



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Escola de verão em Computação Aplicada – LAC-INPE 2008

Introdução aos Problemas Inversos: Aplicações em Pesquisa Espacial

Haroldo Fraga de Campos Velho
Laboratório Associado de Computação e Matemática Aplicada
Instituto Nacional de Pesquisas Espaciais
LAC-INPE – Caixa Postal 515
12245-970 – São José dos Campos (SP)
E-mail: haroldo@lac.inpe.br
home-page: <http://www.lac.inpe.br/~haroldo/>

INTRODUÇÃO AOS PROBLEMAS INVERSOS E APLICAÇÕES EM PESQUISA ESPACIAL

Haroldo F. de Campos Velho

RESUMO

O Mini-Curso apresenta uma introdução sucinta aos problemas inversos. Primeiramente, conceito e classificação de problemas inversos são descritos, para então serem discutidas a técnica de regularização (alguns operadores são apresentados, bem como técnicas de solução de otimização), técnica variacional e redes neurais artificiais. Exemplos de problemas inversos em várias áreas da pesquisa espacial são apresentados e discutidos: ciência espacial (mapas de radiação cósmica de fundo, geofísica espacial), engenharia espacial (detecção de danos em estruturas aero-espaciais, projeto ótimo de radiadores espaciais), oceanografia (propriedade óticas de oceanos), meteorologia (temperatura e umidade atmosféricas a partir de dados de satélites).

EMENTA

1. Introdução	03
2. Métodos Matemáticos para Solução de Problemas Inversos	04
2.1. Técnicas de regularização	06
2.1.1. Regularização de Tikhonov	07
2.1.2. Regularização entrópica	07
2.1.3. Determinação do parâmetro de regularização	08
2.2. Técnicas de Otimização	08
2.2.1. Determinística: Newton e quase-Newton	09
2.2.2. Determinística: Levengerg-Marquardt	10
2.2.3. Determinística: Gradiente conjugado	11
2.2.4. Estocástica: Recozimento simulado com ciclos térmicos	11
2.2.5. Estocástica: Algoritmo genético epidêmico	12
2.3. Método Variacional	13
2.4. Redes Neurais Artificiais	14
3. Aplicações de Problemas Inversos em Pesquisa Espacial	15
3.1. Ciência Espacial: Inversão Magnetotelúrica	16
3.2. Engenharia Espacial	17
3.2.1. Projeto de radiadores espaciais	17
3.2.1. Detecção de danos em estruturais espaciais	18
3.3. Aplicações Espaciais: Oceanografia	19
3.4. Aplicações Espaciais: Meteorologia	21
4. Considerações Finais	22
Referências	23

1. Introdução

A partir de certo estágio do desenvolvimento da sociedade humana, o conhecimento tornou-se cada vez mais compartimentado, primeiramente a ciência é separada em grandes áreas: culturais, biomédicas e exatas. Estas por sua vez se subdividem em dois grandes grupos: ciências básicas e aplicadas; para se subdividirem mais ainda. Por exemplo, nas áreas de ciências exatas tem-se a matemática, a física e a química como exemplos de ciências básicas e as engenharias, geociências e astronomia podem ser encaradas como áreas de aplicação destas ciências. Todavia, cada uma destas ciências, sejam básicas e/ou aplicadas, também podem ser subdivididas em *pura* e *tecnológica*, teórica e experimental. O conhecimento humano é hoje um grande mosaico de especialidades. Há, porém, áreas de estudo que requerem conhecimento de várias especialidades: são as áreas (ciências) multidisciplinares. Problemas inversos (PIs) são exemplos de área multidisciplinar.

A distinção entre o que seja um problema direto ou inverso para um dado fenômeno, está ligada a nossa cultura, isto é, trata-se do que se interpreta como causa e efeito! É atribuído a Oleg Mikailivitch Alifanov (www.me.ua.edu/inverse/whatis.html), proeminente pesquisador russo na área de problemas inversos, a afirmação “a solução de um problema inverso consiste em determinar causas baseado na observação dos seus efeitos”. Do ponto de vista prático, convencionou-se chamar problema direto àquele em que o estudo antecedeu-se historicamente. Tal ambigüidade (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: $B(f) = u \Rightarrow B^{-1}(u) = f!$

Como se percebe, a própria definição de PI pode apresentar controvérsias. Entretanto, nestas notas, a observação de Oleg M. Alifanov é a base para a conceituação de PIs. Uma definição, bem como classificações de PIs, é apresentada na próxima seção. Algumas técnicas de solução de problemas inversos são apresentadas. Estes métodos são exemplificados em algumas aplicações: geofísica, meteorologia, oceanografia e transferência de calor. As últimas seções deste texto estão voltadas a técnicas novas, como redes neurais e algoritmo genético.

Este texto é fortemente baseado na experiência do autor na área e no *Curso de Problemas Inversos*, disciplina regular do Curso de Pós-graduação em Computação Aplicada (CAP) do Instituto Nacional de Pesquisas Espaciais (INPE) (<http://www.lac.inpe.br/cap/>) - ministrada pelo Dr. Fernando M. Ramos. A primeira versão deste texto (mais resumida) foi apresentada como um minicurso do Encontro de Modelagem Computacional em 2001, promoção do Instituto Politécnico da Universidade do Estado do Rio de Janeiro (IPRJ-UERJ, Nova Friburgo), a convite do Prof. Antônio José da Silva Neto. Texto e *slides* do mini-curso do IPRJ-UERJ estão disponíveis na internet (<http://www.lac.inpe.br/~haroldo/scip.html>) - este material do minicurso tornaram-se disponíveis graças ao trabalho voluntário de Elcio H. Shiguemori e Ana Paula Abrantes de Castro (na época, estudantes de doutorado e mestrado, respectivamente, do curso de pós-graduação em Computação Aplicada do INPE).

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), Antônio José da Silva Neto (IPRJ-UERJ), Domenico Anfossi (ISAC-CNR), Ezzat S. Chalhoub (LAC-INPE), Fernando Manuel Ramos (LAC-INPE), Gervásio A. Degrazia (Física, UFSM), João Batista C. Silva (Geofísica, UFPA), José Demísio S. da Silva (LAC-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), Nandamudi L. Vijaykumar (LAC-INPE), Nelson J. Ferreira (CPTEC-INPE), Pedro P. B. de Oliveira (Universidade McKenzie), Stephan Stephany (LAC-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 Fávero 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, Ricardo Varela Correa, Roberto Afonso da Costa Júnior, Roberto L. Galski, Rosângela S. C. Cintra, Sabrina B. M. Sambatti, Wagner B. Muniz. A todos, meu especial agradecimento. Deixo claro, entretanto, que somente ao autor devem ser creditados quaisquer erros ou imperfeições no texto.

As contribuições do grupo de Problemas Inversos do INPE não se referem somente às aplicações da metodologia de problemas inversos em pesquisas espaciais. O grupo tem contribuído também com o desenvolvimento da teoria de problemas inversos (concepção de novas técnicas de regularização), bem como em novos desenvolvimentos em meta-heurísticas e no método variacional.

2. Métodos Matemáticos para Solução de Problemas Inversos

Considera-se o astrofísico georgiano Viktor Amazaspovich Ambartsumian (para mais informações sobre este cientista consultar: <http://www.phys-astro.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. (1996): “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 abaixo 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.

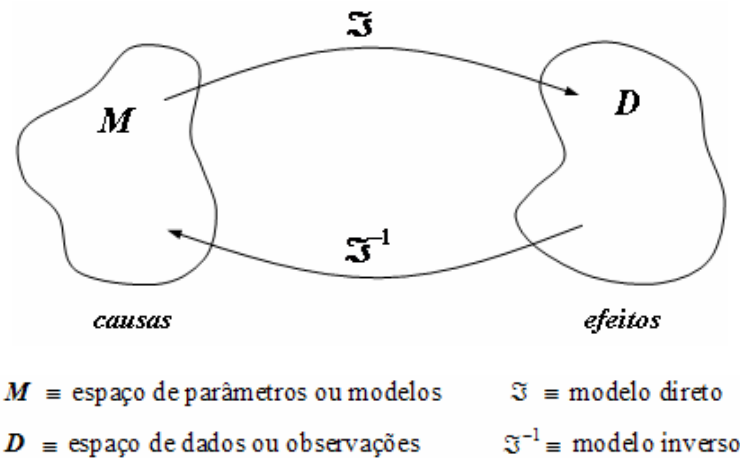


Fig. 1 - Representação esquemática de problemas direto e inverso.

Aliás, o tipo de “causa” a ser determinada pode ser usado para classificar PI. Contudo, outras classificações são possíveis:

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

Nota-se que a classificação dos itens 1 e 2 estão ligadas aos métodos de solução do PI; segundo a do item 3, como já falado, tipifica-se o problema inverso pela *causa* a ser determinada. A classificação do item 4 foi formulada pelo Prof. J. 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 estimção de parâmetros e não estimção 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 $X=[c_1 c_2 \dots c_K]^T$, noutro caso $X=[k, S]^T$. Um caso distinto é estimar uma *função amostrada*: $f(t_i) \equiv f_i$ ($i=1, 2, \dots, N$), nesse caso: $X=[f_1 f_2 \dots f_N]^T$.

A classificação indicada no item 5 foi proposta recentemente (1999) 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.

Matematicamente, problemas inversos pertencem à classe de problemas malpostos. No início deste século, 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 malposto 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:

$$2x - 4 = 0. \tag{1}$$

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

$$ax + b = 0, \tag{2}$$

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

$$ax^2 - 2x + 1 = 0 \tag{3}$$

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. (3) torna-se: $x_{1,2} = 1 \pm 0,1i$, sendo i a unidade dos números imaginários. Ou seja, 1% de ruído em a , faz com que a Eq. (3) não tenha mais solução no campo dos números reais.

