

APLICAÇÕES DE ALGORITMOS BASEADOS NA DISTÂNCIA DE BREGMAN
PARA A SOLUÇÃO DE PROBLEMAS INVERSOS EM TRANSFERÊNCIA
RADIATIVA

Mariella Janette Berrocal Tito

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS
EM ENGENHARIA NUCLEAR

Aprovada por:

Prof. Nilson Costa Roberty, D.Sc.

Prof. Hécio Rangel Barreto Orlande, Ph. D.

Prof. Jorge Passamani Zubelli, Ph. D.

Prof. Antônio Carlos Marques Alvim, Ph. D.

Prof. Joaquim Teixeira de Assis, D. Sc.

Prof. Fernando Carvalho da Silva, D. Sc.

RIO DE JANEIRO, RJ – BRASIL

JANEIRO DE 2006

TITO, MARIELLA JANETTE BERROCAL

Aplicações de Algoritmos Baseados na
Distância de Bregman para a Solução de
Problemas Inversos em Transferência Radiativa.

[Rio de Janeiro] 2006

IX, 136p. 29,7cm (COPPE/UFRJ, D.Sc.
Engenharia Nuclear, 2006)

Tese – Universidade Federal do Rio de
Janeiro, COPPE

1. Problemas inversos
2. Distância de Bregman
3. Reconstrução de imagens
4. Máxima entropia
5. Regularização de Tikhonov

I. COPPE/UFRJ II. Título (série)

“Eres parte de nuestra historia,
compartistes nuestra vida.

Saber que estás en algún lugar me
ayuda a no sentir tristeza ni soledad”.

A meu pai, Fortunato Berrocal Diaz, (2002)

A minhas amigas Felicitas Tito Huamani e Gloria Molleda Berrocal

AGRADECIMENTOS

Ao professor Nilson Costa Roberty pela orientação e sua amizade.

Aos professores Joaquim Teixeira de Assis, Antônio Márquez Alvim, Fernando Carvalho da Silva, Jorge Passamani Zubelli e Hécio Barreto Orlande pela orientação, sugestões e correções neste trabalho.

Aos professores Chaska Rocio Coripuma e Angel Fortunato B. Callasaca pelo apoio moral e material, muito abrigada por emprestar-me seu “laptop” .

A Raul Carita Montero, meu eterno amigo, pelo crédito outorgado para ficar em Brasil os dois anos que estive sem bolsa.

Aos meus amigos da Vila do Pinheiro, pela alegria compartilhada nas reuniões dos sábados.

Aos professores e funcionários do Programa de Engenharia Nuclear.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.).

APLICAÇÕES DE ALGORITMOS BASEADOS NA DISTÂNCIA DE BREGMAN
PARA A SOLUÇÃO DE PROBLEMAS INVERSOS EM TRANSFERÊNCIA
RADIATIVA

Mariella Janette Berrocal Tito

Janeiro / 2006

Orientador: Nilson Costa Roberty

Programa Engenharia Nuclear

Neste trabalho empregaram-se famílias de funções estritamente convexas, relacionada às entropias de Shannon (1948), Rényi (1961), Varma (1966), Havrda – Charvát (1967), Sharma-Taneja (1975), e Burn, para a construção de famílias de distâncias de Bregman. As distâncias de Bregman foram empregadas no método de máxima entropia para a solução do problema inverso unidimensional e bidimensional, sendo a funcional relacionada à entropia de Sharma – Taneja ($\eta_{r,s}$), a que ofereceu melhor estimativa. Ela foi empregada, no algoritmo *ART* (Algorithm Reconstruction Technique) e no método de Tikhonov para a reconstrução de imagens tomográficas.

Os casos – testes, apresentados para a reconstrução de imagens tomográficas, mostram que o funcional $\eta_{r,s}$, ofereceu melhor resultados ao ser empregado: (i) no método de Tikhonov em comparação com os métodos de mínimos quadrados e Levemberg Marquardt; e (ii) no algoritmo *ART* em comparação com o algoritmo *q-ART*, desenvolvido por Carita et al., 2004.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

APPLICATIONS OF ALGORITHMS BASED IN THE BREGMAN DISTANCE FOR THE SOLUTION OF INVERSE PROBLEMS IN RADIATIVE TRANSFER

Mariella Janette Berrocal Tito

January / 2006

Advisors: Nilson Costa Roberty

Department: Nuclear Engineering

In this work families of strictly convex functions were used to build Bregman distances related to the entropies of Shannon (1948), Rényi (1961), Varma (1966), Havrda - Charvát (1967), Sharma-Taneja (1975), and Burn (Taneja, 2001; Esteban & Morales, 1995). These distances were applied to the generalized maximum entropy method for the solution of radiative transfer inverse problem in one and two-dimensional media. Since the function ($\eta_{r,s}$) related to Sharma -Taneja entropy gives better coefficients estimates, it was used in both *ART* (Algorithm Reconstruction Technique) algorithm and Tikhonov method for the tomography images reconstruction associated problem.

The tests-cases presented show that the functional $\eta_{r,s}$ offered better results in following situations: (i) Tikhonov method in comparison with the least square and Levenberg Marquardt methods, and (ii) *ART* algorithm in comparison with the q-*ART* algorithms, developed by Carita et al., 2004.

ÍNDICE	Página
1. Introdução	1
2. A distância de Bregman	4
2.1 Função de Bregman	4
2.2 Distância de Bregman	9
3. Problema direto	11
3.1 Formulação aproximada do problema de transporte	11
3.2 Solução do problema unidimensional	16
3.3 Aproximação do bidimensional desde o unidimensional	19
3.4 Solução do problema bidimensional	21
3.5 Problema direto para reconstrução de imagens tomográficas	27
3.5.1 Construção da partição (base natural) para feixes paralelos	28
3.5.2 Formulação do problema direto	31
3.5.3 Solução do problema direto	32
4. Problemas inversos	34
4.1 O método de máxima entropia generalizada	36
4.2 Formulação do problema inverso para reconstrução de imagens	38
4.3 Algoritmos <i>ART</i> com a funcional $\eta_{s,r}$	39
4.3.1 Algoritmo solução	43
4.4 Regularização de Tikhonov	44
4.4.1 Algoritmo solução	48
5. Resultados	49
5.1 Problema unidimensional	50
5.1.1 Tipo e localização dos detectores	50
5.1.2 Validação do problema direto em meio unidimensional	51
5.1.3 Estimativa do coeficiente de absorção Método de máxima entropia generalizada, empregando a função $\eta_{s,r}$	52
5.1.3.1 Estimativa do coeficiente de absorção	52
5.1.3.2 Os valores iniciais σ_{ao}	56
5.1.3.3 Algoritmo 1, Meio homogêneo	60
5.1.3.4 Algoritmo 2, Estimativa de σ_a para um meio formado por duas placas	63

5.1.3.5 Algoritmo 3. Meio formado por dois materiais. Estimativa de espessuras	67
5.1.3.6 Estimativa do coeficiente de espalhamento	70
5.1.4 Estimativa do coeficiente de absorção empregando a função $\eta_r(z) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r$ na construção da distância de Bregman.	71
5.1.4.1 Meio homogêneo	71
5.1.5 Estimativa do coeficiente de absorção empregando a função $\eta_B(z) = -\ln z_i$ na construção da distância de Bregman.	73
5.1.5.1 Meio homogêneo	73
5.2 Problema Bidimensional	74
5.2.1 Validação do problema direto para um meio bidimensional	74
5.2.2 Localização dos detectores para a estimativa dos coeficientes de absorção e espalhamento	74
5.2.2.1 Geometria 1	74
5.2.2.2 Geometria 2	75
5.2.3 Estimativa do coeficiente de absorção e espalhamento, empregando $\eta_{1,r}(z_i) = \frac{(z_i - z_i^r)}{1-r}$ na construção da distância de Bregman.	77
5.2.3.1 Meios homogêneos	77
5.2.3.2 Meios heterogêneos	90
5.3 Reconstrução de imagem. Para tomografia de feixes paralelos	93
5.3.1 Reconstrução com o algoritmo <i>ART</i>	93
5.3.1.1 Exemplo 1, meio homogêneo	93
5.3.1.2 Exemplo 2, meio heterogêneo com uma deformação	95
5.3.1.3 Exemplo 3, meio heterogêneo com duas deformações	97
5.3.1.4 Exemplo 4, dados com ruído randômico	98
5.3.1.5- Comparação entre o algoritmo <i>ART</i> construído com o funcional $\eta_{s,r}$ e o algoritmo q- <i>ART</i>	101
5.3.2 Reconstrução com a regularização de Tikhonov	102
5.3.2.1 Exemplo 1, meio homogêneo	104
5.3.2.2 Exemplo 2, meio heterogêneo com uma deformação	107

5.3.2.3 Exemplo 3, meio heterogêneo com duas deformações	111
6. Conclusões e futuros trabalhos	115
7. Referências Bibliográficas	117
Apêndice I Conceitos básicos dos multiplicadores de Lagrange	
Apêndice II Algoritmos para funções não lineares	

CAPÍTULO 1

INTRODUÇÃO

A análise de problemas inversos constitui uma área de pesquisa multidisciplinar, primeiramente com aplicações em engenharia, medicina, geofísica e astrofísica. Entre os temas mais estudados estão os de transferência de calor (MENGUC *et al.*, 1994), (MILANDRI *et al.*, 2002), sensoriamento remoto (AHMAD *et al.*, 2003), estimação de propriedades radiativas de águas naturais (HAKIM *et al.*, 2002) e tomografia (ARRIDGE *et al.*, 1999), (LOPEZ, *et al.*, 1997), (BRANCO, 2000), (ABDALA, 1995), (CARITA MONTERO, 2000-2004), (RIBEIRO, 1994), (REIS, 1990), entre outros.

Existe uma grande quantidade de trabalhos publicados em problemas diretos e inversos em transferência radiativa em meios participantes, i.e. meios absorvedores, emissores e espalhadores. Como boas referências para aplicações de problemas inversos em diferentes áreas da engenharia, MCCORMICK (2000), recomenda os livros de BECK *et al.* (1985), BERTERO e BOCACCI (1991), HENSEL (1991), HERMAN (1980), IYER e HIRAKARA (1993), OZİSIK e ORLANDE (2000), TRUJILLO e BUSBY (1997) e TWOMEY (1996).

Baseados nas entropias generalizadas de Shannon (1948), Rényi (1961), Varma (1966), Havrda – Charvát (1967), Sharma-Taneja (1975), e de Burn (TANEJA, 2001, ESTEBAN & MORALES, 1995), definimos e desenvolvemos três famílias de funções e distâncias de Bregman que foram empregadas na solução de problemas inversos em transferência radiativa descritas a seguir:

1.- Estimativa dos coeficientes de absorção em um meio participante unidimensional. Considera-se meio participante um meio onde existe absorção e espalhamento. O problema direto é modelado pela equação de Boltzmann e resolvida pelos métodos de ordenadas discretas e diferenças finitas. O problema inverso é resolvido como um problema de otimização, ao minimizar a distância de Bregman sujeita à função erro, como a distância de Bregman é construída a partir de funções convexas relacionadas às entropias acima mencionadas, denominamos este método de máxima entropia generalizada, (BERROCAL *et al.* 2000-2005).

2.- Estimativa do coeficiente de absorção e espalhamento em um meio participante bidimensional. Apresentam-se duas geometrias: a primeira geometria é empregada na estimativa do coeficiente de absorção e a segunda na estimativa do coeficiente de espalhamento. O problema direto da primeira geometria foi desenvolvido na tese de mestrado UNI-PERU, BERROCAL (2002), a segunda geometria foi desenvolvido na tese de mestrado na UFRJ, BERROCAL (2001). O problema direto é modelado pela equação de Boltzmann e resolvida pelos métodos de ordenadas discretas e diferenças finitas. O problema inverso foi resolvido com o método de máxima entropia generalizada.

3.- Reconstrução de imagens tomográficas empregando os algoritmos *ART*. Para o problema direto, ao desprezar o espalhamento, a equação de Boltzmann fica reduzida a um sistema de equações lineares. O problema inverso é baseado em resolver um sistema com mais incógnitas que dados. Para a solução empregou-se a base natural apresentada por CARITA *et al.* (2004), e os algoritmos tipo *ART*, desenvolvido com a distância de Bregman construída com a entropia de Sharma e Taneja.

4.- Reconstrução de imagens tomográficas empregando a regularização de Tikhonov, Neste caso, o problema inverso é baseado em resolver um sistema de equações onde tem-se mais dados que incógnitas. Empregou-se para a solução do problema inverso uma malha quadrada que representa o domínio da imagem e a regularização de Tikhonov, desenvolvida com a distância de Bregman construída com a entropia de Sharma e Taneja.

O objetivo principal deste trabalho de pesquisa consiste no desenvolvimento de novos algoritmos para a solução de problemas inversos. Com a finalidade de encontrar os melhores algoritmos, foi adotado como critério o menor erro porcentual acumulado, tempo computacional ou o número de iterações em cada caso-teste nos quatros problemas acima descritos.

No Capítulo 2, apresentam-se as definições da função e da distância de Bregman. Apresentam-se ainda as três famílias de funções (η) com suas respectiva distâncias de Bregman empregadas nesta tese.

No Capítulo 3 é abordada a formulação e a solução matemática do problema direto, para os três casos-teste estudados neste trabalho. Também é apresentado em forma resumida a discretização da malha usada e os métodos utilizados para a solução do problema direto.

No Capítulo 4 são introduzidas a formulação e a solução do problema inverso, começando com a descrição do problema inverso para cada caso teste. Apresenta-se uma a descrição dos algoritmos *ART*, o método de máxima entropia generalizada e a regularização de Tikhonov empregando a distância de Bregman construída com o funcional de Sharma e Taneja.

No Capítulo 5, são abordados os resultados dos casos testes relacionados aos três problemas inversos aqui formulados. Em cada problema inverso são feitas algumas considerações sobre o tipo de detector empregado, a localização dos mesmos e a rotina computacional desenvolvida para a solução dos problemas direto e inverso. A validação da solução do problema direto é feito comparando os resultados obtidos com aqueles disponíveis na literatura para aplicações em transferência de calor por radiação térmica. Apresentam-se também, casos – testes, para a solução dos problemas inversos, empregando dados exatos e dados com ruído como medidas de detectores.

No Capítulo 6 é apresentada a conclusão baseada nos resultados numéricos.

Os primeiros resultados obtidos nesta tese foram obtidos para a geometria unidimensional, utilizando um computador Pentium III com uma versão do Matlab 6.5, logo depois mudou-se para a versão do Matlab 7.0 e o tempo computacional nos programas diminuiu para um sexto do tempo original. Sem embargo o tempo computacional para resolver o problema bidimensional com espalhamento era superior a 3 horas em alguns casos – testes. Ao empregar um Laptop Pentium IV com tecnologia HT (Hyper-Threading Technology), o tempo computacional diminuiu ate 180 vezes o tempo computacional original. Esta é uma das razões pelas quais nos últimos casos testes consideram-se o número de iterações em vez do tempo computacional.

No apêndice 1, apresenta-se conceitos básicos dos multiplicadores de Lagrange, recomenda-se as leituras de Van Tiel (1984) e a tese de Burachik (1995) que descreve explicitamente conceitos e propriedades da função convexa e distância de Bregman.

No apêndice II, apresenta-se um resumo de algoritmos para funções não lineares que foram empregadas no capítulo 5.

CAPÍTULO 2

A DISTÂNCIA DE BREGMAN

2.1 A função de Bregman (η)

As noções de função e distância de Bregman foram introduzidas por Bregman, 1967, no estudo de algoritmos iterativos para o cálculo de pontos fixos de vários tipos de operadores. No entanto, o termo “função de Bregman” foi empregado primeiramente por CENSOR & LENT, 1981. As funções de Bregman têm sido usadas extensivamente em algoritmos de otimização convexa (DE PIERRÔ & IUSEM, 1986; IUSEM, 1995; KIWIEL, 1997), em espaços de Hilbert, (BURACHIK, 1995; BURACHIK & IUSEM, 1998), e em espaços de Banach (BURACHIK & SCHEIMBERG, 2003; BUTNARIU & IUSEM, 1997).

