é assumir que as propriedades possam ser definidas em termos de separação de variáveis, isto é, a dependência funcional destas propriedades no domínio

espacial e com o comprimento de onda, possam ser escritas como um produto de 2 funções: uma com dependência espacial e outra com dependência espectral, somente.

O passo seguinte foi tentar identificar propriedades óticas inerentes, com o uso da hipótese da separação de variáveis. De fato, nos modelos bio-óticos, os coeficientes de absorção e espalhamento são dados por [91]:

$$a(z, \lambda) = [a_g^w + 0,06 a_g^c C^{0,65}(z)] \times [1 + 0,2 e^{-0,014(\lambda_g - 440)}] \quad (5.2.30)$$

$$b(z, \lambda) = \left(\frac{550}{\lambda_g}\right) 0,30 C^{0,62}(z), \quad (5.2.31)$$

onde a_g^w é o coeficiente de absorção da água pura e $C(z)$ é a concentração de clorofila.

Assim, o problema direto a ser resolvido é a versão multi-espectral da Eq. (5.2.10). O problema inverso é a reconstrução do perfil de concentração de clorofila. A ETR é resolvida discretizando a variável espacial (profundidade) em N_r regiões (ver Fig. 5.4) e função de fase de espalhamento de Henyey-Greenstein [91]:

$$\begin{aligned} \beta_{\text{HG}}(\cos \psi) &= \frac{1}{4\pi} (1 - g^2) [1 + g^2 - 2g \cos \psi]^{-3/2} \\ &\approx \sum_{l=0}^{N_p} \frac{2l+1}{4\pi} g^l P_l(\cos \psi) \end{aligned} \quad (5.2.32)$$

onde os coeficientes da expansão g^l são obtidos da relação de ortogonalidade do polinômio de Legendre $P_l(\cos \phi)$ [15, 117] e ψ é o ângulo de espalhamento entre as direções (θ', φ') e (θ, φ) :

$$\cos \psi = \mu \mu' + (1 - \mu^2)(1 - \mu'^2) \cos(\varphi' - \varphi) \quad (5.2.33)$$

por último, g é um fator de assimetria: quando $g \rightarrow 1$ o feixe de fótons é espalhado da direção de propagação do feixe, $g = 0,9$ foi utilizado.

O problema inverso é bem desafiador. Uma técnica implícita foi aplicada para a inversão. O método de colônia de formigas (ACO) com regularização intrínseca foi empregada para resolver o problema de otimização. A tabela 5.2 mostra os parâmetros usados na meta-heurística ACO. Porém, como o método ACO é uma técnica estocástica, o sistema foi executado diversas vezes com várias sementes, produzindo algumas soluções inversas pobres [143], o resultado final é uma média dessas execuções. Desta forma foram propostas outras 2 estratégias adicionais para melhorar a solução inversa [143]:

1. Muitos dos perfis gerados no processo de inversão apresentavam soluções com uma curvatura que não é observada (ver Fig. 5.5a), assim selecionou-se automaticamente somente os perfis gerados pela inversão com ACO, com a curvatura adequada, usando um teste de derivada segunda citeSouto:08 (ver Fig. 5.5b);
2. O critério adicional da derivada segunda melhorou os resultados, porém foi possível melhorar ainda mais o resultado *congelando* os valores da inversão dos perfis de concentração de clorofila da superfície até o valor máximo de clorofila e submetendo o problema para nova inversão – ou seja, o processo completo de inversão é com 2 passos, o resultado é mostrado na Fig. 5.5c.

A clorofila oceânica faz parte da cadeia trófica dos oceanos, é componente importante para fito-plancton, que por sua vez é a base da cadeia alimentar de pequenos moluscos, pequenos peixes, até a escala dos grandes predadores. Ou seja, a concentração de clorofila nos oceanos é uma medida da saúde do planeta. Desta forma, o monitoramento da concentração de clorofila nos oceanos é um importante parâmetro ambiental no planeta Terra. Este

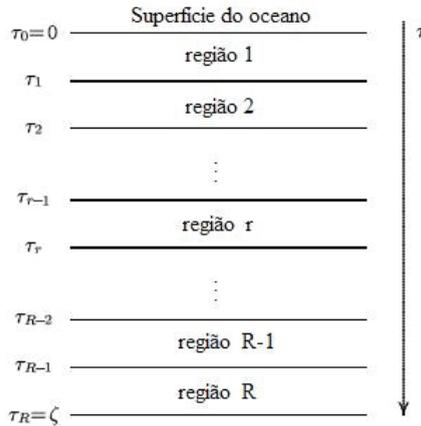


Figura 5.4: Discretização do domínio vertical do oceano: aqui $\zeta = \tau_{\max}$.

sementes	ns	np	na	mit	ρ	q_0
{ 17, 33, 55, 81, 99 }	10	3000	90	500	0,03	0

Tabela 5.2: Parâmetros usados na meta-heurística ACO.

monitoramento para se tornar eficiente em escala planetária necessita de um imageamento por satélite.

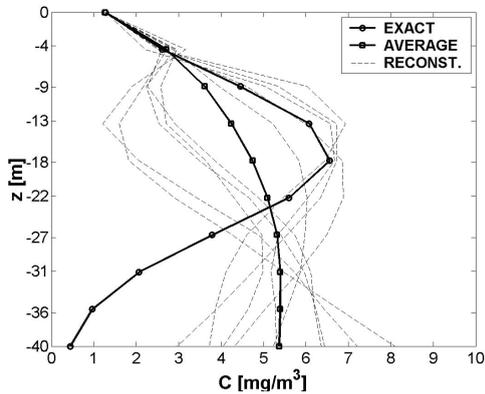
Em julho de 2008, divulgamos um procedimento onde se poderia processar uma tomografia 3D da concentração de clorofila oceânica a partir de dados de satélites [142]. Esta inversão requer cerca de 4 horas de processamento *por pixel* da imagem, onde foram feitas várias análises para diminuir o esforço computacional. Tal procedimento consistiu numa avaliação da redução do grau de anisotropia ($N_\mu = N_\beta = 173$), uso de processamento paralelo (8 processadores) e aplicação de *regularização intrínseca* para o ACO – que reduz 80% do custo computacional efetivo. Sem a melhoria do desempenho numérico usando regularização da forma padrão e aplicação de ACO com um único processador, a inversão é realizada em *aproximadamente 39 dias*.

Nos exemplos executados, os perfis de concentração de clorofila foram representados por curvas gaussianas:

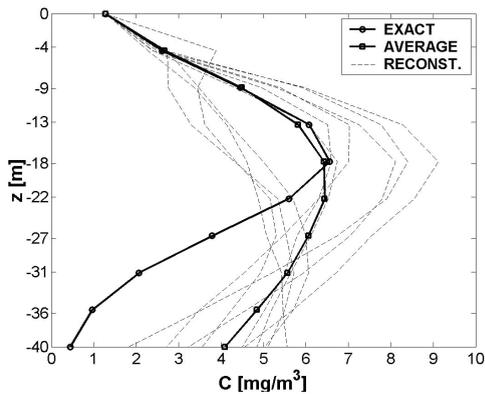
$$C(z) = C_{bg} + \frac{h}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z - z_{max}}{\sigma}\right)^2\right] \quad (5.2.34)$$

onde z é a profundidade em metros e $C(z)$ é dado em mg/m^3 . A tabela 5.3 mostra valores numéricos para a obtenção de 3 perfis (o perfil-1 é a concentração verdadeira mostrada na Fig. 5.5).