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 pelo ponto de vista metodológico. A lista a seguir apresenta as áreas correlatas aos PI:

- Identificação de sistemas;
- Controle ótimo em sistemas estocásticos;
- Álgebra linear computacional em problemas de posto incompleto;
- Restauração 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 (Hu, 2000), 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 bilhões de dólares: 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 tem sido solucionado por software (métodos matemáticos de Ps - ver Hanisch e White, 1993). 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, 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 (Kasibhatla et al., 2000).

2.1. Técnicas de regularização

Não vamos tratar da teoria matemática de equações diferenciais parciais. Embora as propriedades matemáticas sejam relevantes na exposição e desenvolvimento da metodologia, optou-se por uma apresentação sem muitos detalhes matemáticos, onde as técnicas são aplicadas e, espera-se, compreendidas com os exemplos do texto. Uma lista não exaustiva de métodos de solução de PI segue:

- Inversão direta;
- Decomposição em valores singulares;
- Mínimos quadrados e variantes (mínimos quadrados ponderados);
- Métodos de regularização;
- Métodos variacionais;
- Outros (molificação, métodos bayesianos, filtros digitais, redes neurais, etc).

Métodos explícitos, ou de inversão direta, não são métodos gerais. Há um interesse acadêmico numa inversão direta, mas não se constitui num esquema metodológico geral a ser seguido. Para exemplificar e para apresentar um problema que foi usado para testar novas técnicas, seja o problema inverso de identificação da condição inicial em condução do calor:

$$T_t = \alpha T_{xx}; \quad \text{em } t > 0, x \in (0, L); \quad (4)$$

$$\text{condição inicial: } T(x, 0) = f(x),$$

$$\text{condições de contorno: } T_x(0, t) = T_x(L, t) = 0,$$

Para o qual a solução exata do problema direto é

$$T(x, t) = \sum_{m=0}^{\infty} e^{-\beta_m^2 t} \frac{X(\beta_m, x)}{N(\beta_m)} \int_0^1 X(\beta_m, x') f(x') dx', \quad (5)$$

onde $X(\beta_m, x)$ e $N(\beta_m)$ são autofunção e norma características do problema (Ösizik, 1980; Muniz et al, 1999). A solução exata do problema inverso para um perfil de temperatura $T(x, t)$ medida num tempo $t = \tau$ é:

$$f(x) = \sum_{m=0}^{\infty} e^{\beta_m^2 \tau} \frac{X(\beta_m, x)}{N(\beta_m)} \int_0^1 X(\beta_m, x') T(x', \tau) dx'. \quad (6)$$

Pode-se mostrar que a solução (6) define o PI como um problema mal posto, pois viola a 3ª condição de Hadamard (Muniz et al., 1999).

Outro problema resolvido explicitamente é a solução de Burgraff para identificação de fluxo superficial em problemas de condução de calor (ver Beck et al., 1985, página 67).

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 (1962), D.L. Phillips (1962) e S. Twomey (1963) merecem destaque, mas foi com o trabalho de Andrei Nikolaevich Tikhonov em 1963, o início de uma formulação geral para problemas malpostos, chamada *regularização* 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 mais 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 malposto num problema bem-posto (ver Figura 2).

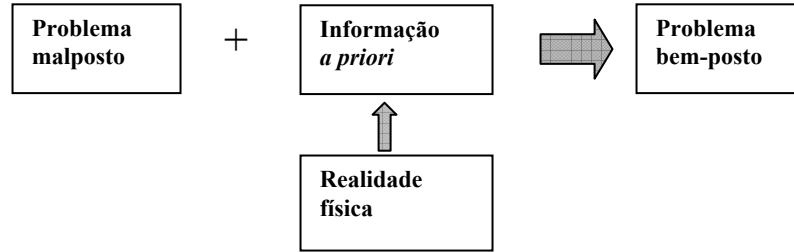


Fig. 2 - Idéia básica para solução de problemas mal-postos: informação adicional (*a priori*) é incorporada ao problema e o problema modificado (muito similar ao problema original) é um problema bem-posto!

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

$$\min_{u \in U} \|A(u) - f^\delta\|_2^2 \quad \text{sujeito a: } \mathcal{Q}[u] \leq \rho, \quad (14)$$

onde $A(u) = f^\delta$ representa o modelo direto e $\mathcal{Q}[u]$ é o operador de regularização (Tikhonov e Arsenin, 1977). 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 (15) - e de regularidade (suavidade) - segundo termo da expressão abaixo - exigida da quantidade desconhecida:

$$\min_{u \in U} \left\{ \|A(u) - f^\delta\|_2^2 + \alpha \|\mathcal{Q}[u]\| \right\}, \quad (15)$$

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.

Vários tipos de operadores de regularização têm sido investigados desde o trabalho dos pioneiros na área. Aqui somente serão comentadas duas classes destes operadores de regularização.

2.1.1. Regularização de Tikhonov

Aqui o operador é expresso por:

$$\mathcal{Q}[u] = \sum_{k=0}^p \mu_k \|u^{(k)}\|_2^2 \quad (16)$$

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

$$\mathcal{Q}[u] = \|u^{(k)}\|_2^2;$$

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). Já na regularização de 1ª ordem, toma-se $u^{(1)} \approx 0$, ou seja, u é aproximadamente constante.

2.1.2. Regularização entrópica

De maneira similar à 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 (1957) como um critério geral de inferência, baseado na teoria matemática da informação de Shannon (Shanon e Weaver, 1949). A Figura 3 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 3a), enquanto o estado de mínima

entropia é ilustrado na Figura 3b: 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:

$$S(u) = \sum_{q=1}^N s_q \log(s_q), \quad \text{com } s_q = r_q^{(k)} / \sum_{q=1}^N r_q^{(k)}; \quad (17)$$

válida para o caso discreto. Neste contexto, as diferentes ordens do operador são (Campos Velho e Ramos, 1997; Ramos et al., 2000, Campos Velho et al., 2000):

$$r_q^{(k)} = \begin{cases} u_q & \text{para } k = 0 \\ u_{q+1} - u_q + (u_{\max} - u_{\min}) + \zeta & \text{para } k = 1 \\ u_{q+1} - 2u_q + u_{q-1} + 2(u_{\max} - u_{\min}) + \zeta & \text{para } k = 1 \end{cases} \quad (18)$$

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 (Muniz et al., 2000).

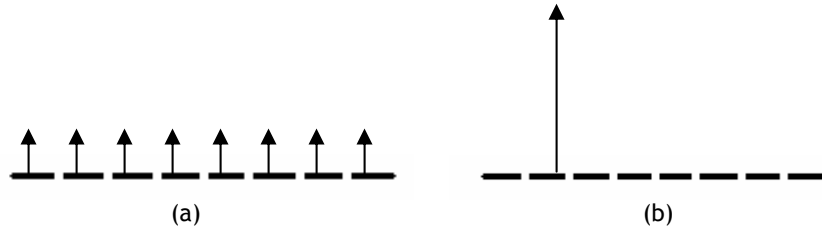


Fig. 3 - (a) estados igualmente prováveis, (b) todos estados tem probabilidade nula de ocorrência exceto um único estado com probabilidade de 100% de ocorrência.

2.1.3. Determinação do Parâmetro de Regularização.

Para se ter uma teoria de regularização completa, é necessário um método para calcular o parâmetro α na Eq. (15). O parâmetro de regularização é que vai realizar o balanço entre o termo da diferença quadrática entre os dados e o modelo - 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 (Bertero e Bocaccio, 1998), como o método da curva L e o da validação cruzada. Outra metodologia empregada está baseada no critério da discrepância de Morosov.

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 erro no processo de medida, α^* é a raiz da equação

$$\|A(u) - f^\delta\|_{\alpha^*} = \delta. \quad (19)$$

O critério da discrepância foi validado também para o caso do operador de entropia de alta ordem (Muniz et al., 2000).

2.2. Técnicas de Otimização

A solução inversa no método de regularização é obtida resolvendo-se o problema de otimização expresso pela Eq. (15). 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 (AS: *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*), etc.
- Métodos híbridos: Combinam a estratégia de busca global dos métodos estocásticos com busca local dos métodos deterministas (GAPlex: GA + Simplex, SAPlex: AS + Simplex, GAGC: GA + Gradiente Conjugado).