A função de Bregman foi inicialmente definida em espaços de dimensão finita e posteriormente para dimensões infinitas (CENSOR & LENT, 1981); estas funções se caracterizam por ser diferenciáveis e estritamente convexas (RAMOS, 2001)

Conceitos e propriedades relacionados à definição de uma função de Bregman são apresentados em Van Tiel (1984) e a tese de Burachik (1995)

Neste trabalho se empregou três famílias de funções diferenciável e estritamente convexas (η). Estas funções estão relacionadas a seis entropias: entropia de Shannon (SHANNON, 1948), entropia de Havrda – Charvát (HAVRDA & CHARVÁT, 1967), entropia de Sharma e Taneja (SHARMA & TANEJA, 1975), entropia de Rényi (RÉNYI, 1961), entropia de Varma (VARMA, 1966) e a entropia de Burn, (ESTEBAN & MORALES, 1995).

Considerando que $z = \{z_1, z_2, \dots, z_p\}$ é o vetor de todas as incógnitas do sistema e tem valores positivos.

1.- Entropia de Shannon (SHANNON, 1948)

$$H_1(z) = - \sum_{i=1}^p z_i \ln z_i \tag{2.1}$$

é relacionada com a função: (KAPUR & KESAVAN, 1992)

$$\eta_1(z_i) = z_i \ln z_i \quad (2.2)$$

2) A entropia de Havrda – Charvát (HAVRDA & CHARVÁT,1967)

$$H_2(z) = \frac{1}{1-s} \left(\sum_{i=1}^p z_i^s - z_i \right) \quad , \quad s \neq 1, \quad s > 0, \quad (2.3)$$

é relacionada com a função: (KAPUR & KESAVAN, 1992)

$$\eta_2(z_i) = \frac{z_i^s - z_i}{s-1} \quad s \neq 1, \quad s > 0, \quad (2.4)$$

3) A entropia de Sharma e Taneja (SHARMA & TANEJA, 1975)

$$H_3(z) = \frac{1}{r-s} \sum_{i=1}^p (z_i^s - z_i^r), \quad r \neq s, \quad r > 0, \quad s > 0, \quad (2.5)$$

é está relacionada com a função: (KAPUR & KESAVAN, 1992)

$$\eta_3(z_i) = \frac{z_i^s - z_i^r}{s-r} \quad s > 1, \quad r \leq 1, \quad \text{ou } s < 1, \quad r \geq 1 \quad (2.6)$$

Observações:

1. A entropia de Havrda - Charvát pode ser obtida da entropia de Sharma e Taneja quando $r = 1$
2. A entropia de Shannon pode ser obtida da entropia de Sharma e Taneja quando $r \rightarrow 1$ e $s \rightarrow 1$.
3. Da função $\eta_3(z_i)$ Eq. (2.6), podem ser obtidas as funções $\eta_1(z_i)$ e $\eta_2(z_i)$.
- 4 Sem embargo os intervalos dados para os parâmetros s e r na Eq.(2.5) não garantam a convexidade da Eq. (2.6). Observe a Fig. 2.1, a Eq. (2.6) avaliada para $s = 0.2$ e $r = 0.1$

Logo, é necessário definir os intervalos dos parâmetros r e s para garantir que a função $\eta_3(z_i)$ seja estritamente convexa e possa ser empregada neste trabalho. Para isso ela tem que satisfazer a seguinte condição:

$$\eta(\lambda \cdot z_i + (1-\lambda) \cdot z_o) < \lambda \cdot \eta(z_i) + (1-\lambda) \cdot \eta(z_o) \quad \text{com } 0 < \lambda < 1 \quad (2.7)$$

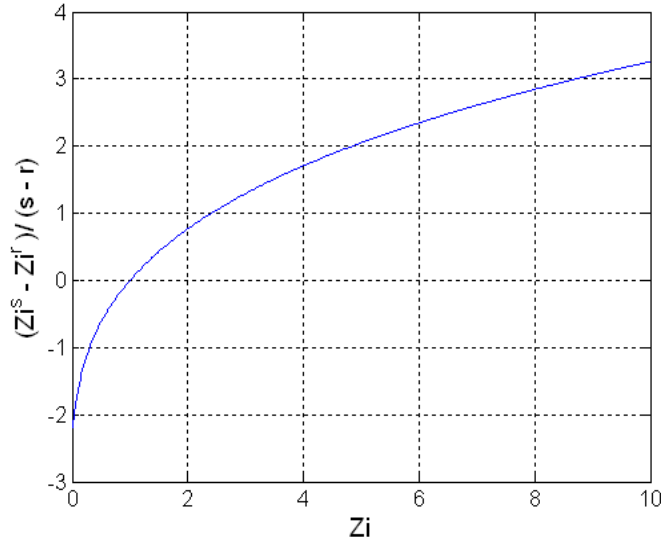


Figura 2.1- A função $\eta_3(z_i)$ avaliada para $s = 0.2$ e $r = 0.1$

Considerando:

- 1.- $z = \{z_1, z_2, \dots, z_p\}$ um vetor formado por todas as incógnitas do sistema,
- 2.- $z \geq 0$,
- 3.- valores de s e r maiores ou iguais a zero,
- 4.- $z_o = 0$

Logo, considerando $z_o = 0$ na Eq. 2.7, tem-se;

$$\eta(\lambda z_i) < \lambda \eta(z_i) \quad (2.8)$$

empregando a Eq. 2.6, a seguinte relação é obtida

$$\frac{1}{s-r} [(\lambda z_i)^s - (\lambda z_i)^r] < \frac{1}{s-r} [\lambda(z_i^s - z_i^r)] \quad (2.9)$$

Para $s - r > 0$, tem-se

$$(\lambda z_i)^s - (\lambda z_i)^r < \lambda(z_i^s - z_i^r) \quad (2.10)$$

$$z_i^s (\lambda^s - \lambda^r z_i^{r-s}) < z_i^s (\lambda - \lambda z_i^{r-s}) \quad (2.11)$$

Da relação da Eq. (2.11), considerando $z_i \geq 0$ tem-se: $\lambda^s < \lambda$ e $\lambda^r z_i^{r-s} \geq \lambda z_i^{r-s}$, logo:

Se $0 < \lambda < 1$, então os valores de r e s que fazem $\eta_3(z_i)$ uma função convexa estão no intervalo: $s > 1$ com $r \in (0,1]$.

Para $s - r < 0$, da Eq. 2.9, tem-se os valores de r e s $r > 1$, com $s \in (0,1]$.

Portanto famílias de funções convexas são definidas como:

$$\eta_{s,r}(z_i) = \frac{z_i^s - z_i^r}{s - r} \text{ com } s > 1 \text{ e } r \in (0,1] \text{ ou } r > 1 \text{ e } s \in (0,1], \quad (2.12)$$

Observe que,

a.- s e r têm o mesmo comportamento na função $\eta_{sr}(z_i)$

b.- Se $s \rightarrow 1$ e $r \rightarrow 1$ obtemos $\eta_1(z_i) = z_i \ln z_i$ dada na Eq. (2.4)

c.- Se $s \rightarrow 1$ e $r > 0$ obtemos $\eta_r(z_i) = \frac{1}{r-1}(z_i - z_i^r)$

4) A entropia de Rényi (RÉNYI,1961)

$$H_4(p) = \frac{1}{1-r} \ln \sum_{i=1}^p z_i^r, \quad r \neq 1, \quad r > 0 \quad (2.13)$$

5) A entropia de Varma (VARMA, 1966)

$$H_5(z) = \frac{1}{m-r} \ln \sum_{i=1}^p z_i^{r-m+1}, \quad m-1 < r < m, \quad m \geq 1 \quad (2.14)$$

Seja r^*

$$r^* = r - m + 1. \quad (2.15)$$

substituindo r^* na Eq. (2.15) obtemos

$$H_5(z) = \frac{1}{1-r^*} \ln \sum_{i=1}^p z_i^{r^*} \quad \text{com,} \quad 0 < r^* < 1, \quad (2.16)$$

As entropias de Varma e de Rényi são iguais, a diferença está em como se define o intervalo do parâmetro r .

Seja a função $\eta_s(z_i)$ igual à entropia negativa de Varma ou de Rényi:

$$\eta_s(z_i) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r \quad (2.17)$$

Observe que, para $r = 1.25$ a função $\eta_s(z_i)$ é côncava e para $r = 0.25$ ela é convexa, (vide figura 2.2). Fazendo uma análise de convexidade determinamos que $\eta_s(z_i)$ é estritamente convexa se $0 < r < 1$. Portanto, definimos como uma função estritamente convexa:

$$\eta_r(z_i) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r \quad \text{com} \quad 0 < r < 1 \quad (2.18)$$

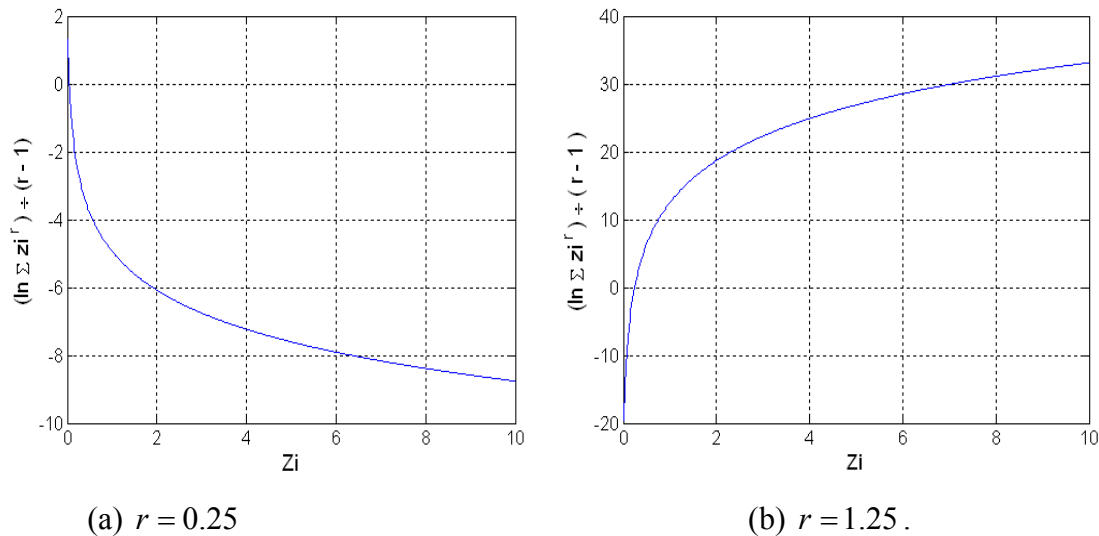


Figura 2.1- Função dada pela Eq. (2.17)

6) A entropia de Burn

$$H_6(z) = \sum_{i=1}^p \ln z_i \quad (2.19)$$

Seja a função $\eta_B(z_i)$ uma função estritamente convexa dada por:

$$\eta_B(z_i) = -\ln z_i \quad (2.20)$$

As funções convexas a ser empregadas, nesta tese, para construir distâncias de Bregman são:

$$1) \quad \eta_{s,r}(z_i) = \frac{z_i^s - z_i^r}{s - r} \quad \text{com } s > 1 \text{ e } r \in (0,1], \text{ ou } r > 1 \text{ e } s \in (0,1], r = 1 \text{ e } s = 1, \quad (2.21)$$

$$2) \quad \eta_r(z_i) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r \quad 0 < r < 1 \quad (2.22)$$

$$3) \quad \eta_B(z_i) = -\ln z_i \quad (2.23)$$

2.2 A distância de Bregman (D)

A distância de Bregman (BREGMAN, 1967) é definida como:

$$D(z, z_o) = D(\eta(z), \eta(z_o)) = \eta(z) - \eta(z_o) - \langle \nabla \eta(z_o), z - z_o \rangle \quad (2.24)$$

onde ∇ é o gradiente. Neste trabalho z e z_o são vetores maiores e iguais a zero.

$$\langle \nabla \eta(z_o), z - z_o \rangle = \left. \frac{\partial \eta}{\partial z} \right|_{z_o} \cdot (z - z_o) \quad (2.25)$$

A distância de Bregman está relacionada com o termo de resto (T_r) da série de Taylor truncada no primeiro termo. Observe que;

$$\eta(z_o) = \eta(z) - \left. \frac{\partial \eta}{\partial z} \right|_{z_o} \cdot (z - z_o) + T_r \quad (2.26)$$

Logo, $D(\eta(z), \eta(z_o)) = T_r$

As distâncias de Bregman para cada função definida pelas Eqs. (2.21)-(2.23) são apresentadas a seguir:

$$1) \quad \eta_{s,r}(z_i) = \frac{z_i^s - z_i^r}{s - r}$$

tem-se

$$\frac{\partial \eta_{s,r}}{\partial z_i} = \frac{1}{s-r} (s \cdot z_i^{s-1} - r z_i^{r-1}) \quad (2.27)$$

e

$$D_{s,r}(z, z_o) = \frac{1}{s-r} \left[\sum_{i=1}^p (z_i^s - z_i^r) - \sum_{i=1}^p (z_{oi}^s - z_{oi}^r) - \sum_{i=1}^p (s \cdot z_{oi}^{s-1} - r z_{oi}^{r-1})(z_i - z_{oi}) \right] \quad (2.28)$$

e no caso em que $s \rightarrow 1$ e $r \rightarrow 1$

$$D_{1,1}(z, z_o) = - \sum_{i=1}^p (z_i - z_{oi} - z_i \ln \frac{z_i}{z_{oi}}) \quad (2.29)$$

$$2. \eta_r(z_i) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r$$

tem-se

$$\frac{\partial \eta_r}{\partial z_i} = \left(\frac{1}{r-1} \right) \frac{r z_i^{r-1}}{\sum_{j=1}^i z_j^r} \quad (2.30)$$

e

$$D_r(z, z_o) = \frac{1}{r-1} \left[\ln \sum_{i=1}^p z_i^r - \ln \sum_{i=1}^p z_{oi}^r - \sum_{i=1}^p r z_{oi}^{r-1} (z_i - z_{oi}) \frac{1}{\sum_{j=1}^i z_j^r} \right] \quad (2.31)$$

$$3) \eta_B(z_i) = - \ln z_i$$

tem-se

$$\frac{\partial \eta_t}{\partial z_i} = - \frac{1}{z_i} \quad (2.32)$$

e

$$D_B(z, z_o) = - \sum_{i=1}^p \ln z_i + \sum_{i=1}^p \ln z_{oi} + \sum_{i=1}^p z_{oi}^{-1} (z_i - z_{oi}) \quad (2.33)$$

CAPÍTULO 3

PROBLEMAS DIRETOS

Considere o problema de transporte de partículas não carregadas, ou de radiação eletromagnética, no interior de um meio participante. Quando a geometria, as propriedades do meio tais como o coeficiente de absorção e o coeficiente de espalhamento, a intensidade da fonte de radiação distribuída no interior do meio e as condições de contorno são conhecidas, pode ser calculado o fluxo angular de partículas ou a intensidade de radiação em qualquer ponto do domínio. Este problema é conhecido como *problema direto* de transferência radiativa .

Quando uma ou mais propriedades e/ou a fonte distribuída não são conhecidas, e medidas experimentais da intensidade da radiação ou fluxo angular de partículas neutras obtidas com detectores posicionados na fronteira do meio estão disponíveis, pode-se tentar estimar as grandezas desconhecidas. Este problema é conhecido como *problema inverso*.

Neste capítulo apresentam-se os problemas diretos de transferência radiativa de 3 casos testes estudados neste trabalho:

1. Estimativa dos coeficientes de absorção e espalhamento em um meio unidimensional
2. Estimativa dos coeficientes de absorção e espalhamento em um meio bidimensional
3. A reconstrução de imagem para tomógrafos de feixes paralelos.