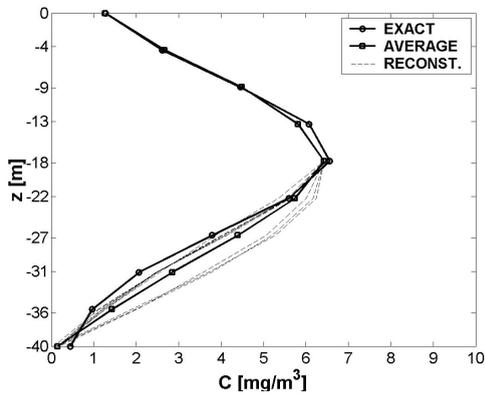
Uma reconstrução tomográfica 3D foi desenvolvida usando a técnica de computação em grade [48], sob a plataforma OurGrid [105]. A grade se constituiu de 3 sistemas computacionais geograficamente distribuídos: no Departamento de Computação da UFSM (SGI



(a)



(b)



(c)

Figura 5.5: Perfis de concentração de clorofila: (a) ACO e regularização intrínseca (RI), (b) ACO-RI com critério da 2a. derivada, (c) ACO-RI com 2 passos.

Perfil	C_{bg}	h	σ	z_{\max}
1	0,2	144	9	17
2	0,2	144	9	25
3	0,2	144	12	17

Tabela 5.3: Parâmetros dos perfis de clorofila para a Eq. (5.2.34).

λ [nm]	410	520	580	600	640	680	700
a_{λ}^w	0,017	0,048	0,108	0,245	0,330	0,450	0,650
a_{λ}^c	0,828	0,528	0,291	0,236	0,334	0,502	0,215

Tabela 5.4: Coeficientes a_{λ}^w e a_{λ}^c da Eq. (5.2.30), de acordo com [91].

P1	P1	P1	P1	P1	P1
P1	P3	P3	P3	P3	P1
P1	P3	P2	P2	P3	P1
P1	P3	P2	P2	P3	P1
P1	P3	P3	P3	P3	P1
P1	P1	P1	P1	P1	P1

Figura 5.6: Representação de pixels de uma imagem com resolução de $10 \text{ Km} \times 10 \text{ Km}$ por pixel, onde P1, P2 e P3 são os perfis construídos com os valores da Tabela 5.3.

Altix XE com 8 cores 2 GHz), Instituto de Informática da UFRGS e LABAC-INPE (Cray XD1 com 4 e 8 processadores AMD Opteron 2.8 GHz, respectivamente). Este problema é um bom exemplo onde a computação em grade pode ser eficiente, pois a inversão a partir de radiâncias (sintéticas, no exemplo) medidas em N_λ canais de satélite é realizada de maneira independente para cada pixel da imagem 5.6.

A figura 5.7 é o resultado da tomografia 3D para a concentração de clorofila para águas do tipo-1 (longe da costa). Bons resultados para inversão são obtidos até a profundidade de 22 m aproximadamente, para um corpo d'água de 40 m de profundidade. Esta reconstrução tomográfica representa uma busca de quase 15 anos de pesquisa em ótica hidrológica inversa.

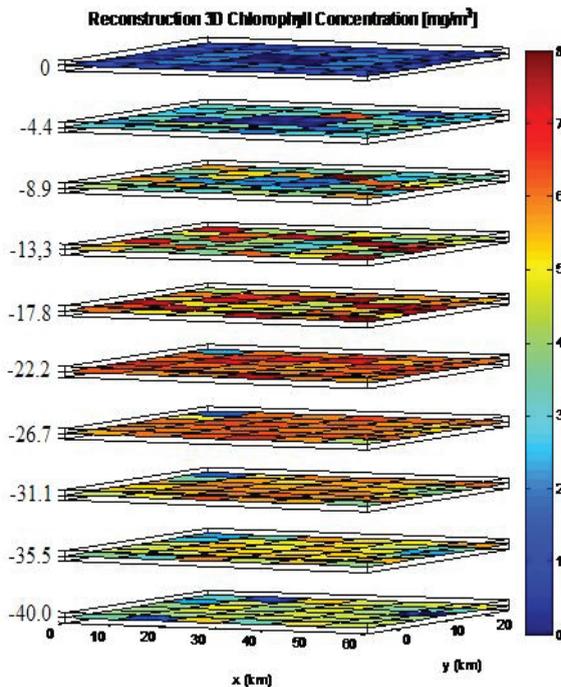


Figura 5.7: Reconstrução 3D da concentração de clorofila em águas oceânicas longe da costa.

Capítulo 6

Considerações Finais

A área espacial é uma rica fonte para aplicações da metodologia de problemas inversos. Um problema típico é a estimativa do fluxo de calor ao qual um veículo espacial estará sujeito na reentrada da atmosfera. Neste texto, foi realizada uma introdução de alguns métodos matemáticos para a solução de problemas inversos, e alguns problemas da área espacial foram selecionados para ilustrar como essas técnicas poderosas são aplicadas no cálculo de soluções inversas.

Dois pontos importantes não foram tratados neste texto. Na parte de metodologia, não foi apresentada a técnica de decomposição truncada em valores singulares (TSVD: *Truncated Singular Value Decomposition*) e nem a técnica estatística de estimação Bayesiana (a referência mais conhecida nesta abordagem é o texto do prof. Tarantola [149]). Uma aplicação importante refere-se a restauração de imagens, na área espacial este é um tema muito importante, tanto em sensoriamento remoto (imagens de satélite), como em imagens astronômicas. Porém, há uma extensa literatura sobre estes temas [2, 3, 7, 108, 135, 154].

É pertinente ressaltar que muito do que o grupo de problemas inversos do LABAC-INPE vem desenvolvendo não foi visto aqui. Como destaque de desenvolvimentos significativos pode-se citar o novo algoritmo para otimização extrema [145]: GEO (*Generalized Extreme Optimization*) e sua aplicação em problemas multi-objetivos (M-GEO) e o emprego em sistemas espaciais: o projeto térmico de radiadores da Plataforma Multimissão (PMM) do INPE e a obtenção da configuração de uma constelação de satélites de sensoriamento remoto [50]. O operador de epidemia para GA, foi proposto na tese de mestrado de Medeiros [88] e seu emprego em problemas inversos foi realizado por Chiwiacowsky e Campos Velho [27]. A versão paralela do EGA foi o tema da tese de mestrado de Sambatti [124] – ver também [18, 125].

Outro desenvolvimento que merece destaque é a aplicação do método de colônia de formigas para otimização (ACO). Na aplicação do ACO em problemas inversos, a idéia de regularização intrínseca foi concebida, onde somente caminhos mais suaves são considerados - a suavidade do caminho é avaliada por uma norma de Tikhonov [139]. Esta estratégia ainda faz uso de regularização, mas não faz uso do parâmetro de regularização. Claramente, a estratégia pode ser adotada em outras meta-heurísticas.

Aplicações importantes são novos algoritmos para produção de mapas de radiação cósmica de fundo, onde foi usada uma versão do SA com ciclos térmicos [37], e o projeto de arranjo de antenas para o projeto BDA (*Brazilian Decimetric Array*) do INPE, onde a estratégia do ACO foi usada para maximizar uma função objetivo da entropia da imagem do telescópio BDA [46]. Um problema relevante nos dias atuais é a determinação de fluxos superficiais do ciclo bio-geo-químicos de gases de efeito estufa. Roberti [120] desenvolveu uma metodologia para atacar este problema, maximizando entropia de 2a. or-

dem, utilizando o método quase-Newton para determinar a solução ótima. Neste contexto, usou-se a estratégia de identificar os valores funcionais de uma função amostrada. O método de Levenberg-Marquardt, que é muito eficiente em problemas inversos de estimação de parâmetros, não conseguiu bons resultados neste caso. Isso indica que o problema de estimação de funções (mesmo no caso de função amostrada) tem uma natureza matemática distinta da estimação de parâmetros. O mesmo problema foi resolvido usando uma técnica de otimização estocástica: enxame de partículas, melhorando o resultado obtido com o método quase-Newton [83]. Os resultados foram ainda melhores com a aplicação de redes neurais [107].

Do ponto de vista teórico, Campos Velho et al. [17] usaram o princípio da máxima entropia não-extensiva para unificar a regularização de Tikhonov e a regularização entrópica.