2.2.1. 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:

$$Q(X + \Delta X) = Q(X) + \nabla Q(X) \cdot \Delta X + \frac{1}{2!} \nabla[\nabla Q(X) \cdot \Delta X] \cdot \Delta X + O(\Delta X^3). \quad (20)$$

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

$$X_{k+1} = X_k + [\nabla \nabla J(X_k)]^{-1} \nabla J(X_k). \quad (21)$$

Do ponto de vista computacional não calculamos explicitamente a inversa da matriz hessiana da função $F(X)$, 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. Resolva o sistema linear de equações algébricas: $[\nabla \nabla J(X_k)] \Delta X_{k+1} = -\nabla J(X_k)$
2. Atualize o vetor de parâmetros: $X_{k+1} = \Delta X_{k+1} + X_k$ (22)
3. Avaliação da convergência: $\frac{\|X_{k+1} - X_k\|}{\|X_{k+1}\|} < \varepsilon$ $\begin{cases} \text{sim: processo finaliza} \\ \text{não: retorna ao passo 1} \end{cases}$

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

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

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*, como muitas vezes é denominado na literatura (Luenberger, 1973). O processo implementado na rotina E04UCF da biblioteca numérica NAG (1995) é o seguinte:

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

$$H_k = H_{k-1} + \frac{b_k (b_k)^T}{(b_k)^T u_k} - \frac{H_{k-1} u_k (u_k)^T H_{k-1}}{(u_k)^T H_{k-1} u_k} \quad (23)$$

onde

$$b_k = X_k - X_{k-1} = \Delta X_k$$

$$u_k = \nabla J(X_k) - \nabla J(X_{k-1}).$$

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

$$\text{Minimize } [\nabla J(X_k)]^T d^k + \frac{1}{2} (d^k)^T (H_k) d^k, \text{ sujeito a: } l_q - p_q^n \leq d_q \leq u^n - p_q^n.$$

5. Fixe, $X_{k+1} = X_k + \beta^k d^k$, onde o passo de comprimento β^k minimiza $J(X_k + \beta^k d^k)$.

6. Testa a convergência: finaliza ou retorna ao Passo 1.

É possível demonstrar que a solução ótima do processo acima, converge para solução do método de Newton (Luenberger, 1973). 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 à escolha da estimativa inicial é o método da máxima descida, que tem convergência linear (Luenberger, 1973).

2.2.2. 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 (1944) e Marquardt (1963), que combina as estratégias de máxima descida com o método de Newton.

Denotando $J(X) \equiv [A(X) - f^0]$, onde $A(\cdot)$ representa o modelo direto, f^0 representa os dados experimentais e $Q \equiv \nabla J$ (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:

$$M \rightarrow \hat{M}; \text{ onde: } \hat{M}_{ij} = M_{ij} + \lambda M_{ii} \delta_{ij}. \quad (24)$$

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:

$$X_{k+1} = X_k - [Q^T Q + \lambda_k I]^{-1} Q^T J. \quad (25)$$

A rotina DBCLSF da biblioteca numérica IMSL (1991) implementa uma versão do método de Levenberg-Marquardt.

2.2.3. 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 como, por exemplo, em Luenberger (1973). Outra referência a ser mencionada é Chiwiacowsky (2005): além da dedução do MGC, o método variacional é descrito com detalhes na sua aplicação a problemas inversos (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): $\gamma^k = \frac{\|\nabla J(X_k)\|_2^2}{\|\nabla J(X_{k-1})\|_2^2}$;
5. Cálculo da direção de busca:
$$p^k = \begin{cases} -J(X_k) & \text{para } k=0 \\ -J(X_k) + \gamma^k p^{k-1} & \text{para } k > 0 \end{cases};$$
6. Realizar uma busca em linha na direção p^k para calcular $\beta^k = \min_{\beta} \{J(X_k + \beta p^k)\}$;
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*: $(p^{k-1})^T A p^k = 0$.

2.2.4. 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, à 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. (1983) e é uma adaptação do algoritmo de Metropolis, um método de Monte Carlo empregado em sistemas termo-dinâmicos (Metropolis et al., 1953). O algoritmo a seguir foi implementado por Press et al. (1992):

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: $J(X_t^m)$;
3. Escolher novas soluções candidatas aleatoriamente e recalculando 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: se $J(X_t^m) < J(X_{t-1}^m)$ a solução é aceita, caso contrário se um valor aleatório $rand$ satisfizer: $rand > e^{-\Delta E/T_t}$, onde $\Delta E \equiv |J(X_t^m) - J(X_{t-1}^m)|$, a solução é aceita;
5. Redução da temperatura T : $T_{t+1} = r * T_t$ (r é o fator de redução);
6. Teste de convergência: finaliza ou retorna ao passo-1.

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

2.2.5. 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.

Os procedimentos nos GA são concebidos a partir da versão moderna da teoria da evolução - a *grande síntese* (Holland, 1975). 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 (Chiwiacowsky e Campos Velho, 2003; Medeiros, 2005). 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^m ($m = 1, 2, \dots, M$);
 2. Cálculo da aptidão (função objetivo) de cada indivíduo: $J(X^m)$;
 3. Ativa operador de seleção (estratégia torneio, ver: Chiwiacowsky e Campos Velho, 2003);
 4. Ativa o operador de cruzamento: $z_i = x_i^\mu y_i^{1-\mu}$ (x_i e y_i são 2 pais selecionados ($x_i, y_i \in X^m$) e z_i um indivíduo da nova geração, $\mu=0,25$ é o parâmetro usado, que equivale igual peso para ambos os pais);
 5. Ativa operador de mutação: $z_i' = \begin{cases} z_i + \Delta(t, l_{\text{sup}} - z_i) & \text{se: dígito aleatório binário é 0} \\ z_i - \Delta(t, z_i - l_{\text{inf}}) & \text{se: dígito aleatório binário é 1} \end{cases}$;
- onde: $\Delta(t, y) \equiv y \left[1 - \text{rand}^{(1-t/T)^b} \right]$ (rand =número aleatório $[0,1)$, t é passo da iteração atual, T é o número máximo de iterações e $b=5$);
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 (Krishnakumar, 1989). Cuco et al. (2008) discutem a similaridade entre as duas abordagens.

2.3. 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. 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 $\hat{u}(x) = u(x) + \varepsilon\eta(x)$, a variação da função $u(x)$ é dada por:

$$\delta u = \hat{u}(x) - u(x) = \varepsilon\eta(x) , \quad (26)$$

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 $u(x)$. De forma mais explícita: du se refere a uma mudança infinitesimal de $u(x)$ causada pela variação infinitesimal do argumento dx , isto é, $du = u(x+dx) - u(x)$; enquanto que δu é uma variação infinitesimal de $u(x)$ que produz uma nova função: $\delta u = \varepsilon\eta(x)$.

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 $F(\cdot)$:

$$\delta I = \delta \int_a^b F[u(x), u'(x), x] dx = \int_a^b \delta F dx = \varepsilon \int_a^b \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial u'} \eta' \right) dx, \quad x \in (a, b); \quad (27)$$

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

$$\delta I = \varepsilon \int_a^b \eta \left(\frac{\partial F}{\partial u} - \frac{d}{dx} \frac{\partial F}{\partial u'} \right) dx + \left(\varepsilon \eta(x) \frac{\partial F}{\partial u'} \right)_a^b. \quad (28)$$

Para determinar o ponto estacionário do funcional ($\delta I / \varepsilon = 0$), 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:

$$\frac{\partial F}{\partial u} - \frac{d}{dx} \frac{\partial F}{\partial u'} = 0. \quad (29)$$

A equação (29) 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, não é possível 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 para os multiplicadores de Lagrange são as *equações adjuntas*.

2.4. Redes Neurais Artificiais

Redes neurais artificiais (RNAs) são arranjos de elementos de processamento chamados neurônios (Figura 2b). O modelo de neurônio artificial consiste de uma combinação linear seguida de uma função de ativação (ver Figura 4a):

$$y_k = \varphi(v_k) = \varphi \left(\sum_{j=1}^n w_{kj} x_j + b_k \right) \quad (30)$$

onde w_{kj} são os pesos das conexões e b_k é o viés (Haykin, 1993). A função de ativação pode assumir várias formas funcionais. As mais usadas são as funções sigmóide e tangente hiperbólica:

$$\varphi(x) = \begin{cases} (1 + e^{-ax})^{-1} & \text{(a)} \\ \tanh(x) = (1 - e^{-ax}) / (1 + e^{-ax}) & \text{(b)} \end{cases} \quad (31)$$

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 $Y^{\text{rede}}(W)$, onde W é a matriz dos pesos das conexões. Para determinar a matriz, o objetivo é minimizar a diferença quadrática:

$$J(W) = \| Y_i^{\text{alvo}} - Y_i^{\text{RN}}(W) \|_2^2. \quad (32)$$

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

$$W_{kj}^m = W_{kj}^{m-1} + \Delta W_{kj}; \text{ onde: } \Delta W_{kj} = \eta \frac{\partial E}{\partial W_{kj}} = \eta \delta_j x_{jk}, \quad (33)$$

$$E = \sum_{n \in \text{saída da RN}} [y_n^{\text{alvo}} - y_n^{\text{RN}}(W_{kj})]^2,$$

$$\frac{\partial E}{\partial W_{kj}} = \frac{\partial E}{\partial Y_j^{\text{RN}}} \frac{\partial Y_j^{\text{RN}}}{\partial W_{kj}} = \frac{\partial E}{\partial Y_j^{\text{RN}}} x_j \Rightarrow \delta_j \equiv \frac{\partial E}{\partial Y_j^{\text{RN}}}.$$

Aqui η é chamada de taxa de aprendizagem. A quantidade E representa na verdade o erro entre o valor de saída da rede (Y^{RN}) e o valor de referência (y^{alvo}). Na verdade, qualquer técnica de otimização estudada até aqui (como aquelas vistas nas seções anteriores) pode ser usada para encontrar a matriz de pesos de conexão W .

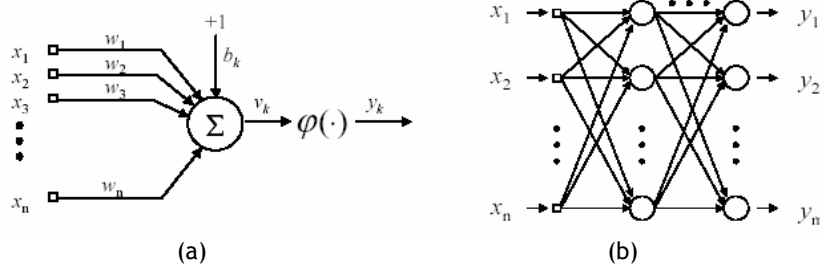


Fig. 4 - (a) neurônio artificial simples, (b) rede neural de múltiplas camadas.