3.1. Formulação aproximada do problema de transporte

A formulação matemática geral adequada para problema de transferência radiativa ou de transporte de partículas não carregadas é a equação de Boltzmann (DUDERSTADT, 1979),

$$\frac{1}{v} \frac{\partial I(r, E, \Omega, t)}{\partial t} + \Omega \cdot \nabla I(r, E, \Omega, t) + \sigma_t(r, E) I(r, E, \Omega, t) = \int_{S^2} \int_0^\infty \sigma_s(r, E' \rightarrow E, \Omega' \rightarrow \Omega) I(r, E', \Omega', t) d\Omega' dE' + S(r, E, \Omega, t) \quad (3.1)$$

onde:

$I(r, E, \Omega, t)$ é a intensidade da radiação, ou o fluxo angular de partículas neutras,

$\sigma_a(r, E)$ é o coeficiente de absorção,

$\sigma_s(r, E' \rightarrow E, \Omega' \rightarrow \Omega)$, $\sigma_s(r, E)$ é o coeficiente de espalhamento,

$\sigma_t(r, E) = \sigma_a(r, E) + \sigma_s(r, E)$ é o coeficiente de extinção total,

$S(r, E, \Omega, t)$ é a fonte distribuída no interior do meio (emissão de radiação térmica devido à temperatura ou fonte de nêutrons resultante de fissão),

r é a variável espacial,

Ω é a direção angular de movimento da radiação,

t é o tempo,

S^2 é o conjunto das direções Ω' em um meio tridimensional,

E é a energia da radiação e

v é a magnitude da velocidade característica da radiação.

Em coordenadas cartesianas temos

$$r = x \cdot \hat{e}_x + y \cdot \hat{e}_y + z \cdot \hat{e}_z \quad \text{e} \quad \Omega = \mu \cdot \hat{e}_x + \xi \cdot \hat{e}_y + \eta \cdot \hat{e}_z.$$

Em transferência radiativa para aplicações em meios tridimensionais é usual denotar $S^2 = 4\pi$. Considerando um sistema sem dependência espectral, ou seja, monoenergético, então a dependência na energia (E) de cada função desaparece, o que leva à conhecida formulação da equação de transporte a uma velocidade. Se considerarmos ainda um sistema estacionário, o primeiro termo da esquerda da Eq. (3.1) é nulo, e a dependência temporal de cada função desaparece. Levando em consideração todas as hipóteses aqui mencionadas, a Eq. (3.1) é simplificada resultando em

$$\Omega \cdot \nabla I(r, \Omega) + \sigma_t(r) I(r, \Omega) = \int_{S^2} \sigma_s(r, \Omega' \cdot \Omega) I(r, \Omega') d\Omega' + S(r, \Omega) \quad (3.2)$$

onde $I(r, \Omega)$, recordemos, é a intensidade da radiação monocromática, i.e. sem dependência espectral, ou o fluxo angular de radiação a uma velocidade no ponto r , e na direção Ω , sendo absolutamente contínua na direção Ω e mensurável na direção Ω' . É possível escrever o coeficiente de espalhamento da seguinte maneira

$$\sigma_s(r, \Omega' \cdot \Omega) = \sigma_s(r) \frac{\phi(r, \Omega' \cdot \Omega)}{4\pi} \quad (3.3)$$

onde $\phi(r, \Omega' \cdot \Omega)$ é denominada função de fase, sendo usualmente representada por uma expansão em uma série de polinômios de Legendre. Observe que estamos aqui usando o mesmo símbolo σ_s nos dois lados da Eq.(3.3). Isto foi feito visando simplificar a notação. Em todos os casos considerados neste trabalho foi usada a separação da dependência nas coordenadas espacial e angular, conforme a representação feita na Eq.(3.3). Considera-se então que o coeficiente de espalhamento possui dependência espacial representada por $\sigma_s(r)$ e a sua dependência angular é representada pela função de fase $\phi(r, \Omega' \cdot \Omega)$.

A Eq. (3.2) pode, portanto, ser escrita como

$$\Omega \cdot \nabla I(r, \Omega) + \sigma_t(r)I(r, \Omega) = Q(r, \Omega) \quad (3.4)$$

onde

$$Q(r, \Omega) = \frac{\sigma_s(r)}{4\pi} \int_{S^2} \phi(r, \Omega' \cdot \Omega) I(r, \Omega') d\Omega' + S(r, \Omega) \quad (3.5)$$

representa um termo fonte, onde o primeiro termo na Eq. (3.5) representa a contribuição da radiação vindo da direção Ω' que é espalhada na direção Ω , "in-scattering" e $S(r, \Omega)$ representa uma fonte interna distribuída no meio.

Uma solução para o problema de transporte é obtida aproximando a equação de transporte por diferenças finitas na variável espacial, e o termo integral (in-scattering) sobre ângulos sólidos é aproximado por quadraturas numéricas resolvendo a equação de transporte radiativa ao longo de direções discretas, estendidas sobre o alcance total do ângulo sólido de medido 4π , no caso tridimensional

Pelo método de ordenadas discretas, o domínio angular $\Omega \in S^2$ é discretizado em L_0 direções, isto é

$$\Omega \in \{\Omega_l\} \text{ com } l=1,2,\dots,L_0$$

Podemos aproximar a integral em Ω por uma quadratura numérica que tem a forma genérica

$$\int_{S^2} f(\underline{\Omega}) d\Omega \approx \sum_{l=1}^{L_0} w_l f(\Omega_l) \quad (3.6)$$

onde w_l é o peso da quadratura associado à direção Ω_l , e onde o somatório dos pesos w_l , $l=1,2,\dots,Lo$, é igual a 4π . O termo integral no lado direito da Eq. (3.5) pode ser então aproximado pela quadratura

$$\int_{4\pi} \phi(r, \Omega' \cdot \Omega) I(r, \Omega') d\Omega' \approx \sum_{m=1}^{Lo} w_m I(r, \Omega_m) \phi(r, \Omega_m \cdot \Omega_l) \quad (3.7)$$

com $l=1,2,\dots,Lo$.

Como trabalharemos com a geometria cartesiana retangular, lembremos então que Ω_l pode ser decomposto nas direções sobre os três eixos coordenados

$$\Omega_l = \mu_l \cdot \hat{e}_x + \xi_l \cdot \hat{e}_y + \eta_l \cdot \hat{e}_z \quad \text{com} \quad \mu_l^2 + \xi_l^2 + \eta_l^2 = 1 \quad (3.8)$$

e além disso

$$\underline{r} = x \cdot \hat{e}_x + y \cdot \hat{e}_y + z \cdot \hat{e}_z \quad (3.9)$$

e

$$\nabla = \frac{\partial}{\partial x} \hat{e}_x + \frac{\partial}{\partial y} \hat{e}_y + \frac{\partial}{\partial z} \hat{e}_z \quad (3.10)$$

As Eqs. (3.4) e (3.5) são então aproximadas por

$$\mu_l \frac{\partial I_l}{\partial x} + \xi_l \frac{\partial I_l}{\partial y} + \eta_l \frac{\partial I_l}{\partial z} + \sigma_l I_l = Q_l \quad \text{com} \quad l=1,2,\dots,Lo. \quad (3.11)$$

e

$$Q_l = \sum_{m=1}^{Lo} \frac{\sigma_s}{4\pi} w_m \phi_{lm} I_m + S_l \quad (3.12)$$

O termo ϕ_{lm} que representa a função de fase é escrito (FIVELAND, 1991), na seguinte forma

$$\phi_{lm} = \sum_{n=0}^M a_n \cdot P_n(\theta) \quad (3.13)$$

onde $P_m(\theta)$ são os polinômios de Legendre que dependem dos cossenos diretores, μ_l, ξ_l, η_l , e M é a ordem de expansão. Em transferência radiativa os coeficientes a_n podem ser determinados com a Teoria de MIE (DUDERSTADT, 1979), válida numa região intermediária entre a óptica geométrica, quando o comprimento de onda da radiação é menor que o diâmetro das partículas do meio com as quais ela interage, e a Teoria de Rayleigh válida no caso oposto.

Considerando o caso mais simples de anisotropia, onde o espalhamento é anisotrópico linear, a representação da função de fase em polinômios de Legendre dada pela Eq. (3.13) e por FIVELAND, (1984), é reduzida a

$$\phi_{lm} = 1 + A_1[\mu_l\mu_m + \xi_l\xi_m + \eta_l\eta_m] \quad (3.14)$$

onde a ordem de expansão é $M=1$ e o valor $a_0=1$, A_1 é o fator de assimetria, ou seja, o coeficiente de anisotropia, e varia na faixa $-1 \leq A_1 \leq 1$; onde os valores $A_1 < 0$ denotam que o espalhamento é para trás (backward), o valor $A_1 = 0$ representa o espalhamento isotrópico e os valores $A_1 > 0$ o espalhamento é para frente (forward). Para espalhamento isotrópico, tem-se então

$$\phi_{lm} = 1 \quad (3.15)$$

Associado a cada direção Ω_l tem-se um peso w_l . Os pesos associados representam áreas sobre uma esfera unitária, e a soma total dos mesmos corresponde, portanto à área de uma esfera unitária. É desejável a invariância das ordenadas e a simetria dos pesos a 90 graus, i.e se (μ_l, ξ_l, η_l) representam uma direção, no primeiro oitante, a condição de simetria requer a existência dos pontos $(-\mu_l, \xi_l, \eta_l)$ no segundo oitante, $(-\mu_l, -\xi_l, \eta_l)$ no terceiro oitante, $(\mu_l, -\xi_l, \eta_l)$ no quarto oitante, $(\mu_l, \xi_l, -\eta_l)$ no quinto oitante, $(-\mu_l, \xi_l, -\eta_l)$ no sexto oitante, $(-\mu_l, -\xi_l, -\eta_l)$ no sétimo oitante, e $(\mu_l, -\xi_l, -\eta_l)$ no oitavo oitante. Desta forma a descrição dos pontos de colocação em um oitante, permite a imediata identificação dos pontos de colocação em todos os oitantes.

Os pontos são colocados sobre $N/2$ níveis relativos a cada eixo como ilustra a Figura 3.3, onde N é a ordem da quadratura. As interseções de cada nível com os níveis

relativos aos outros eixos formam os vértices de triângulos sobre a superfície da esfera. O número total de vértices por cada oitante é igual a $N(N+2)/8$, e define a aproximação S_N nos oitantes.

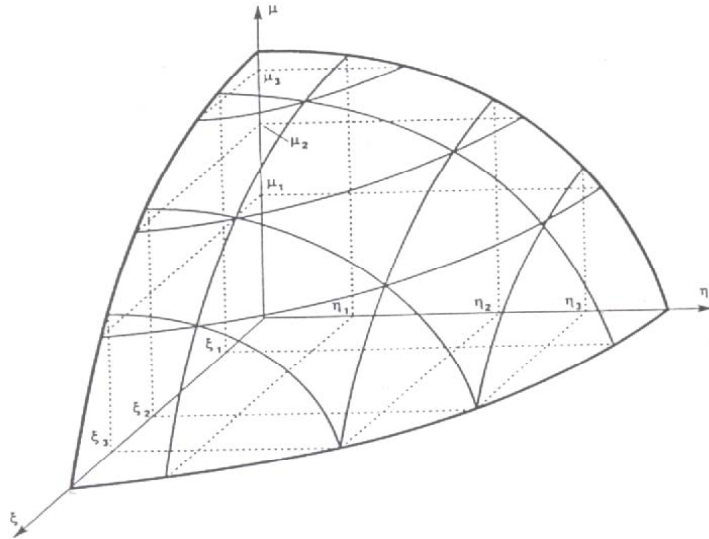


Figura 3.3. Quadratura com nível de simetria S_6 (FIVELAND, 1991),

3.2 Solução do problema unidimensional

A figura 3.4 apresenta um problema físico unidimensional, onde considera-se: uma placa larga de espessura L , uma fonte de intensidade I_b e dois detectores D' e D'' . O detector D' mede o fluxo transmitido e o detector D'' o fluxo refletido pelo meio. O meio é isotrópico, não existem fontes internas, nem qualquer contribuição de fluxos radiativos de outras fontes exteriores que a mencionada.

A Fig. 3.5 apresenta o domínio unidimensional discretizado em i_{\max} intervalos regulares de Δx , L é o comprimento do meio. Produto da discretização do domínio espacial geram-se segmentos espaciais (E_i), onde o sub índice i indica sua posição dentro do meio.

Para a geometria da Fig. 3.4 e da Eq. (3.11) a formulação matemática da equação de transporte unidimensional, em coordenadas cartesianas, em regime estacionário, com simetria azimutal, sem dependência espectral, em um meio isotrópico, sem fontes internas, é dado por::

$$\mu_l \frac{dI_{l(i)}}{dx} + \sigma_{t(i)} I_{l(i)} = Q_{l(i)} \quad Q_{l(i)} = \sum_{m=1}^{L_0} \frac{\sigma_{S(i)}}{4\pi} w_m I_{m(i)} \quad (3.16)$$

para $l = 1, 2, \dots, L_0$; $0 \leq x \leq L$ e $-1 \leq \mu_l \leq 1$

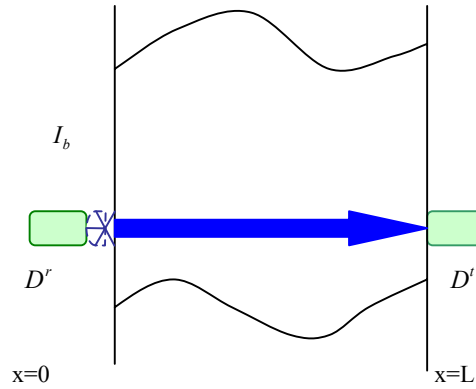


Figura 3.4 – O problema unidimensional

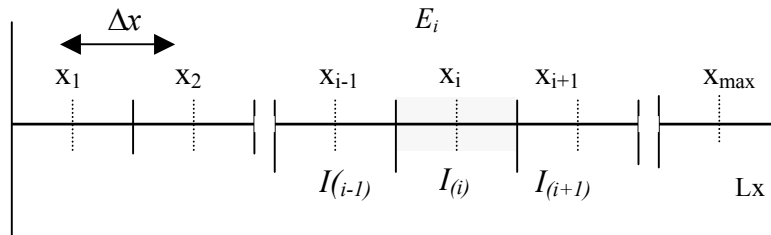


Figura 3.5- Discretização unidimensional

De um procedimento similar ao feito por FIVELAND (1987), para um meio unidimensional, obtemos que a intensidade de radiação pode ser escrita a partir dos fluxos que entram ou saem de suas fronteiras. Na Fig. 3.6 observamos os fluxos que ingressam ou saem pelas fronteiras do segmento $E(i)$, na direção l .

$$\Delta x = x_i - x_{i-1}, \quad \Delta I_{l(i)} = I_{e_{l(i)}} - I_{w_{l(i)}} \quad (3.17)$$

considerando a Eq. (3.17), as derivadas da Eq. (3.16) são aproximadas por diferenças finitas

$$\mu_l \frac{I_{e_{l(i)}} - I_{w_{l(i)}}}{\Delta x} + \sigma_{t(i)} I_{l(i)} = Q_{l(i)} \quad \text{com } l=1,2,\dots,L_0 \quad (3.18)$$

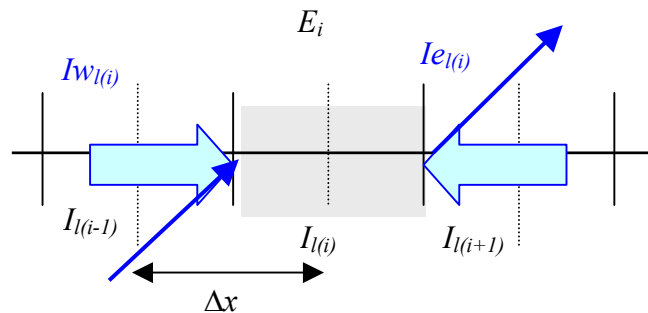


Figura 3.6 Aproximação por diferenças finitas

$I_{l(i)}$ pode ser aproximado (MODEST, 1993), como

$$I_{l(i)} = \gamma Ie_{l(i)} + (1 - \gamma)Iw_{l(i)} \quad (3.19)$$

com $\gamma = 0,5$ (se $I_{l(i)}$ é calculado no médio do elemento segmento

No método de ordenadas discretas o domínio angular é discretizado em Ω_l direções, (com $l = 1, 2, \dots, L_0$), formando um conjunto de quadratura (μ_l , e w_l). Os valores de μ_l descrevem dois quadrantes (vide a Fig. 3.7); o quadrante I com $0 < \mu_l < 1$ e o quadrante II com $-1 \leq \mu_l < 0$, onde o valor de $I_{l(i)}$ vai ser calculado.