A maior parte dos resultados obtidos e grande esforço de desenvolvimento, estão ligados a concepção tradicional de resolução de problemas inversos: métodos de solução de problemas de otimização para obter soluções regularizadas. Não há nada de errado em seguir por este caminho, pelo contrário, os resultados mostram que esta é uma linha de pesquisa com muito a oferecer. Por outro lado, a aplicação de redes neurais é uma metodologia realmente nova neste campo de pesquisa, permitindo atacar problemas em que sequer o problema direto esteja matematicamente formulado. A aplicação da técnica de redes neurais em problemas inversos é um campo de pesquisa muito promissor. Nosso estudo iniciou-se com a identificação de condições iniciais e de contorno em problemas simples de condução do calor. Como comentado na Seção 3.4, as redes neurais foram aplicadas com sucesso na recuperação de perfis verticais de temperatura e umidade atmosféricas a partir de dados de satélite meteorológicos com sensores na banda espectral do infra-vermelho. Além da rede perceptron de múltiplas camadas, outras redes podem ser utilizadas com resultados igualmente excelentes [132]. Inversões com redes neurais podem ser aplicadas com outros tipos de sinais de satélites, como sinais de GPS [34]. Porém, talvez o aspecto mais relevante nesta nova técnica seja sua potencialidade de implementação em dispositivos de hardware. Neurocomputadores podem representar um salto fantástico no ganho de desempenho de processamento.

Apêndice A

Conceitos e Propriedades Matemáticas

A.1 Conceitos Preliminares

Espaço com produto interno, espaço normado e espaço métrico.

Definicao A.1. *Espaço com produto interno: é um espaço vetorial, denotado por exemplo \mathcal{H} , onde está definida uma função complexa bilinear $\langle \cdot, \cdot \rangle$, que satisfaz as seguintes propriedades:*

- (i) $\langle f, g \rangle = \overline{\langle g, f \rangle}$, onde $f, g \in \mathcal{H}$;
- (ii) $\langle f, \alpha_1 g_1 + \alpha_2 g_2 \rangle = \alpha_1 \langle f, g_1 \rangle + \alpha_2 \langle f, g_2 \rangle$, onde α_1 e α_2 são números complexos;
- (iii) $\langle f, f \rangle \neq 0$, se $f \neq 0$; e $\langle f, f \rangle = 0 \Leftrightarrow f = 0$.

Definicao A.2. *Espaço normado: uma norma sobre um espaço vetorial \mathcal{B} é um funcional linear $\|f\|$ com as seguintes propriedades:*

- (i) $\|f\| > 0$, para todo $f \in \mathcal{B}$ e $\|f\| = 0 \Leftrightarrow f = 0$;
- (ii) $\|\alpha f\| = |\alpha| \|f\|$, para todo α escalar (número complexo);
- (iii) $\|f + g\| \leq \|f\| + \|g\|$, para $f, g \in \mathcal{B}$.

Um espaço linear \mathcal{B} equipado com uma norma $\|f\|$ é chamado um *espaço linear normado*.

Todo o produto interno, de um espaço de produto interno, induz uma norma dada por:

$$\|f\| = \sqrt{\langle f, f \rangle}.$$

Espaços com produto interno e espaços normados podem conter uma *função de distância*, dada pela norma da diferença:

$$\|f - g\| \equiv d(f, g)$$

esta norma da diferença é chamada de distância entre as funções f e g e pela propriedade (iii) da norma, tem-se que:

$$\|f - h\| \leq \|f - g\| + \|g - h\| \Leftrightarrow d(f, h) \leq d(f, g) + d(g, h).$$

Definicao A.3. Espaço métrico: dado o par (\mathcal{X}, d) , para um conjunto \mathcal{X} e uma função de distância d , de valor real $d(f, g)$ definida para $f, g \in \mathcal{X}$ com as seguintes propriedades:

- (i) $d(f, g) \geq 0$ para todo $f, g \in \mathcal{X}$ e: $d(f, g) = 0 \Leftrightarrow f = g$;
- (ii) $d(f, g) = d(g, f)$;
- (iii) $d(f, h) \leq d(f, g) + d(g, h)$.

Todo espaço linear normado são espaços métricos, pois: $d(f, g) = \|f - g\|$; entretanto nem todo o espaço métrico é um espaço normado.

As definições acima, vão servir para tornar preciso outros conceitos. Por exemplo, diz-se que a seqüência de elementos f_1, f_2, \dots de um espaço métrico converge para f se

$$\lim_{n \rightarrow \infty} f_n = f.$$

e o elemento f é chamado de *limite* da seqüência $\{f_n\}$. Mais precisamente, diz-se que a distância

$$d(f_n, f) = \|f_n - f\|$$

pode ser arbitrariamente pequena quando $n \rightarrow \infty$.

Supondo que uma seqüência $\{f_n\}$ seja uma seqüência convergente, isto é, $\lim_{n \rightarrow \infty} f_n = f$, desta forma considerando $f_1, f_2, \dots, f_m, \dots, f_n, \dots$ pode-se escrever:

$$\|f_n - f_m\| = \|(f_n - f) + (f - f_m)\| \leq \|f_n - f\| + \|f - f_m\| \rightarrow 0 \text{ quando } m, n \rightarrow \infty$$

ou, para um espaço métrico:

$$d(f_n, f_m) \leq d(f_n, f) + d(f, f_m) \rightarrow 0 \text{ quando } m, n \rightarrow \infty.$$

Desta forma, quando:

$$\lim_{n, m \rightarrow \infty} \|f_n - f_m\| = 0 \text{ num espaço vetorial} \quad (\text{A.1.1})$$

ou

$$\lim_{n, m \rightarrow \infty} d(f_n, f_m) = 0 \text{ num espaço métrico} \quad (\text{A.1.2})$$

é uma seqüência $\{f_n\}$ cujos os elementos satisfazem esta condição limite e é chamada de *seqüência de Cauchy*. Deste modo, toda a seqüência convergente é uma seqüência de Cauchy.

O prof. Gerlach [51] mostra um exemplo simples e interessante para mostrar que um espaço de funções contínuas $C[a, b] = \{f(x) : f \text{ continua em } x \in [a, b]\}$, não é completo, ou seja, o limite de uma seqüência de Cauchy não é um elemento deste espaço. O exemplo é o seguinte: considere a seqüência:

$$u_k(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(kx) \quad x \in [-1, 1].$$

Considerando o produto interno e a norma: $\langle f, f \rangle = \|f\|^2 = \int_{-1}^1 \bar{f}f \, dx$, o espaço $C[-1, 1]$ é *incompleto* em relação a esta norma- L_2 . Pois:

1. $\lim_{p, q \rightarrow \infty} \|u_p - u_q\|^2 = \lim_{p, q \rightarrow \infty} \int_{-1}^1 (u_p - u_q)^2 \, dx = 0$. Ou seja, a seqüência $\{u_k : k = 1, 2, \dots\}$ é uma seqüência de Cauchy.

2. Para x fixo:

$$\lim_{k \rightarrow \infty} u_k(x) = v(x) = \begin{cases} 1 & 0 < x \leq 1 \\ 1/2 & x = 0 \\ 0 & -1 \leq x < 0 \end{cases}$$

que é uma função descontínua: $v(x) \notin C[-1, 1]$.

$$3. \lim_{k \rightarrow \infty} \|v - u_k\|^2 = \lim_{k \rightarrow \infty} \int_{-1}^1 [v(x) - u_k(x)]^2 dx = 0.$$

4. Pode ser mostrado que não existe qualquer função contínua $w(x)$ tal que: $\|w - u_k\| \rightarrow 0$ quando $k \rightarrow \infty$.

Este exemplo é para ilustrar que o espaço de funções contínuas $C[a, b]$ é *Cauchy incompleto*, no sentido que o limite de uma seqüência de Cauchy é um elemento que pode não pertencer ao espaço.