Porque a regra δ funciona? A resposta é simples. Se a direção de busca (ΔW) na equação (33) fosse dada em termos do método de Newton, na expressão ter-se-ia o produto da inversa da matriz hessiana 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: $H^{-1} = \eta I$, onde I é a matriz identidade.

3. Aplicações de Problemas Inversos em Pesquisa Espacial

A área espacial tem 4 grandes pilares: (i) ciência espacial; (ii) engenharia ou tecnologia espacial; (iii) aplicações espaciais (sensoriamento remoto, meteorologia por satélites, sistemas globais de navegação por satélites - GNSS: *Global Navigation Satellite Systems*, telefonia celular e muitas outras), (iv) medicina espacial. O INPE atua em 3 das áreas mencionadas e existe um grupo de estudos em medicina espacial no Brasil na PUC do Rio Grande do Sul. O grupo de problemas inversos realiza pesquisa e desenvolvimento em todas as áreas de atuação do INPE. Como exemplos de aplicações de PI no INPE, citamos os seguintes:

- **Ciência espacial:** inversão magnetotelúrica (geofísica espacial), produção de mapas da radiação cósmica de fundo (astrofísica), projeto de configuração da distribuição de antenas de radio-telescópio interferométrico;
- **Tecnologia espacial:** projeto ótimo de radiadores espaciais, projeto ótimo do controle térmico da plataforma multi-missão (PMM) de satélites do INPE, detecção de danos em estruturas espaciais;
- **Meteorologia:** identificação de perfis de temperatura e umidade atmosféricas a partir de dados de satélites, detecção de fontes de poluição atmosféricas, estimação de propriedades turbulentas em fluxo da camada limite atmosférica;
- **Oceanografia:** propriedades óticas de águas naturais;
- **Assimilação de dados:** aplicações em meteorologia, oceanografia e clima espacial.

Não serão tratados todos os exemplos citados acima. Selecionamos alguns para ilustrar a aplicação das metodologias de problemas inversos.

3.1. Ciência Espacial: Inversão Magnetotelúrica (IM)

Áreas clássicas de problemas inversos são: geofísica, transferência de calor, reconstrução de imagens. A área de geofísica está ligada ao maior negócio do planeta, a área do petróleo, em particular em prospecção de petróleo. Há outras aplicações relevantes em geofísica, como mineração e busca de água subterrânea. A metodologia mais empregada nesta área é o uso de métodos sísmicos, onde cargas (excitação artificial no sistema) são detonadas e registros são captados por geofones. Inversão magnetotelúrica está baseada em fontes eletromagnéticas naturais - raios cósmicos, por exemplo - onde magnetômetros são usados para registrar o campo geomagnético na superfície da Terra. A formulação matemática do problema direto é estabelecida pelas equações de Maxwell do eletromagnetismo num domínio bidimensional (2D) com condições de contorno adequadas (Ramos e Campos Velho, 1996; Campos Velho e Ramos, 1997).

O domínio do problema físico é esboçado na Figura 5, onde Ω^+ e Ω^- correspondem às zonas condutiva ($z < 0$) e de espaço livre ($z > 0$), respectivamente. O campo oscila com um período $2\pi\omega$, suficientemente longo para permitir que correntes de deslocamento possam ser ignoradas. A permeabilidade é tomada como unitária. As equações do eletromagnetismo tornam-se:

$$\nabla \times \vec{H} = 4\pi\sigma\vec{E} \quad (34a)$$

$$\nabla \times \vec{E} = -i\omega\vec{H} \quad (34b)$$

O fator $\exp(i\omega t)$ é assumido em todas as quantidades de campo: $\Phi(y, z, t) = \exp(i\omega t)\phi(y, z)$ ($\Phi = E, H$) e $\sigma = \sigma(y, z)$ é a condutividade elétrica. Visto que H e E não dependem da direção x , as equações acima são reduzidas a um conjunto de 2 equações de Poisson independentes (polarização- H e polarização- E) - o desenvolvimento detalhado pode ser visto em Campos Velho e Ramos (1997). Por simplicidade, considera-se somente o problema da polarização- H na análise inversa:

$$\nabla^2 H_x = \frac{\partial^2 H_x}{\partial y^2} + \frac{\partial^2 H_x}{\partial z^2} = i\eta^2 H_x \quad ; \quad \text{onde: } \eta^2 = 4\pi\sigma\omega \quad (35)$$

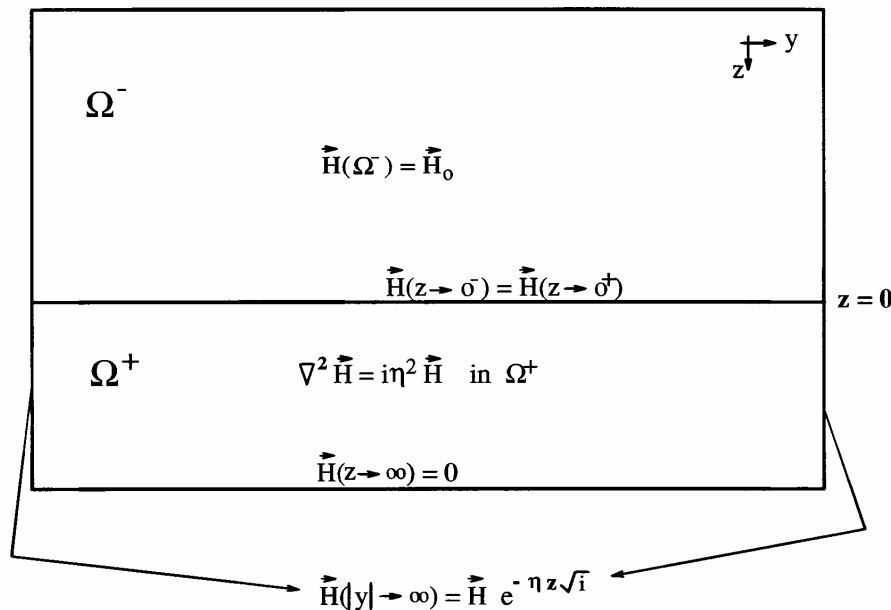


Figura 5: Geometria e condições de contorno do problema direto.

O problema inverso para se determinar o campo de condutividade elétrica $\sigma(y,z)$ é formulado como um problema de otimização com restrições. O operador de regularização é uma combinação da máxima entropia de ordem-zero com a *minimização* da entropia de 1ª ordem - equações (17) e (18) (Ramos e Campos Velho, 1996; Campos Velho e Ramos, 1997). A solução inversa é obtida com a solução ótima do seguinte funcional:

$$J(\sigma) = [H_x^{Exp}(y, z=0) - H_x^{Mod}(y, z=0, \sigma)]^2 - \gamma_0 (S_0(\sigma)/S_{max}) + \gamma_1 (S_1(\sigma)/S_{max}). \quad (36)$$

A Figura 6 mostra os resultados da inversão sem regularização em (6b), mínima entropia de 1ª ordem em (6c) e máxima entropia de ordem-zero em (6d) - para comparação, a solução verdadeira é mostrada em figura 6a.

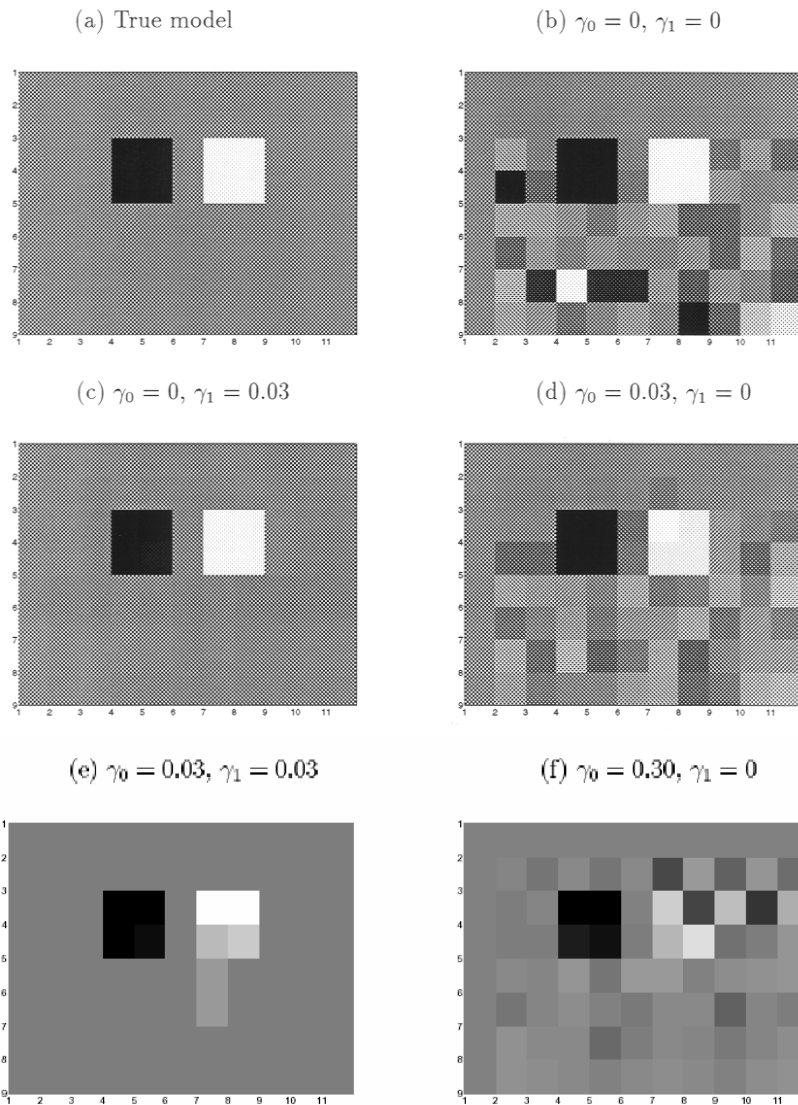


Figura 6: Mapas de condutividade elétrica em IM: (a) modelo verdadeiro, (b) inversão sem regularização, (c) regularização MinEnt-1, (d) regularização MaxEnt-0, (e) e (f) inversões com 2 tipos de entropia.