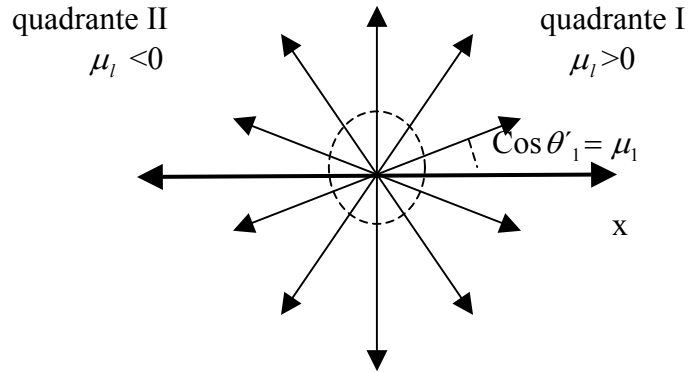


Figura 3.7 Dos quadrantes descritos pelos co-senos diretores μ_l .

QUADRANTE I: $\mu_l > 0$, tem-se que a direção $I_{l(i)}$ esta na direção de Iw , logo da Eq. (3.19), tem-se $Ie_{l(i)} = 2 \cdot I_{l(i)} - Iw_{l(i)}$, inserida na eq. (3.18) dá;

$$I_{l(i)} = \frac{2\mu_l \cdot Iw_{l(i)} + \Delta x \cdot Q_{l(i)}}{2\mu_l + \sigma_{l(i)} \cdot \Delta x} \quad \text{com } l = 1, 2, \dots, L_0 \text{ e } i = 1, 2, \dots, i_{\max} \quad (3.20a)$$

considerando as condições de fronteira entre os elementos de área adjacentes, se $i < i_{\max}$ temos $Iw_{(i+1)} = Ie_{(i)}$

QUADRANTE II: $\mu_l < 0$, tem-se que a direção $I_{l(i)}$ esta na direção de Ie , logo da Eq. (3.18) tem-se $Iw_{l(i)} = 2 \cdot I_{l(i)} - Ie_{l(i)}$ QUE inserida na Eq. (3.17) dá:

$$I_{l(i)} = \frac{-2\mu_l \cdot Iw_{l(i)} + \Delta x \cdot Q_{l(i)}}{-2\mu_l + \sigma_{l(i)} \cdot \Delta x} \quad \text{com } l = 1, 2, \dots, L_0 \text{ e } i = i_{\max}, i_{\max} - 1, \dots, 1 \quad (3.20b)$$

Considerando as condições de fronteira entre os elementos de área adjacentes, se $i < i_{\max}$ temos $I_{e(i-1)} = I_{w(i)}$.

Para determinar $I_{(i)}$ precisamos de $Q_{l(i)}$. Observa-se que $Q_{l(i)}$ depende de $I_{(i)}$, sendo necessário usar um procedimento iterativo para a estimativa de $I_{(i)}$. Seja h um contador de iterações. Então a Eq. (3.20a) para o quadrante I, pode ser escrita como:

$$I_{l(i)}^h = \frac{2\mu_l \cdot I_{w_{l(i)}}^h + \Delta x \cdot Q_{l(i)}^{h-1}}{2\mu_l + \sigma_{l(i)} \cdot \Delta x} \quad (3.21a)$$

$$\text{com } Q_{l(i)}^{h-1} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i)}}{4\pi} w_m \cdot I_{m(i)}^{h-1} \text{ para } l=1,2,\dots,L_o \quad (3.21b)$$

A Tabela 3.1 apresenta a quadratura LSH (Level Symmetric Hybrid) de ordem 10 utilizada neste trabalho.

Tabela 3.1 - Pontos de colocação e pesos da quadratura LSH S_{10} , $N=10$, $L_o=10$.

Ponto Número	Ordenadas μ_l	Pesos w_l	Ponto Número	ordenadas μ_l	Pesos w_l
1	0.1372719	2.0122	6	-0.1372719	2.0122
2	0.5046889	2.1071	7	-0.5046889	2.1071
3	0.7004129	0.5990	8	-0.7004129	0.5990
4	0.8523177	1.1872	9	-0.8523177	1.1872
5	0.9809754	0.3778	10	-0.9809754	0.3778

3.3 Aproximação ao bidimensional desde o unidimensional

Para fazer uma aproximação de um meio unidimensional ao bidimensional considera-se

1. Um sistema formado sempre por uma fonte externa no perímetro do meio com intensidade I_b e dois detectores posicionados segundo a Fig. 3.8. O detector D' medirá o fluxo transmitido e o detector D'' o fluxo por reflexão do meio;
2. Sem fontes internas, nem fluxos externos que ingressem ao meio.

3. O meio é isotrópico
4. O meio bidimensional é discretizado gerando elementos de áreas segundo a Fig. 3.8. Se as áreas sombreadas formam o caminho entre a fonte e o detector, então não se considera qualquer contribuição de fluxos dos elementos de áreas no sombreados (vizinhos) que possam ingressar por espalhamento às áreas sombreadas.

Fazendo uma analogia da Eq. (3.16) para um meio bidimensional temos:

$$\mu_l \frac{dI_{l(i,j)}}{dx} + \sigma_{t(i,j)} I_{l(i,j)} = Q_{l(i,j)} \quad Q_{l(i,j)} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i,j)}}{4\pi} w_m I_{m(i,j)} \quad (3.22)$$

Para a geometria da Fig. 3.8, temos que $Iw_{(i,j)}$ e $Ie_{(i,j)}$ são os único fluxos considerados que ingressam no elemento de área (i,j) que forma a faixa sombreada. Logo, fazendo um desenvolvimento similar às feitas no meio unidimensional temos que as intensidades podem ser aproximadas pelas seguintes equações

Para o quadrante I : $\mu_l > 0$ com $Ie_{l(i,j)} = 2 I_{l(i,j)} - Iw_{l(i,j)}$ temos

$$I_{l(i,j)} = \frac{2\mu_l \cdot Iw_{l(i,j)} + \Delta x \cdot Q_{l(i,j)}}{2\mu_l + \sigma_{t(i,j)} \cdot \Delta x} \quad \text{com } l=1,2,\dots,L_o. \text{ e } i=1,2,\dots,i_{\max}. \quad (3.23)$$

Para o quadrante II : $\mu_l < 0$ com $Iw_{l(i,j)} = 2I_{l(i,j)} - Ie_{l(i,j)}$ temos

$$I_{l(i,j)} = \frac{-2\mu_l \cdot Iw_{l(i,j)} + \Delta x \cdot Q_{l(i,j)}}{-2\mu_l + \sigma_{t(i,j)} \cdot \Delta x} \quad \text{com } l=1,2,\dots,L_o. \text{ e } i=i_{\max}, i_{\max}-1, \dots, 1. \quad (3.24)$$

De forma similar na geometria da Fig. 3.9 tem-se que $In_{(i,j)}$ e $Is_{(i,j)}$ são os únicos fluxos considerados que ingressam no elemento de área (i,j) que forma a faixa sombreada. Fazendo um desenvolvimento similar para a geometria da Fig. 3.8 temos as equações

Para o quadrante I : $\mu_l > 0$ com $In_{l(i,j)} = 2 I_{l(i,j)} - Is_{l(i,j)}$ temos

$$I_{l(i,j)} = \frac{2\mu_l \cdot In_{l(i,j)} + \Delta x \cdot Q_{l(i,j)}}{2\mu_l + \sigma_{t(i,j)} \cdot \Delta x} \quad \text{com } l=1,2,\dots,L_o. \text{ e } i=1,2,\dots,i_{\max}. \quad (3.25)$$

Para o quadrante II : $\mu_l < 0$ com $In_{l(i,j)} = 2I_{l(i,j)} - I_{e_{l(i,j)}}$ temos

$$I_{l(i,j)} = \frac{-2\mu_l \cdot I_{s_{l(i,j)}} + \Delta x \cdot Q_{l(i,j)}}{-2\mu_l + \sigma_{t(i,j)} \cdot \Delta x} \quad \text{com } l=1,2,\dots,L_0 \text{ e } i=i_{\max}, i_{\max}-1, \dots, 1. \quad (3.26)$$

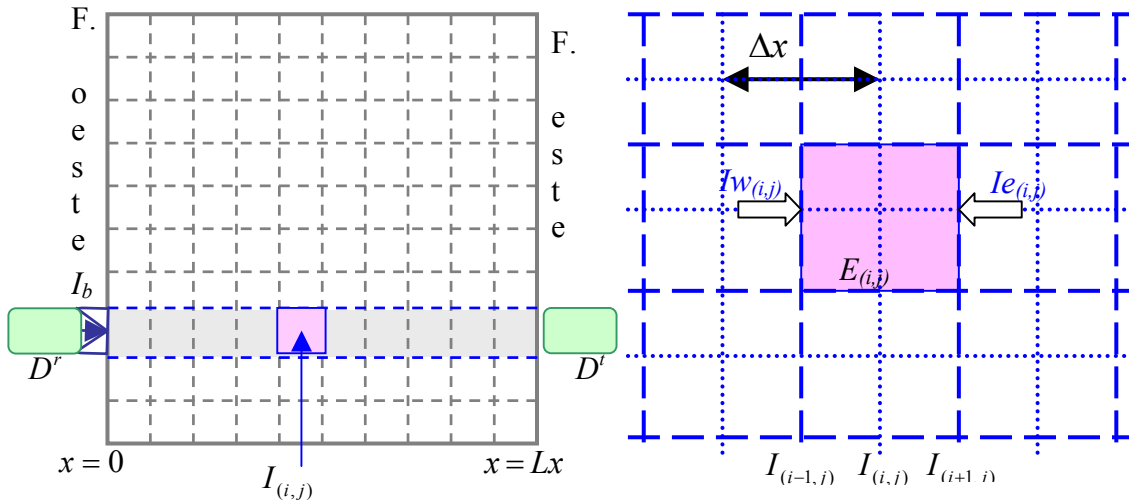


Figura 3.8 Posição da fonte e dos detector. Fluxos I_w e I_e do elemento de área (i,j)

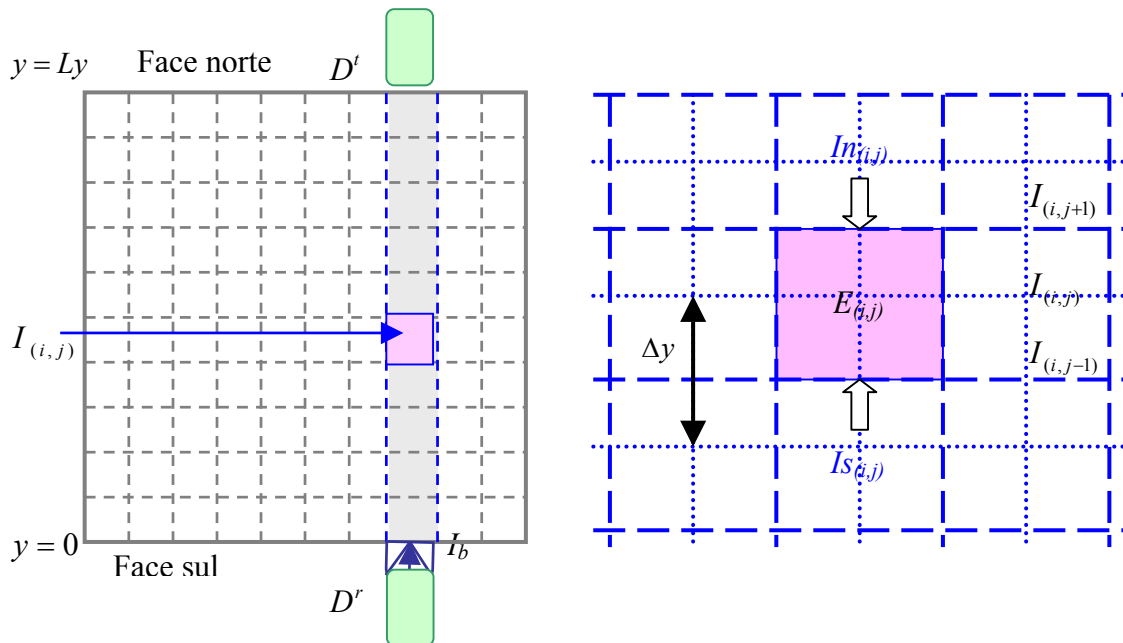


Figura 3.9 Posição da fonte e dos detector. Fluxos I_n e I_s do elemento de área (i,j)

3.4.- Solução do problema bidimensional

Considerando agora a dependência da intensidade da radiação com as coordenadas espaciais x e y e com os ângulos θ e ϕ , obtém-se das Eqs. (3.11).

$$\mu_l \frac{\partial I_l}{\partial x} + \xi_l \frac{\partial I_l}{\partial y} + \sigma_l I_l = Q_l \quad (3.27)$$

com

$$Q_l = \sum_{m=1}^{L_o} \frac{\sigma_s}{4\pi} w_m \phi_{lm} I_m + S_l \quad (3.28)$$

Para $l = 1, 2, \dots, L_o$, $0 \leq x \leq L_x$, $0 \leq y \leq L_y$, $-1 \leq \mu \leq 1$, $-1 \leq \xi \leq 1$

onde L_x e L_y representam os comprimentos ao longo das direções x e y , l representa as direções em que é discretizado o domínio angular, μ_l e ξ_l , são os cossenos diretores com respeito aos eixos cartesianos x e y respectivamente, e w_m são os pesos da quadratura.

O meio é discretizado em uma malha regular com i_{\max} intervalos na direção x e j_{\max} intervalos na direção y , de modo que se obtém uma partição de domínio com $i_{\max} \times j_{\max}$ quadriláteros, denominados elementos $E_{(i,j)}$ (vide Figs. 3.8 e 3.9). A posição de cada elemento de área é identificada pelo índice (i,j) , onde o valor de “ i ” indica o número da coluna, começando à esquerda (face oeste), e o “ j ” indica o número da linha, começando a partir da fronteira inferior do meio (face sul). Os comprimentos das fronteiras do meio são dadas por $L_x = \Delta x \cdot i_{\max}$ e $L_y = \Delta y \cdot j_{\max}$, i.e. face sul. Na Fig.3.10, observa-se que a área de cada elemento é $\Delta A = \Delta x \cdot \Delta y$, os fluxos I_n , I_s , I_w e I_e são aqueles que ingressam ou saem pela face norte, sul, este e oeste, respectivamente, do elemento $E_{(i,j)}$. O espaçamento entre os elementos vizinhos ao longo dos eixos x e y é dado por;

$$\Delta x_i = x_i - x_{i-1}, \quad \Delta y_j = y_j - y_{j-1} \quad (3.29)$$

A variação da intensidade total no centro do elemento de área $E_{(i,j)}$ pode ser escrita como a soma das variações ao longo das direções em x e em y

$$\Delta I_{(i,j)x} = I_{e(i,j)} - I_{w(i,j)} \quad (3.30)$$

$$\Delta I_{(i,j)y} = I_{n(i,j)} - I_{s(i,j)} \quad (3.31)$$

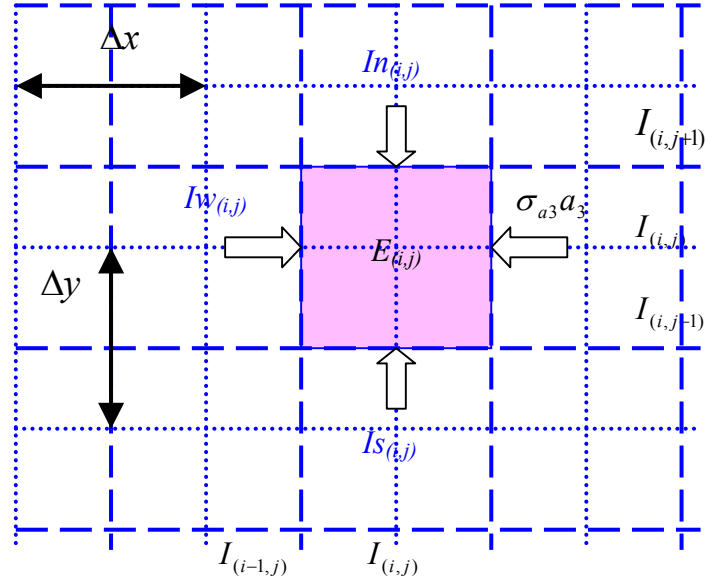


Figura 3.10 Aproximação por diferenças finitas.