Espaços completos com produto interno são espaços de *Hilbert*. Espaços normados lineares completos são espaços de *Banach*.

A.2 A Alternativa de Fredholm

No capítulo 1, vimos que a equação do calor, pode ser formulada como uma equação integral — Eq. (1.2.8) (ver também referência [53]). Portanto, prova de existência e unidade de solução pode ser tentada via estudo de equações integrais. Um resultado importante é o teorema abaixo.

Teorema A.1. Alternativa de Fredholm: *Para a equação integral de Fredholm homogênea*

$$\lambda f(x) - \int_a^b K(x, y)f(y)dy = 0$$

ou a equação integral de Fredholm heterogênea

$$\lambda f(x) - \int_a^b K(x, y)f(y)dy = g(x)$$

a alternativa de Fredholm declara que, para qualquer $\lambda \in \mathbb{C}$ não nulo, ou a equação homogênea tem uma solução não trivial, ou a segunda equação tem solução para todo $g(x)$.

Uma condição suficiente para o teorema é que o núcleo $K(x, y)$ seja de quadrado integrável sobre $[a, b] \times [a, b]$. O resultado para o operador de Fredholm generaliza o resultado para espaços vetoriais de dimensão infinita (espaço de Banach). O operador de Fredholm $T : V \rightarrow V$ é expresso como:

$$T(x, y) = \lambda \delta(x, y) - K(x, y)$$

deste modo,

$$g(x) = \int_a^b T(x, y)f(y)dy.$$

Em termos mais precisos, a alternativa de Fredholm se aplica somente quando K é um operador compacto. Da teoria de Fredholm, núcleos integrais suaves são operadores compactos.

Prova: Consulte [45], página 641.

A.3 Operadores Compactos

Sejam X e Y espaços de Banach.

Definicao A.4. *Operador Compacto: Um operador linear limitado*

$$K : X \rightarrow Y$$

é chamado **compacto** se para cada seqüência limitada $\{u_k\}_{k=1}^{\infty} \subset X$, a seqüência $\{Ku_k\}_{k=1}^{\infty}$ é precompacta em Y , isto é, existe uma subseqüência $\{u_{k_j}\}_{j=1}^{\infty}$ tal que $\{Ku_{k_j}\}_{j=1}^{\infty}$ converge em Y .

Não é difícil perceber que para um espaço de Hilbert real, com produto interno $\langle \cdot, \cdot \rangle$, o operador linear $K : H \rightarrow H$ é compacto e que se $u_k \rightarrow u$, então $Ku_k \rightarrow Ku$.

A.4 Espectro de Operadores Compactos

Se T é um operador linear contínuo de um espaço de Hilbert: $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$, o operador adjunto $T^* : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ é o operador linear contínuo definido por:

$$\langle Tx, y \rangle = \langle x, T^*y \rangle$$

para $x \in \mathcal{H}_1$ e $y \in \mathcal{H}_2$. O posto $R(T)$ e o espaço nulo $N(T)$ de um operador linear com domínio $D(T)$ são definidos por:

$$\begin{aligned} R(T) &= \{Tx : x \in D(T)\} \\ N(T) &= \{x \in D(T) : Tx = 0\}. \end{aligned}$$

Há uma propriedade importante entre o complemento ortogonal destes espaços.

Teorema A.2. *Se $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ é um operador linear contínuo, então $R(T)^\perp = N(T^*)$ e $N(T)^\perp = R(T^*)$.*

O operador T é limitado se:

$$\|T\| = \sup_{\|x\|} \|Tx\| < \infty.$$

O espectro de um operador linear $T : \mathcal{H} \rightarrow \mathcal{H}$ é o conjunto de números complexos $\sigma(T)$ definido por

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ não tem inversa limitada}\}$$

onde I é o operador identidade sobre \mathcal{H} .

Um operador $T : \mathcal{H} \times \mathcal{H}$ é chamado auto-adjunto se $T = T^*$.

Teorema A.3. *Auto-vetores de um operador compacto e simétrico: Se \mathcal{H} é um espaço de Hilbert separável e o operador $T : \mathcal{H} \rightarrow \mathcal{H}$ é compacto e simétrico, então existe uma base ortonormal de \mathcal{H} , consistindo de auto-vetores de T .*

Prova: Consulte [45] página 645, ou [75] página 237.

O teorema espectral para dimensão finita é apresentado em [135]. No caso de dimensão infinita, a equação dos auto-valores é:

$$T\phi_n = \lambda_n\phi_n$$

e fazendo a ordenação dos auto-valores

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq \dots \geq 0$$

com

$$\lambda_n \neq 0 \Rightarrow \dim\{\ker(T - \lambda_n I)\} < \infty.$$

Se $\dim(\mathcal{H}) = \infty$, então: $\lim_{n \rightarrow \infty} \lambda_n = 0$.

Os operadores compactos auto-adjunto tem um espectro simples: cada ponto não nulo do espectro é um ponto isolado que é um auto-valor do operador.

Apêndice B

Termo-estatística Não Extensiva: Propriedades

B.1 Definições

A1: Entropia não extensiva:

$$S_q(p) \equiv \frac{k}{q-1} \left(1 - \sum_{i=1}^{N_p} p_i^q \right). \quad (\text{B.1.1})$$

A2: q -expectativa de um observável:

$$O_q \equiv \langle O \rangle_q = \sum_{i=1}^{N_p} p_i^q O_i. \quad (\text{B.1.2})$$

B.2 Propriedades

1. Se $q \rightarrow 1$:

$$S_1 = k \sum_{i=1}^{N_p} p_i \ln p_i, \quad (\text{B.2.3})$$

$$O_1 = \sum_{i=1}^{N_p} p_i O_i. \quad (\text{B.2.4})$$

2. A entropia não-extensiva é positiva: $S_q \geq 0$.

3. Não extensividade:

$$S_q(A+B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B) \quad (\text{B.2.5})$$

$$\begin{aligned} O_q(A+B) &= O_q(A) + O_q(B) \\ &+ (1-q)[O_q(A)S_q(B) + O_q(B)S_q(A)]. \end{aligned} \quad (\text{B.2.6})$$

4. Máximo de S_q sob a restrição $O_q = \sum_i p_i^q \epsilon_i$ (ensemble canônico):

$$p_i = \frac{1}{Z_q} [1 - \beta(1 - q)\epsilon_i]^{1/(1-q)} \quad (\text{B.2.7})$$

onde ϵ_i é a energia de estado- i , $O_q = U_q$ é a forma não-extensiva para a energia interna e o fator de normalização Z_q (função de partição), para $1 < q < 3$, é dada por

$$Z_q = \left[\frac{\pi}{\beta(1-q)} \right]^{1/2} \frac{\Gamma[(3-q)/2(q-1)]}{\Gamma[1/(q-1)]} . \quad (\text{B.2.8})$$