3.2. Tecnologia Espacial

3.2.1. Projeto de radiadores espaciais

Radiadores espaciais são dispositivos usados para eficiente remoção de calor do interior de satélites ou de estações espaciais para o espaço exterior. Sua função é essencial para o controle térmico de qualquer veículo espacial. Em geral, o radiador consiste de pinturas ou faixas especiais com propriedades termo-

ópticas específicas. A área e o número de radiadores podem variar, dependendo do tamanho da estrutura espacial, entretanto o radiador pode ser projetado independente da estrutura espacial, como na ISS (*International Space Station*).

No problema tratado aqui, o projetista busca por uma combinação adequada de pintura ou faixas adesivas para cobrir diferentes áreas do radiador, de forma que a emissividade efetiva obtida deste arranjo deve estar próxima da emissividade de projeto, ou seja, deseja-se determinar a solução óptima do seguinte funcional (Souza e Ramos, 2000):

$$J(E) = \left[\varepsilon_{proj} - \left(\frac{\sum_{i=1}^N \varepsilon_i A_i}{\sum_{i=1}^N A_i} \right) \right]^2, \text{ com a restrição: } \sum_{i=1}^N A_i = A_R, \quad (37)$$

onde ε_{proj} é a emissividade de projeto e $E = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_N]^T$. No artigo de Souza e Ramos (2000) dois métodos são usados para minimizar o funcional (37): *simulated annealing* (SA) e otimização extrema (EO: *extreme optimization*). A solução encontrada é mostrada na figura 7.

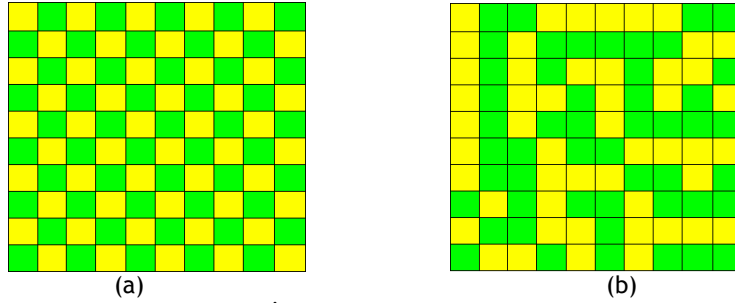


Fig. 7: (a) configuração de referência; (b) configuração obtida para $\min J(E)$ por SA.

3.2.2. Detecção de danos em estruturas espaciais

O problema de detecção de danos é o problema inverso mais importante em teoria de vibrações mecânicas e, por isso, tem sido extensivamente estudado. O uso da técnica variacional nesse tipo de problema é muito recente. Os primeiros trabalhos foram desenvolvidos por Huang (2001; 2002) e Chiwiacowsky & Campos Velho (2003).

Os passos para aplicação da técnica variacional no problema de detecção de danos são os seguintes:

1. Escolher uma configuração inicial para a matriz de rigidez K^0 ;
2. Determine as deformações do sistema (problema direto) e a função objetivo $J(K^0)$;
2. Resolva a equação de sensibilidade:

$$\begin{aligned} \text{modelo direto: } M\ddot{X} + C\dot{X} + KX &= f(t) \\ \text{equação sensibilidade: } M\Delta\ddot{X}_j + C\Delta\dot{X}_j + K\Delta X_j &= \Delta K_j X(t) \end{aligned}$$
3. Resolver a equação adjunta para calcular o gradiente da função objetivo:

$$M\dot{\lambda}(t) - C\lambda(t) + K\lambda(t) = 2[X^{Exp} - X^{Mod}(K)] \quad \text{com a condição final: } \dot{\lambda}(t_f) = \lambda(t_f) = 0,$$

$\lambda(t)$ é o vetor dos multiplicadores de Lagrange;
4. Determine a direção de busca do MGC: $p^m = \begin{cases} -J(K^m) & \text{para } m = 0 \\ -J(K^m) + \gamma^m p^{m-1} & \text{para } m > 0 \end{cases}$,

onde: $\gamma^m = \frac{\|\nabla J(K^m)\|^2}{\|\nabla J(K^{m-1})\|^2}$, sendo: $\nabla_j J(K) = \int_0^{t_f} \lambda^T \Delta K_j X dt$; e $\Delta K_j \equiv \partial[\Delta K]/\partial K_j$
5. Determina o tamanho do passo: $\beta^m = \frac{\int_{t=0}^{t_f} [\Delta X^T (X^{Exp} - X^{Mod})] dt}{\int_{t=0}^{t_f} [\Delta X^T \Delta X] dt}$
6. Atualiza a matriz de rigidez: $K^{m+1} = K^m + \beta^m p^m$;
7. Teste de convergência (princípio da discrepância de Morosov).

Embora Huang (2001; 2002) afirme que esta técnica é totalmente adequada para resolver o problema de identificação de danos com matriz de rigidez dependente do tempo, Chiwiacowsky e Campos Velho (2003) mostraram que a técnica **não funciona** para problemas com um número de graus de liberdade maior do que 5 ou quando o sistema não é tão rígido - por exemplo, tomando-se estruturas em alumínio ao invés de aço (Chiwiacowsky, 2005): **com ou sem ruído** nos dados experimentais. O problema é que, partindo da configuração não danificada, o método variacional a *per se* não consegue calcular uma boa solução inversa. Desta forma, torna-se necessário determinar uma estimativa inicial mais adequada para o passo-1 do algoritmo anterior. Desta forma, Chiwiacowsky e Campos Velho (2003) propuseram uma **técnica híbrida**, onde o algoritmo genético com epidemia é usado para determinar uma configuração inicial para o processo variacional.

A técnica híbrida foi testada com sucesso em vários sistemas (massa-mola, treliça, vigas) e usando diferentes dados experimentais (para domínio do tempo e domínio da frequência) com vários níveis de ruído nestes dados. Aqui vamos aplicar a técnica num exemplo de detecção de danos na estação espacial internacional. A Figura 8 mostra uma foto da ISS e sua representação como uma estrutura com 68 graus de liberdade. Na Figura 8c, os elementos em vermelho são os elementos danificados da estrutura.

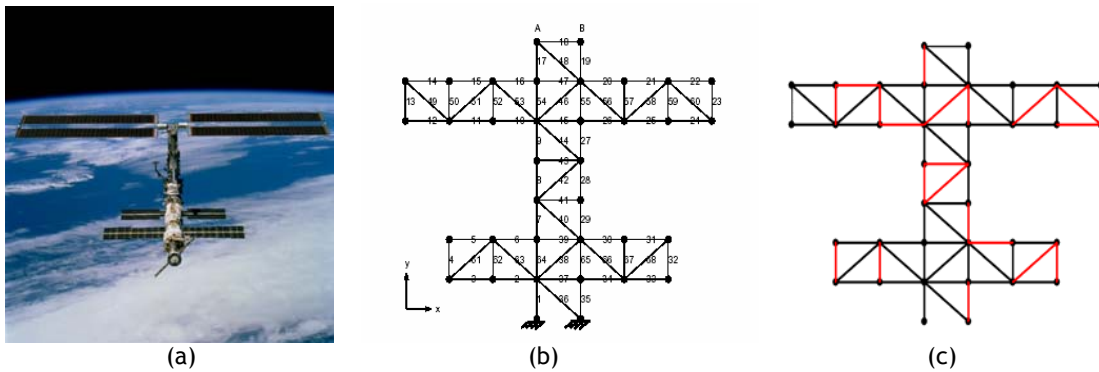


Fig. 8: (a) foto da ISS; (b) representação da estrutura da ISS, (c) danos (vermelho) na ISS.

A Figura 9 mostra a rigidez identificada para cada elemento da representação da ISS através do método híbrido: EGA + variacional (Campos Velho e Chiwiacowsky, 2005; Chiwiacowsky, 2005; Gasbarri et al., 2007). A identificação da rigidez é perfeita no caso de dados sem ruído (Figura 9a) e produz um excelente resultado no caso de dados com ruído (Figura 9b). É importante ressaltar que, no caso de dados com ruído, em todos os casos trabalhados a técnica identificou todas as partes danificadas da estrutura, isto é, até agora não temos um caso registrado em que a estrutura tenha um dano não identificado pela técnica híbrida.