Substituindo as derivadas parciais com relação a x e y , por diferenças obtemos

$$\mu_l \frac{\Delta I_{l(i,j)}}{\Delta x} + \xi_l \frac{\Delta I_{l(i,j)}}{\Delta y} + \sigma_t I_{l(i,j)} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i,j)}}{4\pi} w_m \phi_{lm(i,j)} I_{m(i,j)} \quad \text{para } l=1,2,\dots,L_o \quad (3.32)$$

Levando as Eqs. (3.30) e (3.31) na Eq. (3.32) obtém-se

$$\mu_l \frac{Ie_{l(i,j)} - Iw_{l(i,j)}}{\Delta x} + \xi_l \frac{In_{l(i,j)} - Is_{l(i,j)}}{\Delta y} + \sigma_t I_{l(i,j)} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i,j)}}{4\pi} w_m \phi_{lm(i,j)} I_{m(i,j)}$$

para $l=1,2,\dots,L_o$ (3.33)

O valor de $I_{l(i,j)}$ pode ser aproximado mediante o seguinte artifício (MODEST, 1993)

$$I_{l(i,j)} = \gamma In_{l(i,j)} + (1-\gamma) Is_{l(i,j)} \quad I_{l(i,j)} = \gamma Ie_{l(i,j)} + (1-\gamma) Iw_{l(i,j)}$$

onde é usualmente empregado o valor $\gamma = 0,5$.

O domínio angular Ω_l é discretizado em $l=1,2,\dots,L_o$ nas direções, sendo representada pelos cossenos diretores μ_l , $l=1,2,\dots,L_o$, e ξ_l , $l=1,2,\dots,L_o$, respectivamente com relação aos eixos x e y , de modo que

$$\Omega_l = \mu_l \hat{i} + \xi_l \hat{j} \quad (3.34)$$

onde \hat{i} e \hat{j} são vetores unitários nas direções x e y . Os valores de μ_l e ξ_l , estão compreendidos, portanto nas faixas $-1 \leq \mu_l \leq 1$ e $-1 \leq \xi_l \leq 1$; caracterizando quatro quadrantes ao longo dos quais a radiação pode ser propagada (vide Fig. 3.11), ou sofrer espalhamento. Os valores de μ_l e ξ_l , empregados neste trabalho são apresentados na tabela 3.2. Para cada quadrante calcula-se a intensidade de radiação $I_{l(i,j)}$.

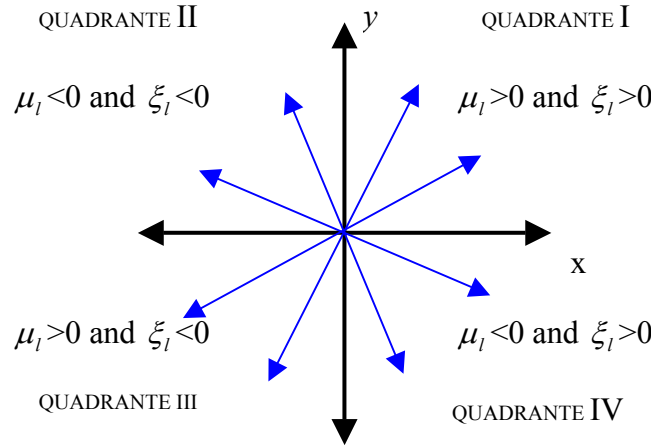


Figura 3.11- O domínio angular e seus 4 quadrantes.

QUADRANTE I: $\mu_l > 0$ e $\xi_l > 0$.

Levando as Eqs. (3.34) na Eq. (3.33), obtém-se

$$I_{l(i,j)} = \frac{\Delta y \cdot \mu_l \cdot Iw_{l(i,j)} + \Delta x \cdot \xi_l \cdot Is_{l(i,j)} + \Delta x \cdot \Delta y \cdot \gamma \cdot Q_{l(i,j)}}{\Delta y \cdot \mu_l + \Delta x \cdot \xi_l + \sigma_{t(i,j)} \cdot \Delta x \cdot \Delta y \cdot \gamma} \quad (3.35)$$

$$\text{com } Q_{l(i,j)} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i,j)}}{4\pi} w_m \phi_{lm(i,j)} I_{m(i,j)} \quad \text{para } l=1,2,\dots,L_o \quad (3.36)$$

desde que $\mu_l > 0$ e $\xi_l > 0$, com $i=1,2,\dots,i_{\max}$, $j=1,2,\dots,j_{\max}$.

Consideremos agora as condições de fronteira entre os elementos de áreas adjacentes:

$$\text{Se } j < j_{\max} \text{ temos } Is_{(i,j+1)} = In_{(i,j)} \text{ e se } i < i_{\max} \text{ temos } Iw_{(i+1,j)} = Ie_{(i,j)} \quad (3.37)$$

QUADRANTE II: $\mu_l < 0$ e $\xi_l > 0$. Levando as Eqs. (3.34) na Eq. (3.33), obtém-se

$$I_{l(i,j)} = \frac{\frac{-\gamma}{(1-\gamma)} \cdot \Delta y \cdot \mu_l \cdot Ie_{l(i,j)} + \Delta x \cdot \xi_l \cdot Is_{l(i,j)} + \Delta x \cdot \Delta y \cdot \gamma \cdot Q_{l(i,j)}}{\frac{-\gamma}{(1-\gamma)} \cdot \Delta y \cdot \mu_l + \Delta x \cdot \xi_l + \sigma_{l(i,j)} \cdot \Delta x \cdot \Delta y \cdot \gamma} \quad (3.38)$$

desde que $\mu_l < 0$ e $\xi_l > 0$, com $i = i_{\max}, i_{\max} - 1, \dots, 1$; $j = 1, 2, \dots, j_{\max}$.

Considerando as condições de fronteira entre os elementos de áreas adjacentes:

$$\text{Se } j < j_{\max} \text{ temos } Is_{(i,j+1)} = In_{(i,j)} \quad \text{e} \quad \text{se } i > 1, \text{ temos } Ie_{(i-1,j)} = Iw_{(i,j)}; \quad (3.39)$$

QUADRANTE III : $\mu_l > 0$ e $\xi_l < 0$. Levando as Eqs. (3.34) na Eq. (3.33), obtém-se

$$I_{l(i,j)} = \frac{\Delta y \cdot \mu_l \cdot Iw_{l(i,j)} - \frac{\gamma}{(1-\gamma)} \cdot \Delta x \cdot \xi_l \cdot In_{l(i,j)} + \Delta x \cdot \Delta y \cdot \gamma \cdot Q_{l(i,j)}}{\frac{-\gamma}{(1-\gamma)} \cdot \Delta y \cdot \mu_l + \Delta x \cdot \xi_l + \sigma_{l(i,j)} \cdot \Delta x \cdot \Delta y \cdot \gamma} \quad (3.40)$$

desde que $\mu_l > 0$ e $\xi_l < 0$, com $j = j_{\max}, j_{\max} - 1, \dots, 1$ e $i = 1, 2, \dots, i_{\max}$.

Considerando as condições de fronteira entre os elementos de áreas adjacentes:

$$\text{Se } j > 1 \text{ temos } In_{(i,j-1)} = Is_{(i,j)} \quad \text{e} \quad \text{se } i < i_{\max} \text{ temos } Iw_{(i+1,j)} = Ie_{(i,j)} \quad (3.41)$$

QUADRANTE IV: $\mu_l < 0$ e $\xi_l < 0$. Levando as Eqs. (3.34) na Eq. (3.33), obtém-se

$$I_{l(i,j)} = \frac{\frac{-\gamma}{(1-\gamma)} \cdot \Delta y \cdot \mu_l \cdot Iw_{l(i,j)} - \frac{\gamma}{(1-\gamma)} \cdot \Delta x \cdot \xi_l \cdot In_{l(i,j)} + \Delta x \cdot \Delta y \cdot \gamma \cdot Q_{l(i,j)}}{\frac{-\gamma}{(1-\gamma)} \cdot \Delta y \cdot \mu_l - \frac{\gamma}{(1-\gamma)} \cdot \Delta x \cdot \xi_l + \sigma_{l(i,j)} \cdot \Delta x \cdot \Delta y \cdot \gamma} \quad (3.42)$$

desde que $\mu_l < 0$ e $\xi_l < 0$, com $i = i_{\max}, i_{\max} - 1, \dots, 1$; $j = j_{\max}, j_{\max} - 1, \dots, 1$.

Considerando as condições de fronteira entre os elementos de áreas adjacentes:

$$\text{Se } j > 1 \text{ temos } In_{(i,j-1)} = Is_{(i,j)} \quad \text{e} \quad \text{se } i > 1 \text{ temos } Ie_{(i-1,j)} = Iw_{(i,j)} \quad (3.43)$$

Para determinar $I_{(i,j)}$ necessitamos de $Q_{l(i,j)}$, no entanto, observamos que $Q_{l(i,j)}$ depende de $I_{(i,j)}$. Pode-se então utilizar um procedimento iterativo para a obtenção de

uma solução aproximada para $I_{(i,j)}$. Seja h um contador de iterações. Então, as Eqs. (3.35) e (3.36), para o quadrante I, podem ser reescritas como

$$I_{l(i,j)}^h = \frac{\Delta y \cdot \mu_l \cdot I_{l(i,j)}^h + \Delta x \cdot \xi_l \cdot I_{l(i,j)}^h + \Delta x \cdot \Delta y \cdot \gamma \cdot Q_{l(i,j)}^{h-1}}{\Delta y \cdot \mu_l + \Delta x \cdot \xi_l + \sigma_{l(i,j)} \cdot \Delta x \cdot \Delta y \cdot \gamma} \quad (3.44)$$

$$Q_{l(i,j)}^{h-1} = \sum_{m=1}^{L_o} \frac{\sigma_{s(i,j)}}{4\pi} w_m \cdot \phi_{lm(i,j)} \cdot I_{m(i,j)}^{h-1} \quad \text{para } l=1,2,..,L_o \quad (3.45)$$

Equações similares das Eqs. (3.44) e (3.45) são obtidas para os quadrantes II, III, IV.

Tabela 3.2 Pontos de colocação e pesos para a quadratura LSH S_{10} , (FIVELAND, 1991)

Ponto	ordenadas		Pesos
Número	μ_l	ξ_l	w_l
1	0.8819171	0.33333333	$\pi/3$
2	-0.8819171	0.33333333	$\pi/3$
3	0.8819171	-0.33333333	$\pi/3$
4	-0.8819171	-0.33333333	$\pi/3$
5	0.33333333	0.8819171	$\pi/3$
6	-0.33333333	0.8819171	$\pi/3$
7	0.33333333	-0.8819171	$\pi/3$
8	-0.33333333	-0.8819171	$\pi/3$
9	0.33333333	0.33333333	$\pi/3$
10	-0.33333333	0.33333333	$\pi/3$
11	0.33333333	-0.33333333	$\pi/3$
12	-0.33333333	-0.33333333	$\pi/3$

3.5.- Problema direto para reconstrução de imagens tomográficas

A tomografia computadorizada é o processo de obter uma imagem bidimensional de uma seção transversal de um corpo (ou tridimensional, se obtivermos varias seções), pelo processamento de uma serie de projeções unidimensionais, obtidas através da radiação. Com ele podemos obter informações sobre a estrutura interna de um corpo sem destruí-lo.

Neste trabalho a reconstrução de imagem é para tomógrafos de feixe paralelo, que é um sistema formado com apenas uma fonte e um detector fazendo movimentos de translação e rotação (Vide Fig. 3.12)

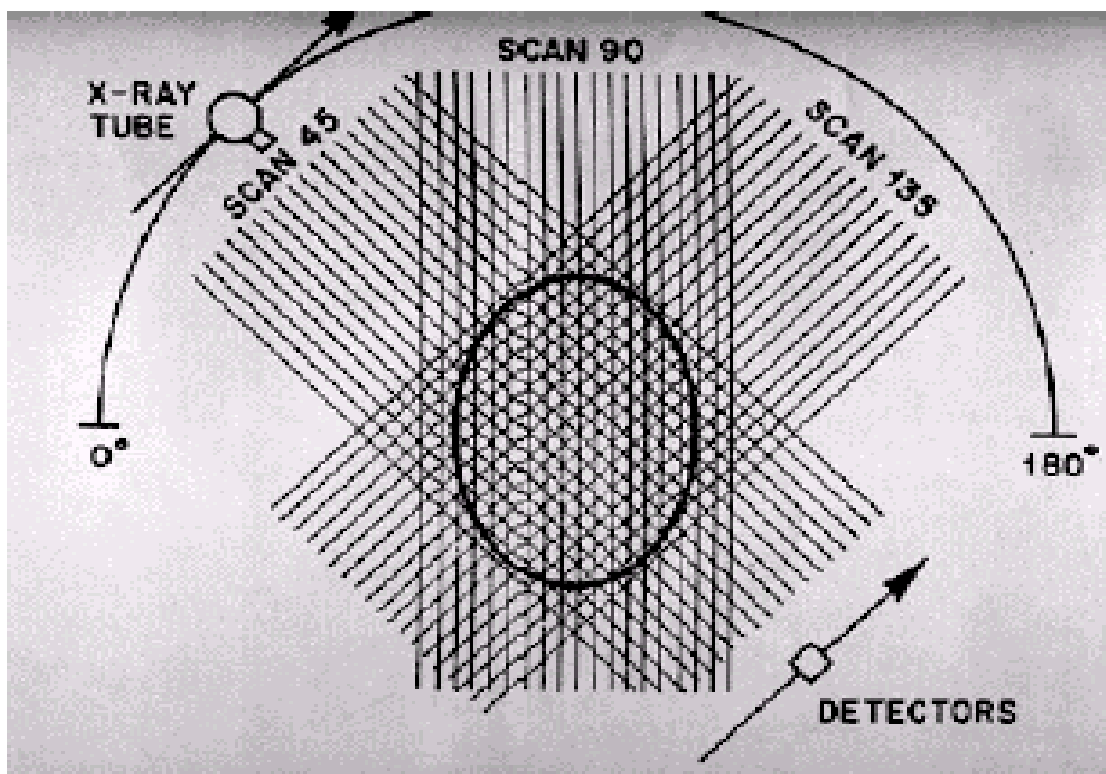


Figura 3.12- Tomógrafo de feixes paralelos

Em muitas técnicas usadas para reconstrução de imagens, o espalhamento no meio não é considerado, levando a problemas lineares de solução numérica mais simples, como o caso de Tomografia Computarizada (Computerized Tomography, CT) e Tomografia Computarizada por Emissão (Single Photon Emission Computerized, SPECT). Para os casos em que o espalhamento não pode ser desprezado, tal como a Tomografia Óptica Infravermelha (Near infrared Optical Tomography, NIROT), o modelo de reconstrução é mais complexo e não linear. Neste trabalho o espalhamento não é considerado.

REIS & ROBERTY, (1992) construíram uma base para a solução do problema inverso de tomografia de transmissão com feixes paralelos, e CARITA MONTERO *et al.*, 1999, 2000, apresentaram a construção de uma base para feixes divergentes. Estas bases (chamadas de bases naturais) representam uma situação física realista em que feixes paralelos de radiação são originados a partir de fontes externas colocadas ao redor do meio sob investigação, cujas intensidades são medidas após terem interagido com o meio absorvedor. A base construída ao longo das faixas é consistente com a radiação dos dados gerados, sendo então simplificada a formulação do problema inverso para a estimativa do coeficiente de absorção em meios participantes heterogêneos, quando um conjunto incompleto de dados experimentais está disponível.

3.5.1 Construção da partição de domínio (base natural) para feixes paralelos

Uma faixa representa o caminho da radiação que vai desde a fonte ao detector como se observa na Fig. 3.13, uma vista é a direção e sentido do caminho da radiação.