Para $q = 1$ tem-se

$$p_i = e^{\beta\epsilon_i} / Z_1 . \quad (\text{B.2.9})$$

Bibliografia

- [1] O.M. Alifanov, Solution of an Inverse Problem of Heat Conduction by Iteration Methods, *J. Eng. Phys.*, **26** (1974), 471–476.
- [2] C.A.Z. Barcelos, M. Boaventura, E.C. Silva Jr, Edge Detection and Noise Removal by use of a Partial Differential Equation with Automatic Selection of Parameters, *Comput. Appl. Math.*, **24** (2005), 131–150.
- [3] C.A.Z. Barcelos, Edge-Preserving Regularization in Image Restoration, *Boletim da Sociedade Brasileira de Matemática Aplicada e Computacional*, vol. **VII**, 31–43, 2006.
- [4] L.B. Barichello, M.T. Vilhena, A General Approach to One Group One Dimensional Transport Equation, *Kernteknik*, **58** (1993), 182–184.
- [5] J.V. Beck, B. Blackwell, C.R. Clair Jr., “Inverse Heat Conduction: Ill-Posed Problems”, John Wiley & Sons, 1985.
- [6] M.N. Berdichevsky, V.I. Dmitriev, “Magnetotellurics in the Context of the Theory of Ill-posed Problems”, Society for Exploration Geophysics, in *Investigations in Geophysics No. 11*, 2002.
- [7] M. Bertero and P. Boccacci, “Introduction to Inverse Problems in Imaging”, Institute of Physics, 1999.
- [8] M.N. Berdichevsky, V.I. Dmitriev, “Magnetotellurics in the Context of the Theory of Ill-posed Problems”, Society for Exploration Geophysics, in *Investigations in Geophysics No. 11*, 2002.
- [9] J. de Boissieu, R. J. Papoular, C. Janot, Maximum entropy method as applied in quasi-crystallography, *Europhysics Letters*, **16** (1991), 343-347.
- [10] E. Bonabeau, M. Dorigo, G. Theraulaz, “Swarm Intelligence: from Natural to Artificial Systems”, Oxford University Press, UK, 1999.
- [11] H. D. Bui, “Inverse Problems in the Mechanics of Materials: An Introduction”, CRC Press, Boca Raton (FL, USA), 1994.
- [12] H. F. Campos Velho, F. M. Ramos, Numerical Inversion of Two-Dimensional Geoelectric Conductivity Distributions from Eletromagnetic Ground Data, *Braz. J. Geophys.*, **15** (1997), 133-143.
- [13] H. F. Campos Velho, M. R. Moraes, F. M. Ramos, G. A. Degrazia, D. Anfossi, An Automatic Methodology for Estimating Eddy Diffusivity from Experimental Data, *Nuovo Cimento C*, **23** (2000), 65-84.

- [14] H. F. Campos Velho, Problemas Inversos: Conceitos Básicos e Aplicações, (Minicurso), Boletim da Sociedade Brasileira de Matemática Aplicada e Computacional, Série III (número especial: Anais do IV Encontro de Modelagem Computacional, IPRJ-UERJ, Nova Friburgo (RJ), novembro-2001).
- [15] H. F. Campos Velho, M. R. Retamoso, M. T. Vilhena, Inverse Problems for Estimating Bottom Boundary Conditions of Natural Waters, *International Journal for Numerical Methods in Engineering*, **54** (2002), 1357-1368.
- [16] H. F. Campos Velho, V. C. F. Barbosa, D. R. Roberti, R. P. Souto, E. H. Shiguemori, Inverse Analysis in Geoscience Problems, *XXVI Iberian Latin American Congress on Computational Methods in Engineering* (CILAMCE-2005), Guarapari (ES), Brasil, (2005), CIL1013-1—CIL1013-13.
- [17] H. F. Campos Velho, F. M. Ramos, E. H. Shiguemori, J. C. Carvalho, A Unified Regularization Theory: The Maximum Non-extensive Entropy Principle, *Comput. Appl. Math.* (CAM-SBMAC), **25** (2006), 307-330.
- [18] H. F. Campos Velho, S. B. M. Sambatti, L. D. Chiwiacowsky, Combining a Parallel Genetic Algorithm with Variational Approach for Assessing Structural Damage, *III European Conference on Computational Mechanics* (ECCM-2006), Mini-Symposium Inverse Engineering, Proc. in CD-Rom: paper code 1558 (09 pages), Abstract Book page 479, Lisbon, Portugal, 5-8 June, 2006.
- [19] H. F. Campos Velho, J. D. S. Silva, E. H. Shiguemori, Hardware implementation for the atmospheric temperature retrieval from satellite data, *Inverse Problems, Design and Optimization Symposium* (IPDO), 2007. vol. 1, pp. 349-353.
- [20] H. S. Carslaw, J. C. Jaeger, "Conduction of Heat in Solids", Oxford at the Clarendon Press, London, 1979.
- [21] J. C. Carvalho, "Inferência de Perfis Verticais de Temperatura Utilizando uma Técnica Iterativa Implícita de Inversão", Tese de Mestrado, Meteorologia, INPE-7228-TDI/682, 1999.
- [22] J. C. Carvalho, "Modelagem e Análise de Sondagens Remotas sobre o Brasil Utilizando-se o Sistema ICI", Tese de Doutorado, Meteorologia, INPE, 2002.
- [23] E. S. Chalhoub, R. D. M. Garcia, The Equivalence between Two Techniques of Angular Interpolation for the Discrete-ordinates Method, *Journal of Quantitative Spectroscopy & Radiative Transfer* (JQSRT), **64** (2000), 517-535.
- [24] E. S. Chalhoub, H. F. Campos Velho, Multispectral Reconstruction of Bioluminescence Term in Natural Waters, *Applied Numerical Mathematics*, **47** (2003), 365-376.
- [25] E. S. Chalhoub, H. F. Campos Velho, M. T. Vilhena, R. D. M. Garcia, A Comparison of Radiances Generated by Selected Methods of Solving the Radiative-transfer Equation, *Transport Theory and Statistical Physics*, **32** (2003), 469-499.
- [26] S. Chandrasekhar, "Radiative Transfer", Dover, New York, 1950.
- [27] L. D. Chiwiacowsky, H. F. Campos Velho, Different Approaches for the Solution of a Backward Heat Conduction Problem, *Inv. Prob. Sci. Eng.*, **11** (2003), 471-494.
- [28] L. D. Chiwiacowsky, H. F. Campos Velho, P. Gasbarri, The Damage Identification Problem: A Hybrid Approach, *2nd Congresso Temático sobre Dinâmica e Controle* (DINCON-2003), SÃO JosÃ© dos Campos/SP, Brazil (2003).

- [29] L. D. Chiwiacowsky, P. Gasbarri, H. F. Campos Velho, A Variational Approach for Solving an Inverse Vibration Problem, *Inverse Problems, Design and Optimization Symposium (IPDO)*, 17-19 March, Rio de Janeiro (RJ), Brazil, Proceedings in CD-Rom, paper code IPDO-078, **Vol. 2**, 290-298.
- [30] L. D. Chiwiacowsky, H. F. Campos Velho, P. Gasbarri, Determining Stiffness Matrix by the Adjoint Method, *International Conference on Computational and Experimental Engineering and Sciences (ICCES-2004)*, 26-29 July, Ilha da Madeira, Portugal - paper code ICCES04-356 (12 pages).
- [31] L. D. Chiwiacowsky, “Método Variacional e Algoritmo Genético em Identificação de Danos Estruturais”, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2005.
- [32] L. D. Chiwiacowsky, H. F. Campos Velho, P. Gasbarri, A Variational Approach for Solving an Inverse Vibration Problem, *Inv. Prob. Sci. Eng.*, **14** (2006), 557-577.
- [33] L. D. Chiwiacowsky, P. Gasbarri, H. F. Campos Velho, Damage Assessment of Large Space Structures Through the Variational Approach, *Acta Astronautica*, **62** (2008), 592-604.
- [34] R. S. Cintra, J. D. S. Silva, H. F. Campos Velho, Retrieval of Humidity Profiles with Radio Occultation Measurements Using an Artificial Neural Networks, *Inverse Problems, Design and Optimization Symposium (IPDO)*, April 16-18, Miami (FL), USA, **vol. 2**, 642-649, 2007.
- [35] S. C. Constable, R. L. Parker, G. C. Constable, Occam’s inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52** (1987), 289-300.
- [36] S. D. Conte, C. de Boor, “Elementary Numerical Analysis: An Algorithmic Approach”, McGraw-Hill, 1987.
- [37] R. V. Corrêa, “Um Estudo de Métodos Estocásticos em Ambiente Paralelo para a Produção de Mapas de Anisotropias da Radiação Cósmica de Fundo em Microondas”, Tese de Doutorado, Computação Aplicada, INPE, São José dos Campos (SP), 2005.
- [38] A. P. C. Cuco, A. J. Silva Neto, F. L. Souza, H. F. Campos Velho, Solution of an Inverse Adsorption Problem with an Epidemic Genetic Algorithm and the Generalized Extremal Optimization Algorithm *Inverse Problems, Design and Optimization Symposium (IPDO-2007)*, Miami (USA), **Vol. 2** (2007), 430–437 (aceito para publicação na revista: *Inverse Problems in Science & Engineering*, 2008).
- [39] R. Daley, “Atmospheric Data Analysis”, Cambridge University Press, 1993.
- [40] S. Doebling, C. Farrar, M. Prime, D. Shevitz, “Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in their Vibration Characteristics: a Literature Review”, Los Alamos National Laboratory (LA-13070-MS), USA, 1996.
- [41] M. Dorigo, V. Maniezzo, A. Colomi, The Ant Optimization: Optimization by a Colony of Cooperating Agents, *IEEE Trans. Syst. Man Cybern. Part B-Cybernetics*, **26** (1996), 29-41.