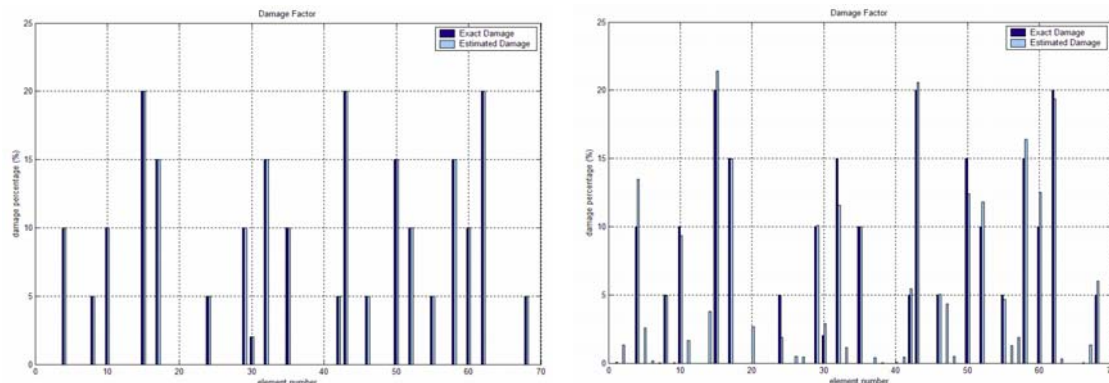


Fig. 9: Identificação da rigidez: (a) dados sem ruído; (b) dados com ruído de 1%.

3.3. Oceanografia: identificação de propriedades óticas em águas naturais

Há duas descrições físicas para a transmissão da luz: (i) da teoria eletromagnética, onde a luz é pensada como uma onda; (ii) da teoria de transporte de partículas sem cargas, se a luz é compreendida como um feixe de partículas (fótons). A segunda formulação é considerada aqui. A equação linear de Boltzmann, também conhecida como a equação de transporte ou equação de transporte radioativo (ETR), para um dado comprimento de onda λ , descreve a distribuição de fótons, expressa por:

$$\mu \frac{dL_\lambda(\zeta, \xi)}{d\zeta} + L_\lambda(\zeta, \xi) = \frac{\overbrace{b}^{\sigma_0}}{a+b} \int_{\Xi} L_\lambda(\zeta, \xi') \beta(\zeta, \xi' \rightarrow \xi) d\Omega + S_\lambda(\zeta, \xi), \quad (38)$$

onde L é a radiância, ζ é a espessura ótica, a e b são os coeficientes de absorção e espalhamento, respectivamente (a razão $b/(a+b)$ é o albedo σ_0), S é o termo de fonte ou sumidouro e ξ é a direção do fóton. O problema físico é esboçado na Figura 10.

Nos estudos de ótica hidrológica inversa, três métodos para resolver o problema direto foram usados e estão disponíveis no Laboratório Associado de Computação e Matemática Aplicada (LAC) do INPE:

- Inserção invariante: Este método é usado no código computacional Hydrolight 3.0, cuja descrição é apresentada por Mobley (1994). Nesta formulação, o problema linear de 2 pontos de contorno é transformado num problema de valor inicial: a equação matricial de Ricatti, resolvida por um integrador numérico bem estabelecido. O software Hydrolight permite a escolha de muitos tipos de fontes internas e diferentes condições de tempo.
- Método S_N Analítico: Primeiramente, a radiância é separada em componentes espalhada e não espalhada. A solução para a componente não espalhada é dada por uma expressão simples, enquanto a radiância espalhada é expandida por soluções elementares das equações de ordenadas discretas. Os coeficientes da expansão são obtidos resolvendo o sistema de equações algébricas resultante. Uma apresentação completa do método é feita por Chalhoub and Garcia (2000).
- Método LTS_N : Este esquema surgiu no início dos anos 90 no contexto de transporte de nêutrons (Barichello e Vilhena, 1993) e foi também aplicado para problemas de transferência radiativa (Segatto e Vilhena, 1994). Uma descrição do método e suas inovações são mencionadas por Segatto e Vilhena (1999). A convergência do método LTS_N foi estabelecida usando a teoria de semigrupo C_0 (Pazos and Vilhena, 1999).

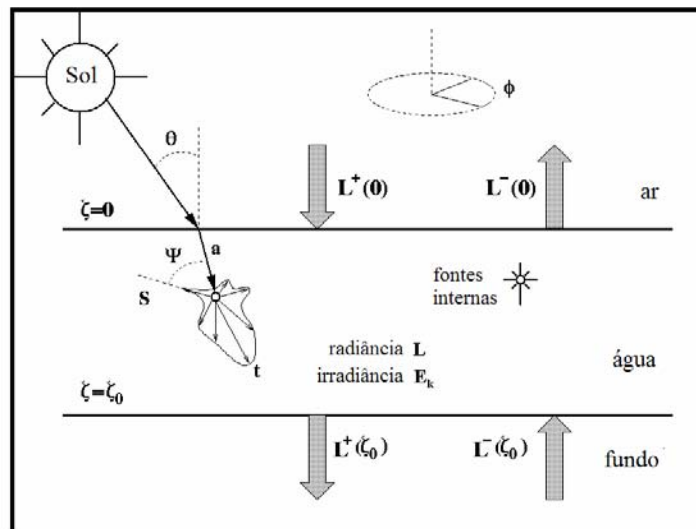


Fig. 10: O problema físico da transmissão de radiação em águas naturais.

Uma comparação entre diferentes métodos de resolução da equação (28) foi realizada por Chalhoub et al. (2003), enquanto o desempenho dos 3 métodos mencionados foi analisado por Souto et al. (2004) num ambiente de computação paralela.

O desenvolvimento de metodologias de inversão para problemas de transferência radiativa e, em particular, ótica hidrológica, é um importante tópico de pesquisa em muitos ramos da ciência e da engenharia (McCormick, 1992; Gordon, 2002). As propriedades de interesse em ótica hidrológica inversa são: termos de fontes internas, propriedades óticas inerentes (POI: função de fase, coeficientes de absorção e espalhamento) e condições de contorno. Reconstruções destas propriedades foram obtidas com dados radiométricos medidos *in situ* e dados de sensoriamento remoto. Vários parâmetros (funções)

associados à ETR tem importância econômica, ecológica e militar. A tabela abaixo ilustra alguns casos resolvidos pelo grupo de problemas inversos do INPE (na tabela, CC significa *condição de contorno*).

Tabela 1: Estimação de propriedades em ótica hidrológica.

Casos	Propriedade	Problema direto	Optimizador	Regularização
(1)	$S(\zeta)$	Ins. Inv.	MQNe (NAG) / GA	-
(2)	$S(\zeta), a, b$	Ins. Inv.	MQNe (NAG)	MaxEnt-0
(3)	$\beta(\xi)$	S_N padrão	MQNe (NAG)	MaxEnt-0
(4)	C.C.	LTS_N	MQNe (NAG)	Tikhonov-0
(5)	$\beta(\xi), \varpi_0$	AS_N	MLM (IMSL)	MLM
(6)	$\beta(\xi), \varpi_0, \tau_0$	AS_N	MLM (IMSL)	MLM
(7)	$S(\zeta)$	AS_N - multi-espectral	MLM (IMSL)	MLM

Um aspecto interessante no Caso-1 é que não foi usado nenhum método de regularização (Stephany et al., 1998), onde o termo de fonte foi aproximado como uma soma de gaussianas com desvio padrão $\sigma = 0,75$. No Caso-2, uma estratégia alternada passo-a-passo foi introduzida, na qual os coeficientes a, b são estimados primeiramente e, após, o termo de fonte é identificado, seguindo-se uma estimacão alternada, até se atingir a convergência. A função de fase foi identificada nos Casos 3 e 4 (Chalhoub et al., 2000; Chalhoub and Campos Velho, 2001), onde a determinação foi possível mesmo com altos níveis de ruído. Estes são casos de estimacão de parâmetros, pois não há dependência entre as propriedades estimadas e a profundidade do oceano. Condições de contorno são identificadas no Caso-5 (Retamoso et al., 2000, 2001; ver também Campos Velho et al., 2002, 2003). Uma estimacão conjunta de função de fase, com alta ordem de espalhamento, albedo e espessura ótica foi realizada no Caso-6 (Chalhoub e Campos Velho, 2002). Finalmente, o Caso-7 representa o primeiro resultado do grupo em ótica hidrológica inversa usando dados de sensoriamento remoto (Chalhoub e Campos Velho, 2003).

3.4. Meteorologia: reconstrução de perfis verticais de temperatura e umidade atmosférica

Uma das conquistas da tecnologia moderna foi a capacidade de melhorar significativamente a previsão do tempo. Uma das etapas essenciais neste processo foi a introdução do uso de satélites meteorológicos, incorporados ao sistema de observação da atmosfera. Na sondagem remota da atmosfera por satélites para identificação de perfis verticais de temperatura e umidade, a atmosfera é aproximada em vários níveis ou camadas de estratificação, nessas camadas, a temperatura e a umidade são consideradas constantes. Usando-se a aproximação do corpo negro para cada camada atmosférica, a energia radiante é transferida de acordo com a equação de Schwarzschild:

$$\tilde{I}_\lambda(\tau_\lambda) = I_\lambda(\tau_\lambda) - I_\lambda(\tau_\lambda^{(s)}) = - \int_{\tau_\lambda^{(s)}}^{\tau_\lambda} B_\lambda(\tau_\lambda) d\tau_\lambda \quad (39)$$

$$B_\lambda(T) = \frac{2h\lambda^3}{c^2 [e^{h\lambda/(KT)} - 1]} \quad (\text{função de Planck})$$

onde λ é comprimento de onda do sensor instalado no satélite, h é a constante de Planck, c é a velocidade da luz no vácuo, k_B é a constante de Boltzmann e T é a temperatura no nível z . Este problema inverso foi resolvido usando 7 canais de satélite para a reconstrução de perfil de temperatura em 4 níveis. O princípio da máxima entropia de 2ª ordem foi usado com regularização (Ramos et al., 1999). Esta metodologia mostrou-se menos sensível em relação ao perfil inicial utilizado do que o código ITPP-5 (Figuras 11a e 11b).