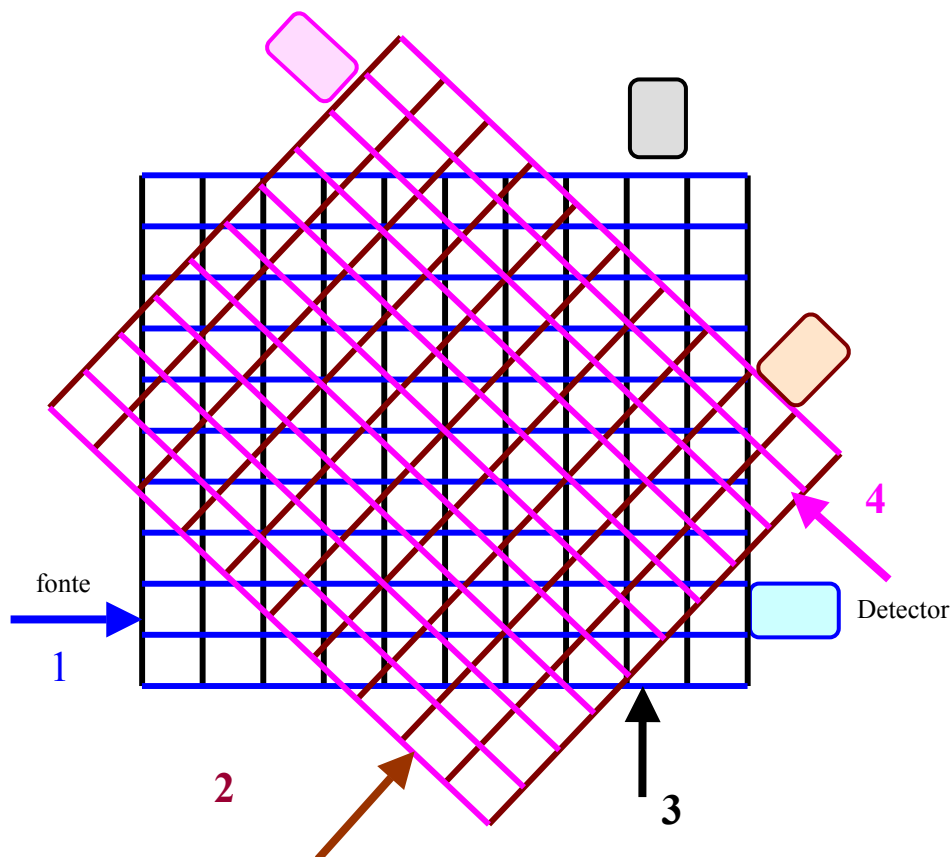


Figura 3.13- Construção da partição do domínio por 10 faixas e 8 vistas

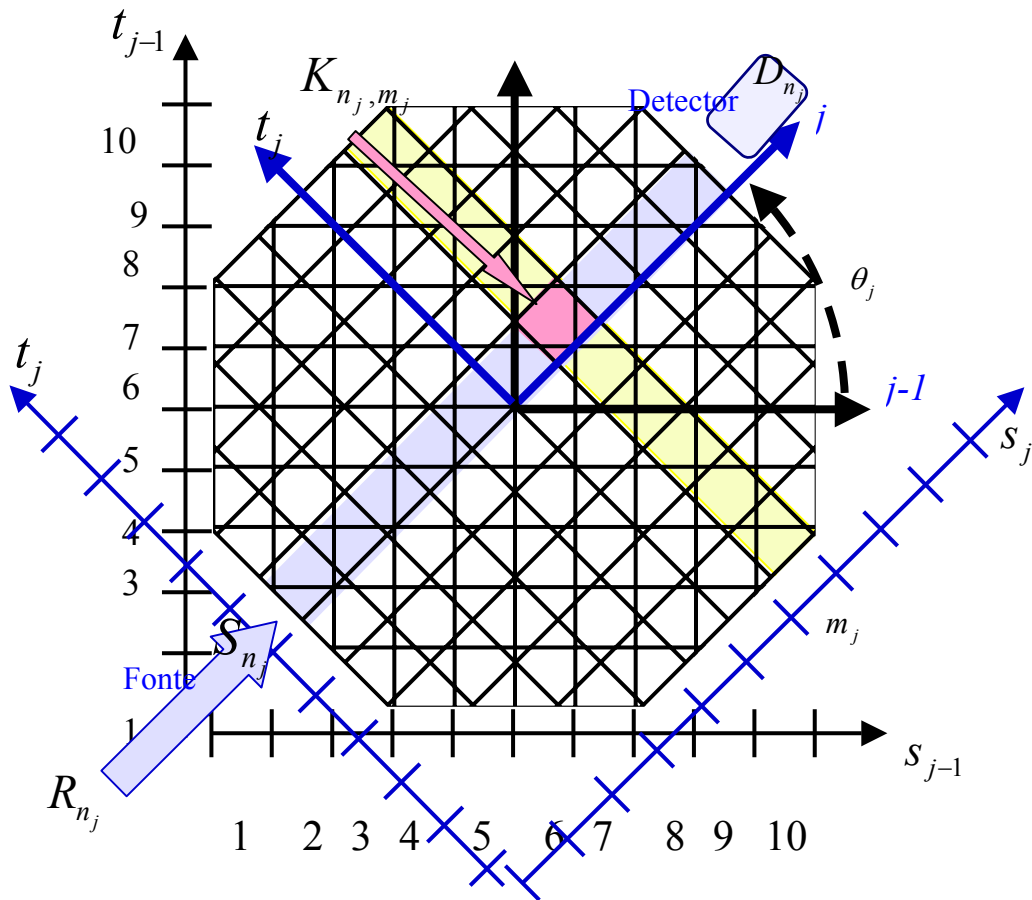


Figura 3.14- Discretização do domínio, sistema de coordenadas giradas. Localização de um elemento de discretização, K_{n_j, m_j} , na partição do domínio

A partição do domínio (Dom) para feixes paralelos é construída pela intercessão de todas as faixas. Observe na figura 3.14 que s_j e t_j formam um sistema de coordenadas giradas, aqui chamadas de malhas de pixels, sendo cada uma delas compatível com uma direção angular Ω_j que indica a direção de propagação da radiação, com $j = 1, 2, \dots, J$ onde J é o número total de giros ou de vistas do sistema.

O caminho por onde a radiação passa desde a fonte até o detector é representado por uma faixa R_{n_j} , assim para cada fonte S_{n_j} , onde n_j indica uma faixa na coordenada t_j com a radiação propagada na direção j . Assim $n_j = 1, 2, \dots, L$, onde L é um número par e indica o número de faixas em t_j . A fonte S_{n_j} , está alocada na

fronteiras entrantes $\partial R_{n_j}^-$, da faixa R_{n_j} e seu respectivo detectores D_{n_j} alocados na fronteiras saínte $\partial R_{n_j}^+$.

Então há um total de J sistemas de coordenadas giradas. Em cada um destas coordenadas giradas o problema de transporte transforma-se na derivada direcional, sendo resolvido como uma equação diferencial ordinária. A interseção completa de todas as malhas de pixels é levada em conta, gerando a partição de domínio, polígonos p_e com $e=1, 2, \dots, E$ onde E é o número total de polígonos (elementos).

Para a construção da base natural, primeiro são calculados as interseções de todas as linhas (observe que cada faixa tem duas linhas limitantes). Os pontos onde as interseções acontecem são os vértices dos polígonos p_e . Então é importante identificar um conjunto de vértices $(x_1, y_1), (x_2, y_2), \dots, (x_L, y_L)$, onde L é o número total de pontos de interseção (vértices que formam um polígono). Um algoritmo foi desenvolvido, no qual para cada pixel retangular $K_{n_j, m}$ (Vide figura 3.14) tem-se que pesquisar seqüencialmente, ao redor de cada um dos pontos de interseção. Obviamente, cada um dos pontos de interseção tem que ser um vértice de todos os polígonos criados ao redor dele. Os polígonos construídos deste modo são caracterizados pelos índices que relacionam os pontos de interseção com os índices de todas as faixas (e respectiva fonte) que deram origem a eles. Depois, no passo de numeração de polígonos, se um polígono é criado idêntico ao polígono criado por outro de seus vértices, o contador de elementos não é acrescido de uma unidade. Deste modo multiplicidade é evitada no processo de numeração de polígonos, e ao mesmo tempo os vértices que pertencem a cada polígono são determinados (CARITA MONTERO *et al.*, 2001).

A saída do código computacional escrito para a construção da base natural consiste numa tabela ordenada de todos os polígonos, suas áreas, opcionalmente seus vértices, e sua matriz de pertinência (matriz de incidência) para sua localização em cada um dos pixels orientados. A matriz de pertinência está definida por dois conjuntos de índices, i.e. n_j^e e m_j^e que localiza cada elemento particular e com respeito ao sistema de coordenadas giradas s_j e t_j , com $j = 1, 2, \dots, J$. E. Portanto, para cada segmento retangular $K_{n_j, m}$ conhecido, encontram-se os elementos que ele contém. Além disso, temos também na saída do código computacional uma listagem dos elementos que pertencem à interseção de duas faixas quaisquer.

3.5.2 Formulação do problema direto

Em tomografia de feixes paralelos ou espalhamento não é considerado, e das Eqs 3.4 e 3.5 e 3.7 tem-se:

$$\Omega \cdot \nabla I(r, \Omega) + \sigma_t(r)I(r, \Omega) = 0 \quad \left\{ V \in \mathbb{R}^2, \Omega \in 2\pi \right\} \quad (3.46)$$

com $I_{j,n_j}(r, \Omega) = I_{in}(r, \Omega)$, para $\Omega \cdot n < 0$ em $\forall r \in \partial V_1$

onde I é o fluxo angular da radiação, I_{in} o fluxo que ingressa ao meio, σ é o coeficiente de absorção no meio, r é a coordenada espacial e Ω é a direção angular, V é o domínio (volume de uma faixa R_{n_j}), ∂V é a fronteira do meio, os sub índices j, n_j indicam a posição de uma fonte e uma faixa da fonte j .

Se a geometria, a intensidade da radiação que entra no meio, e os coeficientes de absorção são conhecidos, a intensidade que sai do meio pode ser calculada resolvendo a Eq. (3.46) e cada sistema girado faixa por faixa.

Dadas a geometria, as propriedades radiativas, e as condições de fronteiras, os valores da intensidade de radiação em cada retângulo da faixa R_{n_j} , com $n_j = 1, 2, \dots, L$ e $j = 1, \dots, J$, podem ser calculados. Para cada faixa R_{n_j} , adotando o sistema de coordenada que gira de acordo com cada direção particular Ω_j , há uma equação diferencial ordinária a ser resolvida, e as $J \times L$ equações formam um sistema acoplado que deve ser resolvido

Uma forma de solucionar o problema direto é com ajuda do método de ordenadas discreta para a discretização do domínio angular, que é feito de modo que seja consistente com os feixes paralelos de radiação originada em fontes externa. Então, usando uma malha de pixels consistente com faixas paralelas de radiação (sistema de coordenadas giradas dado por (s_j, t_j)), que é girado em concordância à direção particular Ω_j , conforme mostrado na Fig. 3.15, a Eq. (3.4) pode ser escrita para cada faixa R_{n_j} da seguinte maneira

$$\frac{dI(r, \Omega)}{ds_j} + \sigma_t(r)I(r, \Omega) = 0$$

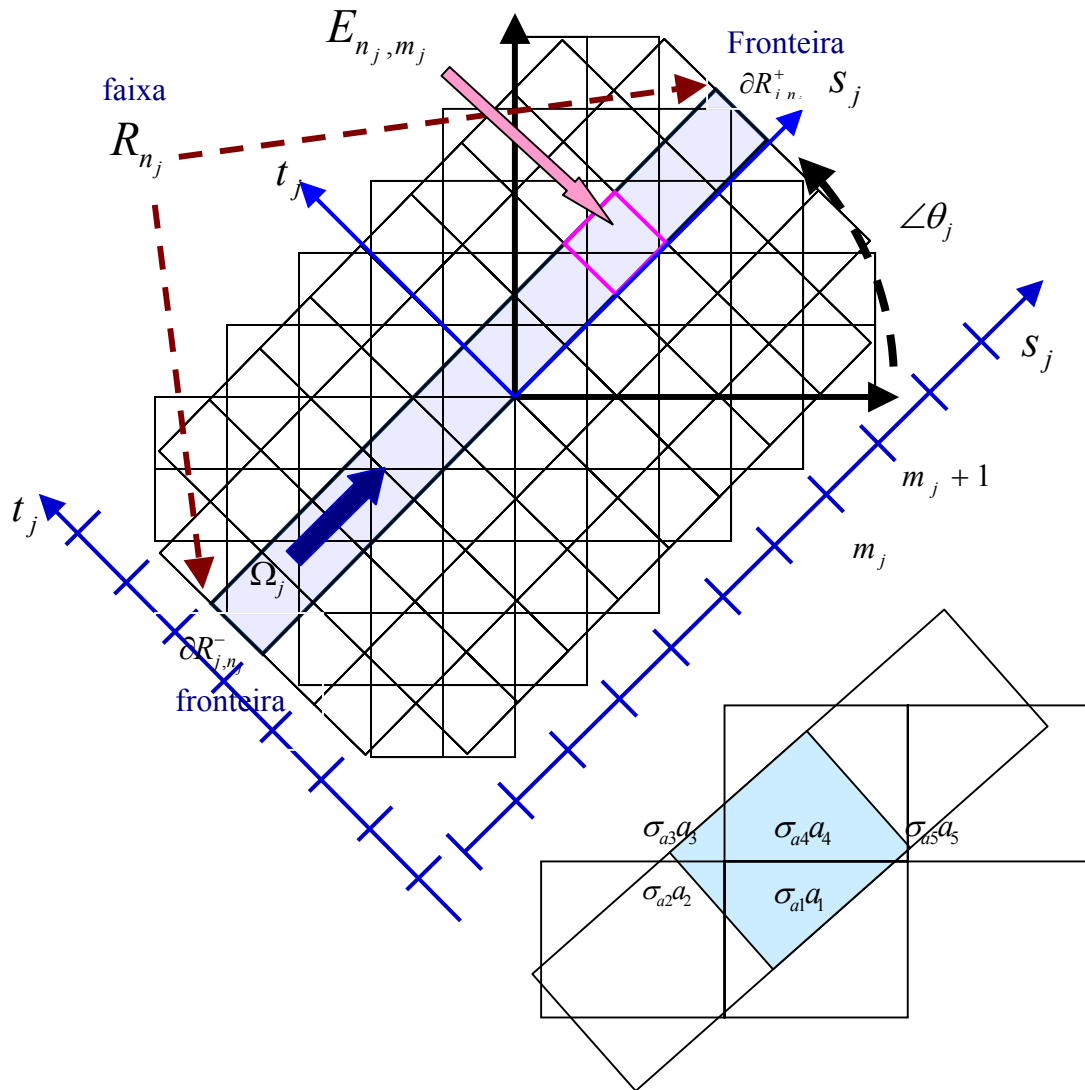


Figura 3.15.- O problema direto na faixa R_{n_j}

3.5.3 Solução do problema direto

Fazendo uma aproximação de diferenças finitas adiantadas para feixes paralelos de radiação e integrando no elemento de área E_{n_j, m_j} que é um elemento da partição de domínio pertencente a faixa n_j da fonte j ,

$$(I_{j, n_j, m_j+1} - I_{j, n_j, m_j}) + I_{j, n_j, m_j+1} \sum_{e \in K_{n_j, m_j}^j} \sigma_{a, e} a_e = 0$$

$$j = 1, 2, \dots, J, \quad n_j = 1, 2, \dots, 2M, \quad \text{e} \quad m_j = 1, 2, \dots, 2M \quad (3.47)$$

m_j é a discretização do eixo girado t_j e a_e umas das áreas que forma E_{n_j, m_j} , Vide Fig.

3.15. logo

$$I_{j,n_j,m_j+1} = \frac{I_{j,n_j,m_j}}{1 + \sum_{e \in K_{n_j,m_j}^j} \sigma_{a,e} a_e} \quad j = 1, 2, \dots, J, \quad n_j = 1, 2, \dots, 2M, \quad e \quad m_j = 1, 2, \dots, 2M \quad (3.48)$$

Outra forma para encontrar os fluxos medidos pelos detectores é descrita a seguir.