- [42] DBCLSF routine, IMSL MATH LIBRARY, Users Manual. Version 2.0, Houston, IMSL, 1991.
- [43] E04UCF routine, NAG Fortran Library Mark 17, Oxford, UK, 1995.
- [44] H. W. Engl, M. Hanke, A. Neubauer, "Regularization of Inverse Problems: Mathematics and its Applications", Kluwer, 1996.
- [45] L. C. Evans, "Partial Differential Equations", American Mathematical Society, Providence, 2000.
- [46] C. Faria, "Uma Nova Estratégia de Otimização de Arranjos Interferométricos *Aplicada ao Brazilian Decimetric Array*", Tese de Doutorado, Computação Aplicada, INPE-14171-TDI/1086, São José dos Campos (SP), 2006.
- [47] M. Fleisher, U. Mahlab, J. Shamir, Entropy optimized filter for pattern recognition, *Applied Optics*, **29** (1990), 2091-2098.
- [48] I. Foster, C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure", Morgan Kaufman Publishers, 1999.
- [49] P. Gasbarri, L. D. Chiwiacowsky, H. F. Campos Velho, A Comparison between Frequency and Time Domain Approaches for Determining a Structural Damage Using the Adjoint Method, *Aerotecnica Missili e Spazio*, **86** (2007) 53-69 (disponível na internet: <http://www.aidaa.it/2-2007/130.pdf>).
- [50] R. L. Galski, "Desenvolvimento de Versões Aprimoradas, Híbridas, Paralela e Multiobjetivo do Método da Otimização Extrema Generalizada e sua Aplicação no Projeto de Sistemas Espaciais", Tese de Doutorado, Computação Aplicada, INPE, São José dos Campos (SP), 2006.
- [51] U. Gerlach, "Linear Mathematics in Infinite Dimensions", Dept. of Mathematics, University of Ohio, (acesso em Maio/2008). <http://www.math.ohio-state.edu/gerlach/>
- [52] H. R. Gordon, Inverse methods in hydrologic optics, *Oceanology*, **44** (2002), 9-58.
- [53] C. W. Groetsche, "The Theory of Tikhonov Regularization for Fredholm Integral Equation of the 1st Kind", Pitman, London, 1984.
- [54] S.F. Gull, G.J. Daniell, Image reconstruction from incomplete and noisy data, *Nature*, **272** (1978), 686-690.
- [55] J. Hadarmard, Sur les problèmes aux dérivées partielles et leur signification physique, *Bull. Princeton Univ.*, **13** (1902), 49-52.
- [56] J. Hadarmard, "Lectures on Cauchy's Problem in Linear Partial Differential Equations", Dover, 1952.
- [57] R. J. Hanisch, R. L. White, The restoration of HST images and spectra II, Space Telescope Science Institute (<http://resources.stsci.edu/>), Baltimore (MA), USA, November 18-19, 1993.
- [58] P. C. Hansen, Analysis of Discrete Ill-posed Problems by Means of the L-curve, *SIAM Review*, **34** (1992), 561-580.

- [59] S. Haykin, "Neural Networks: A Comprehensive Foundation"; , Macmillan, New York, 1993.
- [60] J. D. Hoffman, "Numerical methods for engineers and scientists", McGraw-Hill, 1993.
- [61] J. H. Holland, "Adaptation in Natural and Artificial Systems", MIT Press, 1975.
- [62] S. Hawking, "Os Gênios da Ciência: Sobre os Ombros de Gigantes", Elsevier/Campus, São Paulo, 2005.
- [63] C. S. Hönl, "Análise Funcional e o Problema de Sturm-Liouville", EDUSP, São Paulo, 1978.
- [64] W. Hu, Ringing in the New Cosmology, *Nature*, **404** (2000), 939-940.
C. H. Huang, A Non-linear Inverse Vibration Problem of Estimating the Time-dependent Stiffness Coefficients by Conjugate Gradient Method, *Int. J. Num. Meth. Eng.*, **50** (2001), 1545-1558.
- [65] C. H. Huang, A Non-linear Inverse Vibration Problem of Estimating the Time-dependent Stiffness Coefficients by Conjugate Gradient Method, *Int. J. Num. Meth. Eng.*, **50** (2001), 1545-1558.
- [66] C. H. Huang, An Inverse Vibration Problem for Simultaneously Estimating the Time-dependent Stiffness Coefficients, *4th International Conference on Inverse Problem in Engineering: Theory and Practice (ICIPE-2002)*, Rio de Janeiro, Brazil, (2002).
- [67] V. K. Ivanov, "On Linear Problems which are not Well-posed", *Soviet Math. Dokl.*, **3** (1962), 981-983.
- [68] E. T. Jaynes, Information Theory and Statistical Mechanics, *Physical Review*, **106** (1957), 620-630.
- [69] E. T. Jaynes, On the rationale of the maximum entropy methods, *Proceedings of the IEEE*, **70** (1982), 939-952.
- [70] W. F. Jones, A. T. Price, The perturbations of alternating geomagnetic fields by conductivity anomalies, *Geophys. J. R. Astr. Soc.*, **20** (1970), 317-334.
- [71] A. C. Kak, M. Slaney, "Principles of Computerized Tomography Imaging", IEEE Press, New York, 1988.
- [72] P. S. Kasibhatla (Editor), "Inverse Methods in Global Biogeochemical Cycles", American geophysical Union, Washington, 2000.
- [73] S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, Optimization by Simulated Annealing, *Science*, **220** (1983), 671-680.
- [74] S. Kidder, T. Haar, "Satellite Meteorology: An Introduction", Academic Press, 1995.
- [75] A. N. Kolmogorov, S. V. Fomin, "Elementos da Teoria das Funções e de Análise Funcional", Mir, Moscou, 1982.
- [76] M. L. Krasnov, A. I. Kisseliov, G. I. Makarenko, "Equações Integrais", Mir, Moscou, 1981.