Mesmo obtendo-se boas inversões com métodos de regularização, uma outra abordagem foi executada: a inversão com redes neurais artificiais. São várias as motivações para o estudo de uma nova metodologia: melhorar a inversão do perfil de temperatura, principalmente abaixo do nível de 100 hPa, trabalhar com uma técnica ainda mais rápida. De fato, uma vez treinada, as redes neurais são muito mais eficientes do ponto de vista computacional do que técnicas de dependam da solução recursiva de problemas de otimização. Mais ainda, redes neurais podem ser implementadas em dispositivos de hardware, tornando viável uma inversão quase em tempo real a bordo do satélite. A figura 12a mostra a reconstrução do perfil de temperatura atmosférica por várias técnicas (a melhor estimativa foi obtida com a rede neural - perceptron de multi-camadas) e na figura 12b uma comparação com implementação da rede em VDHL (Shigemori et al., 2007).

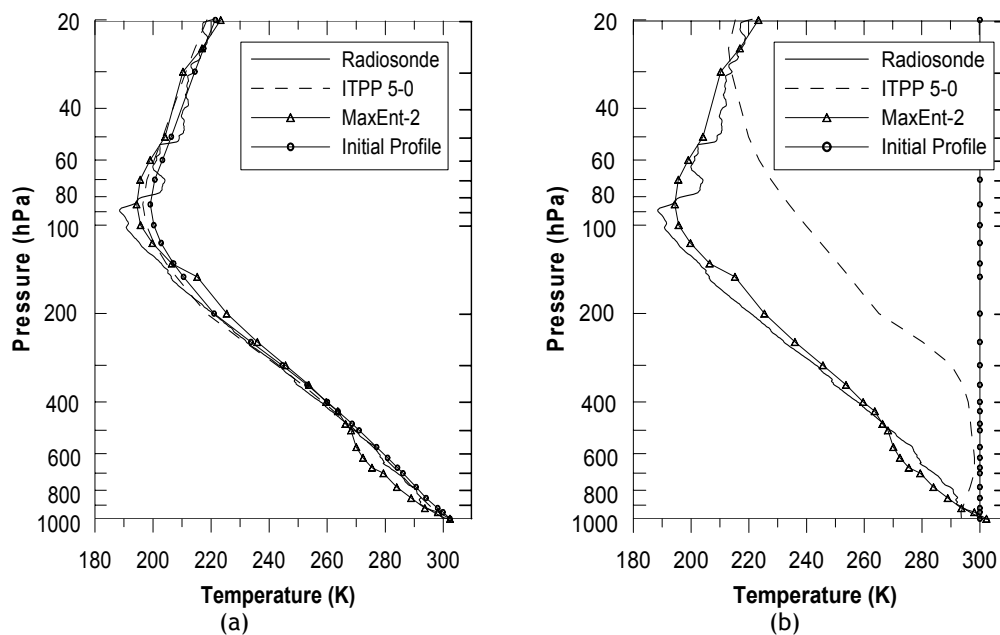


Fig. 11: Perfis de temperatura atmosférica: (a) com perfil inicial de média climatológica, (b) perfil inicial homogêneo com temperatura idêntica a da superfície.

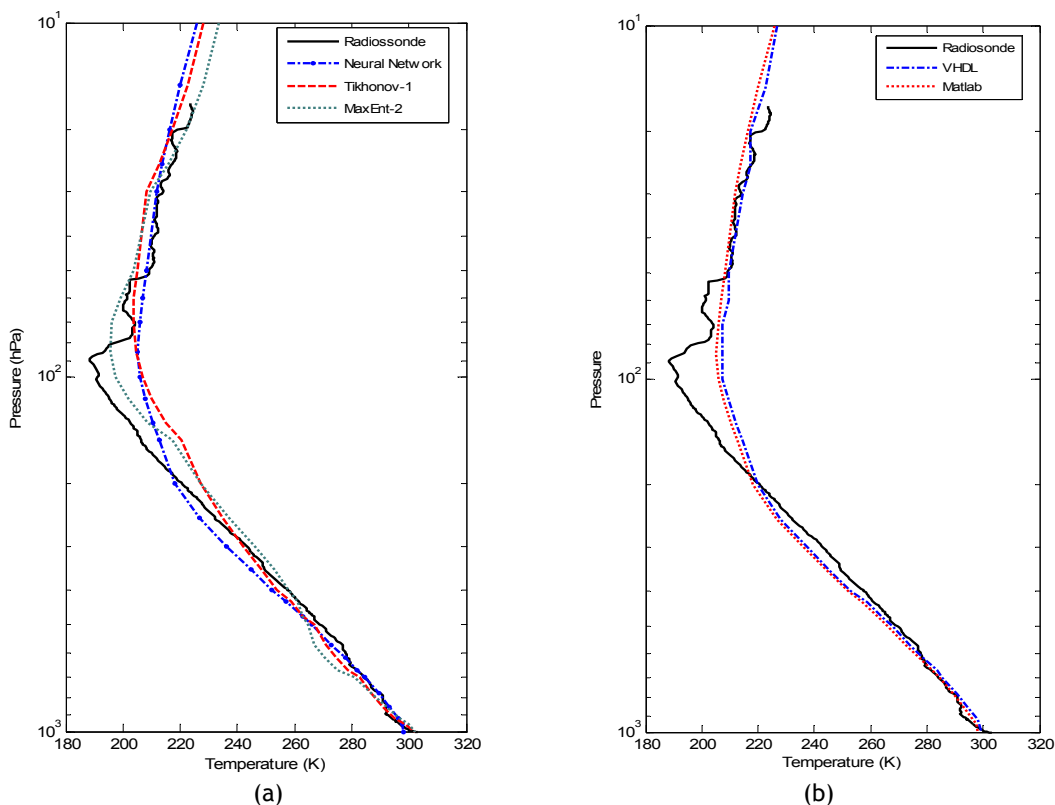


Fig. 12: Reconstrução de perfis de temperatura atmosférica por redes neurais: (a) comparação de várias metodologias, (b) inversão: software versus hardware.

4. Considerações Finais

A área espacial é uma rica fonte para aplicações da metodologia de problemas inversos. Na verdade, muitos desafios em inversão brotaram do desenvolvimento da tecnologia espacial (Alifanov, 1974; Beck et

al., 1985). 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.

É pertinente ressaltar que muito do que o grupo de problemas inversos do LAC-INPE vem desenvolvendo não foi visto aqui. Como destaque de desenvolvimentos significativos pode-se citar o novo algoritmo para otimização extrema (Souza, 2002): 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 (Galski, 2006). O operador de epidemia para GA, foi proposto na tese de mestrado de Medeiros (2005) e seu emprego em problemas inversos foi realizado por Chiwiacowsky e Campos Velho (2003). A versão paralela do EGA foi o tema da tese de mestrado de Sambatti (2004) - ver também (Sambatti et al., 2004; Campos Velho et al., 2006b).

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 (Souto, 2005). 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 (Corrêa, 2005), 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 (Faria, 2007). Um problema relevante nos dias atuais, é a determinação de fluxos superficiais do ciclo bio-geo-químicos de gases de efeito estufa. Roberti (2005) desenvolveu uma metodologia para atacar este problema, maximizando entropia de 2ª ordem, 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.

Do ponto de vista teórico, Campos Velho et al. (2007) 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 está 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 (Shiguemori, 2007). Inversões com redes neurais podem ser aplicadas com outros tipos de sinais de satélites, como sinais de GPS (Cintra et al., 2007). 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.

Bibliografia

- Alifanov, O. M., *Journal of Engineering Physics*, 26: 471-476, 1974.
Barichello, L. B., Vilhena, M. T., *Kernntechnik*, 58: 182-184, 1993.
Beck, J.V., Blackwell, B., St. Clair, C. R. Jr., *Inverse Heat Conduction: Ill-Posed Problems*, John Wiley & Sons, 1985.
Bertero, M. & Boccacci, P., *Introduction to Inverse Problems in Imaging*, Institute of Physics, 1998.
Campos Velho, H. F., Ramos, F. M., Chalhoub, E. S., Stephany, S., Carvalho, J. C., Souza, F. L., *Inverse Problems in Science and Engineering*, 15: 359-372, 2007.
Campos Velho, H. F., Ramos, F. M., Shiguemori, E. H., Carvalho, J. C., *Computational and Applied Mathematics*, 25: 307-330, 2006a.
Campos Velho, H. F., Chiwiacowsky, L. D., Sambatti, S. B. M., *Scientia*, 17: 10-18, 2006b.
Campos Velho, H. F., Chiwiacowsky, L. D., Gasbarri, P. (2005): Some Results on Damage Detection Using Variational Approach, 4o. Congresso Temático de Dinâmica e Controle (DINCON-2005), 6-10 Junho, UNESP, Bauru (SP), Brasil, 1456-1464, 2005.

Campos Velho, H. F., Vilhena, M. T., Retamoso, M. R., R.P. Pazos, R. P., *Progress in Nuclear Energy*, 42: 457-468, 2003.

Campos Velho, H. F., Retamoso, M. R., Vilhena, M. T., *International Journal for Numerical Methods in Engineering*, 54: 1357-1368, 2002.

Campos Velho, H. F., Silva, J. D. S., Nowosad, A. G., Rios Neto, A., *Boletim da SBMAC, Série III* (8 páginas), 2001 - http://www.sbmac.org.br/bol/boletim_2001/default_boletim_01.htm

Campos Velho, H. F., Ramos, F. M., *Brazilian Journal of Geophysics*, 15: 133-143, 1997.