Vide a Figura 3.15 que apresenta a geometria do problema, uma faixa dada por uma fonte j . Para uma faixa n_j da fonte j dentro de uma região de reconstrução, da Eq. (3.46), tem-se :

Se $I_{j,n_j}(r, \Omega)$ é diferente de zero, a Eq. (3.46) pode ser escrita para o domínio discreto como

$$\nabla \cdot (\Omega \ln I_{j,n_j}(r, \Omega)) = -\sigma_e(r) \quad (3.49)$$

Integrando na faixa R_{j,n_j} , usando o teorema da divergência, observando que $\Omega \cdot n = 0$ nas superfícies laterais da faixa (vide Fig. 3.15), a equação anterior fornece o seguinte resultado:

$$\int_{e \in R_{j,n_j}} \sigma_e(r) dr = A \ln \left(\frac{I_{in}(r, \Omega)}{I_{out}(r, \Omega)} \right) = h_{j,n_j} \quad (3.50)$$

onde A é o comprimento ao longo da fronteira de entrada de radiação $(\partial R^-)_{j,n_j}$ e h_{j,n_j} é a relação não linear da intensidade de radiação entrante (I_{in}) e saínte (I_{out}) medida em $(\partial R^+)_{j,n_j}$.

No caso experimental real, a medição de I_{in} e I_{out} sempre tem erro experimental devido à instrumentação empregada, e a solução da Eq. (3.46) é uma aproximação para h_{j,n_j} . Então, sendo conhecido I_{in} é possível calcular I_{out} na faixa R_{j,n_j} , empregando a Eq. 3.50 e

$$h_{j,n_j} = \sum_{\substack{e=1 \\ e \in R_{j,n_j}}}^{E_{j,n_j}} \sigma_e a_e, \quad \forall j = 1, \dots, J \quad e \quad n_j = 1, \dots, N_j \quad (3.51)$$

onde a_e é a área do elemento e que pertence à faixa R_{j,n_j} , σ_e é coeficiente de absorção do elemento e , e E_{j,n_j} é o número total de elementos do cone n_j da fonte j .

CAPÍTULO 4

O PROBLEMA INVERSO

Uma definição simplificada do problema direto e inverso foi feita no capítulo anterior. Uma definição mais completa pode ser obtida em KIRSCH (1996). Segundo esse autor dois problemas (ou funções) são considerados inversos entre si, se a formulação de um deles requerer o conhecimento total (ou parcial) do outro. No problema direto a informação sempre deve ser completa e precisa. No problema da informação pode ser incompleta e imprecisa. Seguindo esta definição, é possível determinar qual problema deverá ser considerado direto ou inverso.

ALIFANOV (1974) nos dá a seguinte definição “*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. 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 4.1 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 direto, como o campo de temperatura, concentração de partículas, corrente elétrica, etc.

Na maioria dos casos, um problema inverso é considerado mal – posto. Para um problema ser considerado bem-posto, três requisitos essenciais deverão ser satisfeitos:

- i. Existência - o problema deverá apresentar solução;
- ii. Unicidade - a solução deverá ser única;
- iii. Estabilidade - a solução deverá exibir dependência contínua em relação aos dados que a originaram

Pode-se matematicamente, obter a existência e a unicidade de uma solução delimitando-se o espaço onde ela, provavelmente, se encontra. Se um problema exhibe múltiplas soluções, as informações que definem o operador são insuficientes para descrevê-lo corretamente. Neste caso, critérios adicionais baseados em informações de compromisso com a modelagem devem ser adotados.

O atributo da estabilidade é tido como o mais importante, já que é praticamente impossível convergir para o espaço de soluções sem considerar os efeitos degenerativos do ruído aditivo sobre os dados, o operador, ou até mesmo as limitações impostas pelo caráter iterativo do processamento numérico.

Caso a solução não dependa continuamente dos dados, os valores computados encontram-se, freqüentemente, distantes da solução esperada. Desta forma, não haveria como superar esta dificuldade a menos que se pudesse fornecer informações adicionais confiáveis acerca da solução. Portanto, *"a falta de informação não pode ser remediada por artifícios matemáticos"* (KIRSCH, 1996).

Consideremos uma função $F_k(z)$ igual a zero, onde $z = \{z_1, z_2, \dots, z_p\}$ é o vetor formado por todas as incógnitas do sistema, sendo p o número total de incógnitas, e $k=1, \dots, K$, expressa uma determinada equação. Logo podemos definir um sistema de equações formado por p incógnitas e K equações.

$$F_k(z) = 0 \text{ com } k = 1, \dots, K \text{ e } z = \{z_1, z_2, \dots, z_p\}$$

No sistema podemos ter:

Se $p < K$, o sistema está geralmente sobre-determinado e a unicidade da solução, no caso em que exista, pode ser encontrada quando se trabalha com dados exatos e consistentes com o modelo, e uma solução aproximada no caso que se considera ruído nas medições.

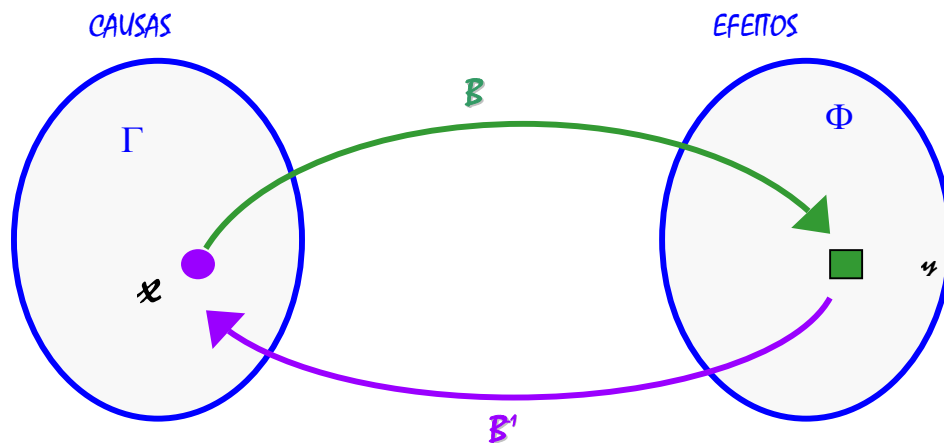


Figura 4.1 Representação esquemática do problema direto e inverso

Caso $p > K$, tem-se um sistema indeterminado que pode ter um conjunto infinito de soluções. Nesta situação, quando os dados são consistentes com o modelo, pode-se dizer que não há dados suficientes para resolver o problema de forma única sendo necessário outro tipo de análise. Não se encontrando a unicidade da solução para a formulação inicial do problema, procura-se dentro do conjunto de soluções o melhor resultado, segundo um critério de otimização (KAPUR & KESAVAN, 1992; CENSOR & LENT, 1981; REIS & ROBERTY, 1992; CARITA et al., 2001; CIDADE *et al.* 1998).

Estas classes de problemas são resolvidas como um problema de otimização que trata de encontrar a “melhor” solução entre todas as soluções possíveis, os métodos mais conhecidos aplicados a transferência radiativa são os métodos de regularização, o funcional de Tikhonov, o algoritmo *ART* e o método de máxima entropia.

4.1.- O método de máxima entropia generalizada

Seja σ o vetor formado por todas as incógnitas do sistema

$$\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_E\} \quad (4.1)$$

onde E é o número total de incógnitas. Os fluxos nos detectores podem ser calculados como uma função $f_k(\sigma)$ da solução do problema direto. Os fluxos medidos experimentalmente são denotados por q_{med}^k com $k = 1, \dots, K$; onde K é o número total de dados experimentais. A função erro é definida como a diferença entre q_{med}^k e $f_k(\sigma)$.

$$F_k(\sigma) = q_{med}^k - f_k(\sigma), \quad (4.2)$$

O método de máxima entropia generalizada está baseado na minimização da distância de Bregman restrita à função erro, sendo a distância de Bregman construída por uma função convexa relacionadas às entropias de Rényi, Varma, Shama-Taneja e Burn.

$$\text{Minimizar } D(\sigma, \sigma_0) \quad \text{sujeito a } F(\sigma) \quad (4.3)$$

Seja σ_0 uma informação inicial do valor de σ . A distância de Bregman entre σ e σ_0 , é definida pela Eq. (2.27), onde η é uma função convexa.

$$D(\sigma, \sigma_0) = \eta(\sigma) - \eta(\sigma_0) - (\nabla \eta(\sigma_0), \sigma - \sigma_0)$$

com

$$\langle \nabla \eta(\sigma_o), \sigma - \sigma_o \rangle = \sum_{e=1}^E \frac{\partial \eta}{\partial \sigma_e} \Big|_{\sigma=\sigma_o} (\sigma_e - \sigma_{oe}) \quad (4.4)$$

Observemos que cada η empregada vai definir uma nova distância de Bregman para ser utilizada na construção da lagrangeana

$$L_{s,r}(\sigma, \sigma_o, \lambda) = D(\sigma, \sigma_o) + \sum_{k=1}^K [\lambda_k F_k(\sigma)] \quad (4.5)$$

onde λ_k são os multiplicadores de Lagrange onde $k = 1, \dots, K$; observe que a Eq. (4.5) apresenta como incógnita o vetor σ formado pelos σ_e onde $e = 1, \dots, E$. O que dá um total de $(E + K)$ incógnitas.

Lembrando que para o funcional $\eta_{s,r}(\sigma, \sigma_o) = \frac{\sigma_e^s - \sigma_e^r}{s - r}$, a distância de Bregman é dada pelas Eq. 2.30 e 2.30^a, destas equações podemos escrever a lagrangeana na forma:

$$L_{s,r}(\sigma, \sigma_o, \lambda) = \frac{1}{s - r} \left[\sum_{e=1}^p (\sigma_e^s - \sigma_e^r) - \sum_{e=1}^p (\sigma_{oe}^s - \sigma_{oe}^r) - \sum_{e=1}^p (s \cdot \sigma_{oe}^{s-1} - r \sigma_{oe}^{r-1})(\sigma_e - \sigma_{oe}) \right] + \sum_{k=1}^K [\lambda_k F_k(\sigma)] \quad (4.6a)$$

se $s \rightarrow 1$ e $r \rightarrow 1$

$$L_{1,1}(\sigma, \sigma_o, \lambda) = - \left[\sum_{e=1}^p (\sigma_e - \sigma_{oe} - \sigma_e \ln \frac{\sigma_e}{\sigma_{oe}}) \right] + \sum_{k=1}^K [\lambda_k F_k(\sigma)] \quad (4.6b)$$

Para resolver este sistema se vai procurar o ponto crítico. Para isso, igualamos a zero as derivadas da lagrangeana com respeito a todas as suas incógnitas.

$$\frac{\partial}{\partial \sigma_e} L(\sigma, \sigma_o, \lambda) = 0 \quad e = 1, \dots, E \quad (4.7)$$

$$\frac{\partial}{\partial \lambda_k} (L(\sigma, \sigma_o, \lambda)) = 0, \quad k = 1, 2, \dots, K. \quad (4.8)$$

As equações (4.7) e (4.8) oferecem $(E+K)$ equações, pelo que agora temos um sistema formado de $(E+K)$ incógnitas e $(E+K)$ equações. Para resolver o sistema formado pelas Eqs. (4.7) e (4.8), foi escrito um código em MATLAB. Este “m-file” faz uso das ferramentas encontradas no “Toolbox” de otimização do MATLAB. A minimização deste sistema foi obtida com os algoritmos Quase-Newton e procura linear, estes algoritmos são descritos no apêndice II. Este método vai ser empregado na construção de outros algoritmos desenvolvido para o problema inverso unidimensional.

4.2.- Formulação e solução do problema inverso para reconstrução de imagens

A possibilidade da reconstrução de uma imagem a partir de diversas projeções foi estabelecida por RADON em 1917, e a aplicação deste problema no desenvolvimento de técnicas numéricas e procedimentos médicos com objetivo de diagnóstico, o desenvolvimento da tomografia computadorizada, foi realizada nos anos 60 e 70 por (CORMACK, 1963) e (HOUNSFIELD, 1972), trabalho pelo qual receberam o prêmio Nobel em 1979.

A reconstrução de imagens é feita a partir da estimativa de propriedades do meio, neste caso do coeficiente de absorção. Diferentes tipos de radiação tais como partículas neutras, raios gama e fóton têm sido usados para a identificação de formas de defeitos em peças, no caso da aplicação em ensaios não-destrutivos na indústria, ou regiões com alterações em tecido biológico como o caso da tomografia.

Para a formulação do problema inverso vamos empregar a Eq. (3.50) dada na seção 3.5.3.

$$\int_{R_{j,n_j}} \sigma(\vec{x}) d\vec{x} = A \ln \left(\frac{\phi_{in}(\vec{x}, \vec{\Omega})}{\phi_{out}(\vec{x}, \vec{\Omega})} \right) = h_{j,n_j} \quad (4.9)$$

onde A é o comprimento ao longo entre as fronteiras de entrada e saída da radiação na faixa j , n_j . Lembrar que ϕ_{out} representa a intensidade de saída que é medida experimentalmente.

Usando a partição de domínio de feixes paralelos, a Eq. (4.9) pode ser escrita na seguinte forma discretizada, que é consistente com o sistema fonte-detector ;

$$\sum_{e \in R_{j,n_j}} \sigma_e a_e = h_{j,n_j}, j= 1,2,\dots,J \text{ e } n_j = 1, 2,\dots, N_j \text{ e } e=1,2,\dots,E, \quad (4.10)$$

onde σ_e são as incógnitas, E é o número total de elementos e N_j é o número de dados disponíveis, cada dado está relacionado a h_{j,n_j} . Sendo $E \gg N_j$, temos que, existe um grande número de possíveis soluções para o sistema representado na Eq. (4.10). O problema inverso de reconstrução tomográfica, i.e. determinação de σ_e , $e=1,2,\dots,E$, é resolvido então como um problema de otimização, onde é definido um critério para a escolha de uma das soluções possíveis.

4.3.- Algoritmos ART com o funcional $\eta_{s,r}$

O funcional convexo empregado para definir a distância de Bregman é $\eta_{s,r}$ os outros funcionais apresentaram dificuldade para a implementação no algoritmo ART.

Agora definimos o vetor de incógnitas com σ , ele é formado por todas as incógnitas do sistema, quer dizer, os coeficientes de absorção de cada elemento de área (σ_e), que formam o meio.

Considerando a distância de Bregman para $\eta_{s,r}$, definida na Eq. (2.30), capítulo 3, definimos a lagrangeana baseada na distância de Bregman e a restrição dada pela Eq. (4.9)

$$L(\sigma, \sigma_o, \lambda) = D(\sigma, \sigma_o) + \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \left(h_{j,n_j} - \int_V \delta_{j,n_j}(\bar{x}) \sigma(\bar{x}) dx \right) \quad (4.11)$$

onde λ representa o multiplicador de Lagrange, e δ_{j,n_j} é a função característica da faixa n_j da fonte j .

$$\delta_{j,n_j}(\bar{x}) = \begin{cases} 1 & \text{se } \bar{x} \in R_{j,n_j} \\ 0 & \text{caso contrário} \end{cases} \quad (4.12)$$

e σ_o é alguma informação anterior relacionada com a incógnita σ .

A distância de Bregman para $\eta_{s,r}$ é

$$D_{s,r}(\sigma, \sigma_o) = \frac{1}{s-r} \int_V \sigma^s - \sigma^r - (\sigma_o^s - \sigma_o^r) - (s\sigma_o^{s-1} - r\sigma_o^{r-1})(\sigma - \sigma_o) dx \quad (4.13)$$

$$L_{s,r}(\sigma, \sigma_a, \lambda) = D_{s,r}(\sigma, \sigma_a) + \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \left(h_{j,n_j} - \int_V \sigma dx \right) \quad (4.14)$$

Resolvendo a lagrangeana, Eq. (4.14) no ponto critico,

$$\frac{\partial L_{s,r}(\sigma, \sigma_o, \lambda)}{\partial \sigma} = 0 \quad (4.15)$$

então:

$$\frac{\partial L_{s,r}(\sigma, \sigma_o, \lambda)}{\partial \sigma_e} = 0 = \frac{1}{s-r} [s\sigma_e^{s-1} - r\sigma_e^{r-1} - s\sigma_{oe}^{s-1} + r\sigma_{oe}^{r-1} + (s-r) \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \delta_{j,n_j,e}] \quad (4.16)$$

Agora desenvolveremos alguns casos especiais; quando $s = 0$ e $s = 1$, lembrar que expressões similares são obtidas para $r = 0$ e $r=1$, pois a dependência de r e s na função η_{sr} são idênticas.