- [77] K. Krishnakumar, Micro-Genetic Algorithms for Stationary and Non-Stationary Function Optimization, *Intelligent Control and Adaptive Systems*, SPIE 1196 (1989).
- [78] C. Lanczos, “The Variational Principles of Mechanics”, Dover, 1970.
- [79] K. Levenberg, A Method for the Solution of Certain Problems in Least Squares, *Quart. Appl. Math.*, **2** (1944), 164-168.
- [80] K.N. Liou, “An Introduction to Atmospheric Radiation”, Academic Press, 1980.
- [81] S. Lipschutz, “Teoria dos Conjuntos - Coleção Schaum”, McGraw-Hill, 1970.
- [82] D. G. Luenberger, “Introduction to Linear and Nonlinear Programming”, Addison Wesley, 1973.
- [83] E. F. P. Luz, “Estimação de Fonte de Poluição Atmosférica usando Otimização por Enxame de Partículas”, Dissertação de Mestrado, Computação Aplicada, INPE, São José dos Campos, 2007.
- [84] E. F. P. Luz, H .F. Campos Velho, J. C. Becceneri, D. R. Roberti, Estimating Atmospheric Area Source Strength Through Particle Swarm Optimization, *Inverse Problems, Desing and Optimization Symposium (IPDO-2007)*, April 16-18, Miami (FL), USA.
- [85] P. Mahey, “Programação Não-Linear: Introdução à Teoria e aos Métodos”, Campus (Projeto Arquimedes), 1987.
- [86] D. Marquardt, An Algorithm for Least-squares Estimation of Nonlinear Parameters, *SIAM J. Appl. Math.*, **11** (1963), 431–441.
- [87] N.J. McCormick, Inverse radiative transfer problems: a review. *Nuclear Science and Engineering*, **112** (1992), 185–198.
- [88] F. Medeiros, “Algoritmo Genético Híbrido como um Método de Busca de Estados Estacionários de Sistemas Dinâmicos”, Dissertação de Mestrado em Computação Aplicada - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2002.
- [89] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equations of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, **21** (1953) 1087–1092.
- [90] M. Mitchell, “An introduction to Genetic Algorithms”, MIT Press, USA, 1996.
- [91] C. D. Mobley, “Light and Water - Radiative Transfer in Natural Waters”, Academic Press, San Diego, 1994.
- [92] C. D. Mobley, “Hydrolight 3.0 User’s Guide”, SRI Intl., Menlo Park, 1995.
- [93] V. A. Morosov, On the Solution of Functional Equations by the Method of Regularization, *Soviet Math. Dokl.*, **7** (1966), 414–417 .
- [94] V. A. Morosov, “Methods for Solving Incorrectly Posed Problems”, Springer Verlag, 1984.
- [95] V. A. Morozov, “Regularization Methods for Ill-Posed Problems”, CRC Press, 1992.

- [96] W. B. Muniz, H. F. Campos Velho, F. M. Ramos, A Comparison of Some Inverse Methods for Estimating the Initial Condition of the Heat Equation, *J. Comp. Appl. Math.*, **103** (1999), 145–163.
- [97] W. B. Muniz, “Um Problema Inverso em Condução do Calor utilizando Métodos de Regularização”, Tese de Mestrado, Matemática Aplicada UFRGS, 1999.
- [98] W. B. Muniz, F. M. Ramos, H. F. Campos Velho, Entropy- and Tikhonov-based Regularization Techniques Applied to the Backwards Heat Equation, *Comp. & Math. Applications*, **40** (2000), 1071-1084.
- [99] NASA. International Space Station Imagery: STS097-704-080. 2000.
- [100] F. Natterer, “The Mathematics of Computerized Tomography”, John Wiley & Sons, New York, 1986.
- [101] J. L. Neves, J. P. Braga, A. P. Braga, M. B. Almeida, Recurent Neural Network Model to Retrieve the Long Range Spherical Potential Energy Function from Second Virial coefficient, *Inverse Problems in Engineering*, **10** (2002), 153–162.
- [102] G. K. Nicholls, “Bayesian Inference and Markov Chain Monte Carlo”, Lectures Notes, in *Bayesian Methods in Inverse Problems*, Summer course in Kuopio, Finland, 21-24 June, 2004.
- [103] S. Nieman, T. Achtor, “ITPP-5.0 User Guide”. University of Wisconsin, 1995.
- [104] F. S. Oliveira, “Inversão Gravimétrica Usando Regularização Entrópica”, Tese de Mestrado em Geofísica, Universidade Federal do Pará, 2005.
- [105] OurGrid, <http://www.ourgrid.org/>
- [106] M.N. Özisik, “Heat Conduction”, Wiley Interscience, New York, 1980.
- [107] F. F. Paes, H. F. Campos Velho, E. F. P. Luz, A. R. Carvalho, Identifying Atmospheric Pollutant Sources Using Artificial Neural Networks, *The Meeting of the Americas, AGU 2008 Joint Assembly*, May 27-30, Fort Lauderdale (FL), USA.
- [108] V. Pereyra, G. Scherer, Least Squares Scattered Data Fitting by Truncated SVDs, *Applied Numerical Mathematics*, **40** (2002), 73–86.
- [109] D. L. Phillips, A Technique for the Numerical Solution of Certain Integral Equations of the First Kind, *J. Assoc. Comput. Mach. (J-ACM)*, **9** (1962), 84–97.
- [110] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, “Numerical Recipes: The Art of Scientific Computing”, Cambridge University Press, 1992.
- [111] R. P. Pazos, M. T. Vilhena, Convergence in Transport Theory, *Applied Numerical Mathematics*, **30** (1999), 79–92.
- [112] A. J. Preto, H. F. Campos Velho, J. C. Becceneri, N. N. Arai, R. P. Souto, S. Stephany, A New Regularization Technique for an Ant-colony Based Inverse Solver Applied to a Crystal Growth Problem, *13th Inverse Problem in Engineering Seminar (IPES-2004)*, 14-15 June 2004, University of Cincinnati, Ohio, USA, 147-153.
- [113] F. M. Ramos, A. Giovannini, Résolution d’un problème inverse multidimensionnel de diffusion de la chaleur par la méthode des éléments analytiques et par le principe de l’entropie maximale, *International Journal of Heat and Mass Transfer*, **38** (1995), 101-111.

- [114] F. M. Ramos, H. F. Campos Velho, Reconstruction of Geoelectric Conductivity Distributions Using a Minimum First-Order Entropy Technique, *2nd International Conference on Inverse Problems on Engineering*, Le Croisic, France, vol. 2, 199–206.
- [115] F. M. Ramos, H. F. Campos Velho, J.C. Carvalho, N.J. Ferreira, Novel Approaches on Entropic Regularization, *Inverse Problems*, bf 15 (1999), 1139–1148.
- [116] L. Rebollo-Neira, J. Fernandez-Rubio, A. Plastino, A Non-extensive Maximum Entropy Based Regularization Method for Bad Conditioned Inverse Problems, *Phys. A*, **261** (1998), 555–568.
- [117] M. R. Retamoso, “Estimação de Condições de Contorno e do Termo de Fonte em Transferência Radiativa de Águas Naturais”, Tese de Doutorado, Engenharia Mecânica, UFRGS, Porto Alegre (RS), 2000.
- [118] D. R. Roberti, H. F. Campos Velho, G. A. Degrazia, Identifying Counter-gradient Term in Atmospheric Convective Boundary Layer, *Inverse Problems in Engineering*, **12** (2004), 329–339.
- [119] D. R. Roberti, D. Anfossi, H. F. Campos Velho, G. Degrazia, Estimação da Taxa de Emissão e Localização de Fontes, *Ciencia e Natura*, (2005), 131–134.
- [120] D. R. Roberti, Problemas Inversos em Física da Atmosfera, Tese de Doutorado em Física, Universidade Federal de Santa Maria, 2005.
- [121] D. R. Roberti, D. Anfossi, H. F. Campos Velho, G. Degrazia, Estimation of Pollutant Source Emission Rate from Experimental Data, *Nuovo Cimento* **30** (2007), 177–186.
- [122] C. D. Rodgers, “Inverse Methods for Atmospheric Sounding: Theory and Practice”, in *Series on Atmospheric, Oceanic and Planetary Physics, Vol. 2*, World Scientific, 2000.
- [123] A. Rytter, “Vibration Based Inspection of Civil Engineering Structures”, Ph.D. dissertation, Dept. of Building Technology and Structural Engineering, Aalborg University, Denmark, 1993.
- [124] S. M. B. Sambatti, “Diferentes Estratégias de Paralelização de um Algoritmo Genético Epidêmico, Aplicadas na Solução de Problemas Inversos”, Tese de mestrado em Computação Aplicada, CAP-INPE, São José dos Campos (SP), 2004.
- [125] S. B. Sambatti, H. F. Campos Velho, L. D. Chiwiacowsky, S. Stephany, A. J. Preto, Some parallel strategies for an epidemic genetic algorithm applied to an inverse heat conduction problem, *International Conference on Computational and Experimental Engineering and Sciences (ICCES)*, 26-29 July, Ilha da Madeira, Portugal - paper ICCES04-435 (6 pages), 2004.
- [126] C. F. Segatto, M. T. Vilhena, Extension of the LTS_N Formulation for Discrete Ordinates Problem without Azimuthal Symmetry, *Annals of Nuclear Energy*, **21** (1994), 701–710.
- [127] C. E. Shannon, W. Weaver, “The Mathematical Theory of Communication”, Univ. of Illinois Press, 1949.