Campos Velho, H. F., Stephany, S., Chalhub, E. S., Ramos, F. M., Retamoso, M. R., Vilhena, M. T., *New approaches on inverse hydrological optics*, *Boletim da SBMAC, Série III* (11 páginas), 2001 - http://www.sbmac.org.br/bol/boletim_2001/default_boletim_01.htm

Chalhub, E. S., Campos Velho, H. F., Vilhena, M. T., Garcia, R. D. M., *Transport Theory and Statistical Physics*, 32:469-499, 2003.

Chalhub, E. S., Campos Velho, H. F., *Applied Numerical Mathematics*, 47: 365-376, 2003.

Chalhub, E. S., Campos Velho, H. F., *J. Quant. Spectrosc. Radiat. Transf.*, 72: 551-565, 2002.

Chalhub, E. S., Campos Velho, H. F., *J. Quant. Spectrosc. Radiat. Transf.*, 69: 137-149, 2001.

Chalhub, E. S., Campos Velho, H. F., Ramos, F.M., J.C.R. Claeysen, J. C. R., *Hybrid Methods in Engineering*, 2: 373-388, 2000.

Chalhub, E. S., Garcia, R. D. M., *J. Quant. Spectrosc. Radiat. Transf.*, 64: 517-535, 2000.

Chiwiacowsky, L. D., *Método Variacional e Algoritmo Genético em Identificação de Danos Estruturais*, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2005.

Chiwiacowsky, L. D., Campos Velho, H. F., Gasbarri, P., *The damage identification problem: a hybrid approach*, 2º Congresso Temático em Dinâmica e Controle (DINCON 2003), 18-22 agosto, São José dos Campos (SP), Brasi, 1393-1402, 2003.

Chiwiacowsky, L. D., Campos Velho, H. F., *Inverse Problem in Science and Engineering*, 11: 471-494, 2003.

Cintra, R. S. C.; Silva, J. D. S. ; Campos Velho, H. F., *Retrieval of humidity profiles with radio occultation measurements using an artificial neural networks*. *Inverse Problems, Design and Optimization Symposium*, Miami, vol. 2, 642-649, 2007.

Corrêa, R. V., *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 em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2005.

Cuco, A. P. C., Silva Neto, A. J., Sousa, F. L., *Inverse Problem in Science and Engineering*, 2008 (aceito).
DBCLSF routine, IMSL MATH LIBRARY, Users Manual. Version 2.0, Houston, IMSL, 1991.

E04UCF routine, NAG Fortran Library Mark 17, Oxford, UK, 1995.

Engl, H. W., Hanke, M., Neubauer, A., *Regularization of Inverse Problems: Mathematics and its Applications*, Kluwer, 1996.

Faria, C., *Uma nova estratégia de otimização de arranjos interferométricos aplicada ao Brazilian Decimetric Array*, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2007.

Galski, R. L., *Desenvolvimento de versões aprimoradas, híbridas, paralela e multiobjetivo do método de otimização extrema generalizada e sua aplicação no projeto de sistemas espaciais*, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2006.

Gasbarri, P., Chiwiacowsky, L. D., Campos Velho, H. F. (2007): *Aerotecnica Missili e Spazio*, 86: 53-69, 2007.

Gordon, H. D., *Oceanologia*, 44: 9-58, 2002.

Groetsche, C. W., *The Theory of Tikhonov Regularization for Fredholm Integral Equation of the 1st Kind*, Pitman, 1984.

Hanisch, R. J., White, R. L., *The Restoration of HST Images and Spectra II*, Space Telescope Science Institute (<http://resources.stsci.edu/>), Baltimore (MA), USA, November 18-19, 1993.

Haykin, S., *Neural Networks: A Comprehensive Foundation*, Macmillan, New York, 1993.

Holland, J. H., *Adaptation in Natural and Artificial Systems*, MIT Press, 1975.

Hu, W., *Nature*, 404: 939-940, 2000.

Huang, C. H., *Int. J. Num. Method. Eng.*, 50: 1545-1558, 2001.

Huang, C.H., *An inverse vibration problem for simultaneously estimating the time-dependent stiffness coefficients*, *Proceedings of the 4th International Conference on Inverse Problems in Engineering*, Rio de Janeiro, Brazil, 2002.

Kasibhatla, P. S. (Editor), *Inverse Methods in Global Biogeochemical Cycles*, Washington: American geophysical Union. 2000.

Kirkpatrick, S. C.D. Gelatt, Jr., M.P. Vecchi, *Science*, 220: 671-680, 1983.

Krishnakumar, K., SPIE 1196, *Intelligent Control and Adaptive Systems*, 1989.

Levenberg K., *Quart. Appl. Math.*, 2:164-8, 1944.

Luenberger, D.G., *Introduction to linear and nonlinear programming*, Addison Wesley, 1973.

Özisik, M. N., *Heat Conduction*, Wiley Interscience, 1980.

F. L. L. Medeiros, F. L. L., Algoritmo genético híbrido como um método de busca de estados estacionários de sistemas dinâmicos, Tese de Mestrado, Computação Aplicada - INPE, São José dos Campos (SP), Brasil, 2005.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E., *J. Chem. Phys.*, 21 : 1087-1092, 1953.

Marquardt D., *J. Soc. Ind. Appl. Math.*, 11:431-41, 1963.

McCormick, N. J., *Nuclear Science and Engineering*, 112: 185-198, 1992.

McCormick, N. M., TEMA: Tendências em Matemática Aplicada e Computacional, 2: 1-12, 2001.

Mobley, C. D., *Light and Water: Radiative Transfer in Natural Waters*, Academic Press, San Diego, USA, 1994.

Muniz, W. B., Ramos, F. M., Campos Velho, H. F., *Computers & Mathematics with Applications*, 40(8/9): 1071-1084, 2000.

Muniz, W. B., Campos Velho, H. F., Ramos, F. M., *Journal of Computational and Applied Mathematics*, 103: 145-163, 1999.

Pazos, R. P., Vilhena, M. T., *Applied Numerical Mathematics*, 30: 79-92, 1999.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 1992.

Ramos, F. M., Campos Velho, H. F., Carvalho, J. C., Ferreira, N. J., *Inverse Problems*, 15: 1139-1148, 1999.

Ramos, F. M., Giovannini, A., *International Journal of Heat and Mass Transfer*, 38: 1139-1148, 1995.

Retamoso, M. R., Vilhena, M. T., Campos Velho, H. F., Ramos, F. M., *Applied Numerical Mathematics*, 40: 87-100, 2001.

Retamoso, M. R., Vilhena, M. T., Campos Velho, H. F., Ramos, F. M., TEMA: Tendências em Matemática Aplicada e Computacional, 1 (parte 2): 455-462, 2000.

Roberti, D. R., Problemas inversos em física da atmosfera, Tese de doutoramento em física, UFSM, Santa Maria (RS), 2005.

Sambatti, S. M. B., Diferentes estratégias de paralelização de um algoritmo genético epidêmico, aplicadas na solução de problemas inversos. Tese de doutoramento em Computação Aplicada, CAP-INPE, São José dos Campos (SP), 2004.

Sambatti, S. M. B., Campos Velho, H. F., Chiwiacowsky, L. D., Stephany, S., Preto, A. J., 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-2004), 26-29 July, Ilha da Madeira, Portugal - paper ICCE504-435 (6 pages), 2004.

Segatto, C. F., Vilhena, M. T., *Annals of Nuclear Energy*, 21: 701-710, 1994.

Shiguemori, E. H., 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 em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2007.

Silva Neto, A. J., Moura Neto, F. D., Escolha de Modelos - Problemas Inversos em Engenharia - Mini-curso, Congresso Nacional de Matemática Aplicada e Computacional, SBMAC, Santos (SP), Brasil, 2000.

Stephany, S., Ramos, F. M., Campos Velho, H. F. & Mobley, C. D., *Computer Modeling and Simulation in Engineering*, 3: 161-165, 1998.

Stephany, S., Campos Velho, H. F., Ramos, F. M., Mobley, C. D., *Journal of Quantitative Spectroscopy & Radiative Transfer*, 67: 113-123, 2000.

Souto, R. P., Recuperação de perfis verticais de propriedades óticas inerentes a partir da radiação emergente da água, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos 2005.

Souto, R. P., Campos Velho, H. F., Stephany, S., Chalhoub, E. S., *Transport Theory and Statistical Physics*, 33: 449-468, 2004.

Sousa, F. L., Otimização extrema generalizada: um novo algoritmo estocástico para projeto ótimo, Tese de doutorado em Computação Aplicada, CAP-INPE, São José dos Campos, Brasil, 2002.

Souza, F. L., Ramos, F. M., *Boletim da SBMAC (Soc. Bras. de Mat. Aplicada e Computacional)*, Série III (10 páginas), 2000 - http://www.sbmac.org.br/bol/boletim_2001/default_boletim_01.htm

Tarantola, A. 1987, *Inverse Problem Theory*, Elsevier.

Tikhonov, A. N., Arsenin, V.Y., *Solution of Ill-posed Problems*, John Wiley & Sons, 1977.

K.A. Woodbury, *What are inverse problems?*, 1995 (<http://www.me.ua.edu/inverse/whatis.html>).

Woodbury, K.A., (www.sbmac.org.br/a_sbmac/publicacoes/frames.htm) *Boletim da SBMAC*, vol. 4, 2000.

Wuensche, C. A., Lubin, P. M., Villela, T., An alternative algorithm for cosmic microwave background radiation sky harmonic analysis, *Astronomical Data Analysis Software and Systems III*, A.S.P. Conference Series, Vol. 61, 1994, Dennis R. Crabtree, R.J. Hanisch, and Jeannette Barned (Eds.), pp. 269-272, 1994.