- [128] E. H. Shiguemori, H. F. Campos Velho, F. M. Ramos, J. D. S. da Silva, A Parametric Study of a New Regularization Operator: Non-extensive Entropy, *4th International Conference on Inverse Problems in Engineering: Theory and Practice* (ICIPE-2002), May 26-31, Angra dos Reis (RJ), Brazil - Proc. in CD-Rom: paper code 093 - Proceedings Book: Inverse Problems in Engineering: Theory and Practice (Editor Helcio R. B. Orlande), **Vol. II**, Part Heat and Mass Transfer, pp. 243–250 (2002) - [=] e-papers, Rio de Janeiro, Brasil.
- [129] E. H. Shiguemori, H. F. Campos Velho, J. D. S. Silva, Estimation of Initial Condition in Heat Conduction by Neural Network, *Inv. Prob. Sci. Eng.*, **12** (2004), 317–328.
- [130] E. H. Shiguemiri, H. F. Campos Velho, J. D. S. da Silva, Generalized Morozov's Principle, *Inverse Problems, Design and Optimization Symposium* (IPDO), 17-19 March, Rio de Janeiro (RJ), Brazil, Proceedings in CD-Rom, paper code IPDO-079 (8 pages), Vol. 2, 290–298 (2004).
- [131] E. H. Shiguemori, J. D. S. Silva, H. F. Campos Velho, J. C. Carvalho, Neural Network based Models in the Inversion of Temperature Vertical Profiles from Satellite Data, *Inv. Prob. Sci. Eng.*, **14** (2006), 543–556.
- [132] E. H. Shiguemori, “Recuperação de Perfis de Temperatura e Umidade da Atmosfera a partir de Dados de Satélite – Abordagens por Redes Neurais Artificiais e Implementação em Hardware”, Tese de Doutorado, Computação Aplicada, INPE-13269, São José dos Campos (SP), 2007
- [133] E. H. Shiguemori, J. D. S. Silva, H. F. Campos Velho, J. C. Carvalho, Atmospheric temperature retrieval by a Radial Basis Function neural network, *International Journal of Information and Communication Technology* **1** (2008), 224–239.
- [134] E. H. Shiguemori, J. D. S. Silva, H. F. Campos Velho, Atmospheric Temperature Retrieval from Satellite Data: New Non-extensive Artificial Neural Network Approach, *Proceedings of the 2008 ACM symposium on Applied computing* (Session: Artificial Intelligence in Space Applications), March 16-20, Fortaleza (CE), Brazil (2008), pp 1688–1692, ISBN 1-59593-753-7.
- [135] A. J. Silva Neto, F. D. Moura Neto, “Escolha de Modelos: Problemas Inversos em Engenharia”, Mini-curso, *Congresso Nacional de Matemática Aplicada e Computacional*, SBMAC, Santos (SP), Brasil, 2000.
- [136] J. B. C. Silva, V. C. F. Barbosa, F. S. Oliveira, H. F. Campos Velho, Gravity Inversion Using Entropic Regularization, *Geophysics*, **74** (2007), 151–160.
- [137] R. T. Smith, C. K. Zoltani, G. J. Klem, M. W. Coleman, Reconstruction of tomographic images from sparse data sets by a new finite element maximum entropy approach, *Applied Optics*, **30** (1991), 573–582.
- [138] H. Sohn, C. Farrar, F. Hemez, D. Shunk, D. Stinemat, B. Nadler, “A Review of Structural Health Monitoring Literature: 1996-2001”, Los Alamos National Laboratory (LA-13976-MS), USA, 2003.
- [139] R. P. Souto, “Recuperação de Perfis Verticais de Propriedades Óticas Inerentes a partir da Radiação Emergente da Água”, Tese de Doutorado, Computação Aplicada, INPE, São José dos Campos (SP).

- [140] R. P. Souto, H. F. Campos Velho, S. Stephany, E. S. Chalhoub, Performance Analysis of Radiative Transfer Algorithms in a Parallel Environment, *Transport Theory and Statistical Physics*, **33** (2004), 449–468.
- [141] R. P. Souto, V. C. F. Barbosa, H. F. Campos Velho, S. Stephany, Determining Chlorophyll Concentration in Off-shore Sea Water from Multi-spectral Radiances by Using Second Derivative Criterion and Ant Colony Meta-Heuristic, *Inverse Problems, Design and Optimization Symposium (IPDO-2007)*, Miami (USA), **Vol. 1** (2007), 341–348.
- [142] R. P. Souto, F. F. Paes, S. Stephany, H. F. Campos Velho, P. O. A. Navaux, A. S. Charão, A. J. Preto, 3D Tomography of Ocean Chlorophyll Applying Grid Computing, *The Tenth International Conference on Integral Methods in Science and Engineering (IMSE-2008)*, July 07-10, Santander, Spain, 2008.
- [143] R. P. Souto, H. F. Campos Velho, S. Stephany, V. C. F. Barbosa, Multi-spectral Inversion for Chlorophyll Concentration in Offshore Sea Water by Using the Ant Colony Optimization and the Second Derivative Criterion, *International Journal of Remote Sensing*, (2008) – submetido.
- [144] F. L. Souza, F. M. Ramos, P. Paglione, R. M. Girardi, New Stochastic Algorithm for Design Optimization, *AIAA Journal*, **41** (2003) 1808–1818.
- [145] F. L. Sousa, “Otimização Extrema Generalizada: Um Novo Algoritmo Estocástico para o Projeto Ótimo”, Tese de Doutorado, Computação Aplicada, INPE-9564-TDI/836, São José dos Campos (SP), 2003.
- [146] S. Stephany, “Reconstrução de Propriedades Óticas Inerentes e Bioluminescência em Águas Naturais”, Tese de Doutorado em Computação Aplicada, Instituto Nacional de Pesquisas Espaciais, 1997.
- [147] S. Stephany, F. M. Ramos, H. F. Campos Velho, C. D. Mobley, A Methodology for Internal Light Sources Estimation, em Computer Modeling and Simulation in Engineering (CMSE), **3** (1998), 161–165.
- [148] S. Subbey, “Regularizing the Volterra Integral Equation - The Capillary Pressure Case”, Ph.D. Dissertation, Dept. of Physics, University of Bergen, Bergen, Norway, 2000.
- [149] A. Tarantola, “Inverse Problem Theory and Methods for Model Parameter Estimation”, SIAM, 2004.
- [150] A.N. Tikhonov, V.Y. Arsenin, “Solution of Ill-posed Problems”, John Wiley & Sons, 1977.
- [151] C. Tsallis, Possible Generalization of Boltzmann-Gibbs Statistics, *J. Statistical Phys.*, **52** (1988), 479–487.
- [152] C. Tsallis, Nonextensive statistics: theoretical, experimental and computational evidences and connections, **Braz. J. Phys.**, **29** (1999), 1–35.
- [153] S. Twomey, On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature, *J. Assoc. Comput. Mach. (J-ACM)*, **10** (1963), 97–101.

- [154] C.R. Vogel, “Computational Methods in Inverse Problems”, *in* *Frontiers in Applied Mathematics*, SIAM, 2002.
- [155] J.P. Zubelli, “An Introduction to Inverse Problems: Examples, Methods, and Questions”, Mini-curso no 22o. Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 1999.