1.- Para $s=0$

Observe, se fazemos $s=0$ na Eq. (4.16) obtemos:

$$0 = -\frac{1}{r} [-r\sigma_e^{r-1} + r\sigma_{oe}^{r-1} - r \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \delta_{j,n_j,e}] \quad (4.17)$$

Logo explicitando σ_e na Eq. (4.17) tem-se para $s=0$

$$\sigma_e = [\sigma_{oe}^{r-1} - \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r-1}} \quad (4.18)$$

Da Eq. (4.10) temos substituindo o valor de σ_e dada na Eq. (4.18), quando $s = 0$, temos

$$h_{j,n_j} = \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} - \lambda_{j,n_j}]^{\frac{1}{r-1}} a_e \quad (4.19)$$

Observamos que σ_e está relacionado com λ_{j,n_j} . Logo, definimos a função $F(\lambda_{j,n_j})$ como

$$F(\lambda_{j,n_j}) = h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} \sigma_e(\lambda_{j,n_j}) a_e \quad (4.20)$$

usando a série de Taylor, $F(\lambda_{j,n_j})$ pode ser escrita a partir de um λ_{oj,n_j} conhecido

$$F(\lambda_{j,n_j}) = F(\lambda_{oj,n_j}) + F'(\lambda_{oj,n_j})(\lambda_{j,n_j} - \lambda_{oj,n_j}) \quad (4.21)$$

Logo, obtemos λ_{j,n_j} quando $F(\lambda_{j,n_j}) \rightarrow 0$,

$$\lambda_{j,n_j} = \lambda_{oj,n_j} + \frac{F(\lambda_{oj,n_j})}{F'(\lambda_{oj,n_j})} \quad (4.22)$$

Para o caso $s=0$, tem-se:

$$F(\lambda_{j,n_j}) = h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} - \lambda_{j,n_j}]^{\frac{1}{r-1}} a_e \quad (4.23)$$

e a derivada

$$F'(\lambda_{j,n_j}) = \frac{1}{(r-1)} \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} - \lambda_{j,n_j}]^{\frac{2+r}{r-1}} a_e \quad (4.24)$$

Substituindo as Eqs. (4.23 e 4.24) na Eq. (4.22), tem-se:

$$\lambda_{j,n_j} = \lambda_{oj,n_j} + \frac{h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} - \lambda_{oj,n_j}]^{\frac{1}{r-1}} a_e}{\frac{1}{(r-1)} \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} - \lambda_{oj,n_j}]^{\frac{2+r}{r-1}} a_e} \quad (4.25)$$

Então é possível aproximar λ_{j,n_j} e σ_e que fazem $F(\lambda_{j,n_j}) \rightarrow 0$, de forma iterativa, a partir de λ_{oj,n_j} e σ_{oe} conhecidos. Quer dizer, sendo k como um contador iterativo, tem-se;

$$\lambda_{(k),j,n_j} = \lambda_{(k-1),j,n_j} + \frac{h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)oe}^{r-1} - \lambda_{(k-1)oj,n_j}]^{\frac{1}{r-1}} a_e}{\frac{1}{(r-1)} \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)oe}^{r-1} - \lambda_{(k-1)oj,n_j}]^{\frac{2+r}{r-1}} a_e} \quad (4.26)$$

Obtemos também da Eq. (4.19):

$$\sigma_{(k)e} = [\sigma_{(k-1)oe}^{r-1} - \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{(k-1)j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r-1}} \quad (4.27)$$

Expressões similares para, $F(\lambda_{j,n_j})$, $F'(\lambda_{j,n_j})$, e λ_{j,n_j} podem ser obtidas para $r=0..$

2. Para $s=1$,

Substituindo $s=1$ na Eq. (4.16);

$$0 = \frac{1}{1-r} [1 - r\sigma_e^{r-1} - 1 + r\sigma_{oe}^{r-1} + (1-r) \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \delta_{j,n_j,e}] \quad (4.28)$$

Logo, explicitando σ_e na Eq. (4.28) tem-se

para $s=1$

$$\sigma_e = [\sigma_{oe}^{r-1} + \frac{(1-r)}{r} \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r-1}} \quad (4.29)$$

A seguir só vamos desenvolver as equações para $s=1$.

Usando da Eq. (4.30), $F(\lambda_{j,n_j})$ pode ser calculada como

$$F(\lambda_{j,n_j}) = h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{oe}^{r-1} + \frac{(1-r)}{r} \lambda_{j,n_j}]^{\frac{1}{r-1}} a_e \quad (4.30)$$

e

$$F'(\lambda_{j,n_j}) = \sum_{e=1}^E \delta_{j,n_j,e} \left[\sigma_{oe}^{r-1} + \frac{(1-r)}{r} \lambda_{j,n_j} \right]^{\frac{2+r}{r-1}} \frac{a_e}{r} \quad (4.31)$$

Então é possível aproximar λ_{j,n_j} e σ_e que fazem $F(\lambda_{j,n_j}) \rightarrow 0$, em forma iterativa a partir de λ_{oj,n_j} e σ_{oe} conhecidos.

$$\lambda_{(k),j,n_j} = \lambda_{(k-1)j,n_j} + \frac{h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \lambda_{(k-1)j,n_j}]^{\frac{1}{r-1}} a_e}{\sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \lambda_{(k-1)j,n_j}]^{\frac{2+r}{r-1}} \frac{a_e}{r}} \quad (4.32)$$

e

$$\sigma_{(k)e} = [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{(k-1)j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r-1}} \quad (4.33)$$

4.3.1 Algoritmo solução

O algoritmo para $s = \mathbf{0}$ implementado no computador seria :

para $k=1,2,\dots$ até que a convergência seja obtida (loop externo)

para a fonte $j = 1$ até J

para a faixa $n_j = 1$ até N_j

para $k=1,2,\dots$ até que a convergência seja obtida

$$\lambda_{(k),j,n_j} = \lambda_{(k-1)j,n_j} + \frac{h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)e}^{r-1} - \lambda_{(k-1)oj,n_j}]^{\frac{1}{r-1}} a_e}{\frac{1}{(r-1)} \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)oe}^{r-1} - \lambda_{(k-1)oj,n_j}]^{\frac{2+r}{r-1}} a_e}$$

loop interno

para todo $e \in R_{j,n_j}$

$$\sigma_{(k)e} = [\sigma_{(k-1)oe}^{r-1} - \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{(k-1)j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r-1}}$$

O algoritmo para $s = \mathbf{1}$ implementado no computador seria :

para $k=1,2,\dots$ até que a convergência seja obtida (loop externo)

para a fonte $j = 1$ até J

para a faixa $n_j = 1$ até N_j

para $k=1,2,\dots$ até que a convergência seja obtida

$$\lambda_{(k),j,n_j} = \lambda_{(k-1)j,n_j} + \frac{h_{j,n_j} - \sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \lambda_{(k-1)j,n_j}]^{\frac{1}{r-1}} a_e}{\sum_{e=1}^E \delta_{j,n_j,e} [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \lambda_{(k-1)j,n_j}]^{\frac{2+r}{r-1}} \frac{a_e}{r}}$$

loop interno

para todo $e \in R_{j,n_j}$

$$\sigma_{(k)e} = [\sigma_{(k-1)e}^{r-1} + \frac{(1-r)}{r} \sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{(k-1)j,n_j} \delta_{j,n_j,e}]^{\frac{1}{r}}$$

Para $s \rightarrow 1$ e $r \rightarrow 1$, obtemos a função relacionada com a entropia de Shannon, neste caso, consideramos agora λ_{j,n_j} constante para todos os elementos na faixa R_{j,n_j} , o algoritmo é também conhecido como MART, e pode ser escrito como

para $k=1, 2, \dots$ até que a convergência seja obtida

para a fonte $j = 1$ até J

para a faixa $n_j = 1$ até N_j

$$\exp(\lambda_{(k)j,n_j}) = \frac{h_{j,n_j}}{\sum_{e=1}^E \delta_{j,n_j,e} \sigma_e a_e}$$

para todo $e \in R_{j,n_j}$

$$\sigma_{(k)e} = \sigma_{(k)oe} \exp \left[\sum_{j=1}^J \sum_{n_j=1}^{N_j} \lambda_{j,n_j} \chi_{j,n_j} \right]$$

Fica evidente que o problema de reconstrução depende apenas dos dados experimentais, de uma estimativa inicial para a incógnita, de uma adequada ordenação de fontes, feixes, elementos e do conhecimento de suas áreas. Assim a partição de domínio adotada neste trabalho, com a base natural, é muito conveniente para a implementação computacional.

4.4.- Método de Tikhonov

O método de 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 mal posto num problema bem –posto.

Tem-se agora um problema similar de otimização: dada pela minimização do funcional erro, definido na Eq. (4.2), elevado ao quadrado sujeito a um operador de regularização TIKHONOV & ARSENIN, (1977).

$$\text{Minimizar } \sum_{k=1}^K [F_k(\sigma)]^2 \quad \text{sujeito a } \alpha D(\sigma, \sigma_0) \quad (4.34)$$

onde α é o parâmetro de regularização e o termo de regularização considerado é a distância de Bregman

$$M = \left\{ \sum_{k=1}^K [F_k(\sigma)]^2 + \alpha D(\sigma, \sigma_0) \right\} \quad (4.35)$$

A técnica dos multiplicadores de Lagrange permite colocar na mesma função custo os objetivos de fidelidade dos parâmetros com o modelo direto e a de regularidade (suavidade) eleita da quantidade desconhecida, da seguinte forma:

$$\text{Minimizar } M = \text{Minimizar } \left\{ \sum_{k=1}^K [F_k(\sigma)]^2 + \alpha D(\sigma, \sigma_0) \right\} \quad (4.36)$$

Note que se $\alpha \rightarrow 0$ o termo de fidelidade dos dados, ou função objetivo, é superestimado enquanto para $\alpha \rightarrow \infty$ toda informação no modelo matemático é perdida.

Como a minimização do funcional descrito pela Eq (4.36) resulta da obtenção de uma solução única, condizente com os valores esperados (σ), emprega-se a equação de ponto crítico a ser resolvida.

$$\frac{\partial M(\sigma)}{\partial \sigma_e} = 0 \quad , \quad N(\sigma) = \sum_{e=1}^E \frac{\partial M(\sigma)}{\partial \sigma_e} = 0 \quad (4.37)$$

Neste sistema se tem um número igual de equações e de incógnitas, e ele pode ser resolvido através do método de Newton-Raphson multivariável (LIM, 1990; PRESS *et al*, 1996; CIDADE ,2000), que é um método iterativo, fazendo

$$\sigma^{k+1} = \sigma^k + \Delta \sigma^k, \quad \text{com } k = 0,1,2,\dots; \quad (4.38)$$

até obter, através de um algoritmo iterativo,

$$N(\sigma^{k+1}) = 0 \quad (4.39)$$

onde k indica o número de iterações e σ^k o valor inicial.

A linearização da Eq. (4.37) foi obtida a partir da retenção dos termos de primeira ordem da expansão de Taylor:

$$N(\sigma^{k+1}) = N(\sigma^k) + \sum_{e=1}^E \frac{\partial N(\sigma)}{\partial \sigma_e} \Big|_{\sigma_e^k} (\Delta \sigma_e^k) \quad (4.40)$$

Então, se $N(\sigma^{k+1}) = 0$, têm-se

$$-N(\sigma^k) = \sum_{e=1}^E \frac{\partial N(\sigma)}{\partial \sigma_e} \Big|_{\sigma_e^k} (\Delta \sigma_e^k) \quad (4.41)$$

É possível calcular σ^{k+1} a partir de um σ^k e um σ^{k-1} , considerando as Eqs. 4.40 e 4.41

$$\sigma_e^{k+1} = \sigma_e^k + \frac{1}{\frac{\partial N(\sigma)}{\partial \sigma_e} \Big|_{\sigma_e^k}} [N(\sigma_e^{k-1}) + \frac{\partial N(\sigma)}{\partial \sigma_e} \Big|_{\sigma_e^{k-1}} (\Delta \sigma_e^{k-1})] \quad (4.42)$$

Neste trabalho tem-se que

$$M_{s,r} = \sum_{j,nj} \left[h_{j,nj} - \sum_{e \in R_{j,nj}} \sigma_e a_e \delta_{e,j,nj} \right]^2 + \alpha D_{rs}(\sigma, \sigma_0) \quad (4.43)$$

logo, para $s = 0$ com $r > 1$

$$N_{0,r}(\sigma_e, \sigma_{eo}) = \frac{\partial M_{0,r}(\sigma_e, \sigma_{eo})}{\partial \sigma_e} = \sum_{j,nj} (-2h_{j,nj} \delta_{e,j,nj} a_e + 2a_e \sum_{e \in R_{j,nj}} \sigma_e a_e \delta_{e,j,nj}) + \alpha (\sigma_e^{r-1} - \sigma_{oe}^{r-1}) = 0 \quad (4.44)$$

para $s \rightarrow 1$ com $r > 0$

$$N_{1,0}(\sigma_e, \sigma_{eo}) = \sum_{j,nj} (-2h_{j,nj} \delta_{e,j,nj} a_e + 2a_e \sum_{e \in R_{j,nj}} \sigma_e a_e \delta_{e,j,nj}) + \alpha \frac{r}{1-r} (\sigma_e^{r-1} - \sigma_{oe}^{r-1}) = 0 \quad (4.45)$$

para $s \rightarrow 1$ e $r \rightarrow 1$

$$N_{11}(\sigma_e, \sigma_{eo}) = \sum_{j,nj} (-2h_{j,nj} \delta_{e,j,nj} a_e + 2a_e \sum_{e \in R_{j,nj}} \sigma_e a_e \delta_{e,j,nj}) + \alpha \ln \frac{\sigma_e}{\sigma_{oe}} = 0 \quad (4.46)$$

as derivadas parciais podem ser obtidas sem dificuldade

$$N'(\sigma_e) = \frac{\partial N(\sigma_e, \sigma_{eo})}{\partial \sigma_e} \quad (4.47)$$

para $s = 0$ com $r > 1$

$$N'_{0,r}(\sigma_e) = \sum_{j,nj} (\delta_{e,j,nj} 2a_e \sum_{e \in R_{j,nj}} a_e) + \alpha(r-1)\sigma_e^{r-2} = 0 \quad (4.48)$$

para $s \rightarrow 1$ com $r > 0$

$$N'_{1,r}(\sigma_e) = \sum_{j,nj} (2a_e \delta_{e,j,nj} \sum_{e \in R_{j,nj}} a_e) + \alpha r \sigma_e^{r-2} = 0 \quad (4.49)$$

para $s \rightarrow 1$ e $r \rightarrow 1$

$$N'_{1,1}(\sigma_e) = \sum_{j,nj} (2a_e \delta_{e,j,nj} \sum_{e \in R_{j,nj}} a_e) + \alpha \ln \frac{1}{\sigma_e} = 0 \quad (4.50)$$

Estas derivadas podem ser obtidas numericamente, de

$$N'(\sigma_e) = \frac{N(\sigma_e + vh) - N(\sigma_e)}{vh}, \quad (4.51)$$

nos empregamos $vh = 0.0005$, devido que as derivadas numéricas se aproximam às derivadas analíticas com um erro de 10^{-5} .

O emprego do método iterativo mostrou-se apropriado, devido ao seu caráter de correção sucessiva (KREYSZIG, 1972). Portanto, utilizando as Eqs. (4.44) e (4.50) pode-se calcular o valor dos incrementos e o valor de σ^{k+1} dada pela Eq (4.42).

4.4.1. Algoritmo solução com $s=0$ e $r=1$

1. passo; inicialização do método escolha uma estimativa inicial σ_{eo} , com $e = 1, \dots, E$
2. passo; calcular $N^*(\sigma_{eo}) = \sum_{j,nj} (-2h_{j,nj} \delta_{e,j,nj} a_e + 2a_e \sum_{e \in R_{j,nj}} \sigma_{eo} a_e \delta_{e,j,nj})$
3. passo, calcular $N^{*'}(\sigma_{eo})$ da Eq. 4.51

4. passo $\sigma_e^1 = \sigma_{e0} + \frac{N^*(\sigma_{e0})}{N'^*(\sigma_{e0})}$

5. passo calcular $N(\sigma_e^1)$ e $N'(\sigma_e^1)$ das Eqs. 4.44 e 4.51

6. passo; fazer $k= 1$ ate 60000

$$\sigma_e^{k+1} = \sigma_e^k + \frac{1}{N'(\sigma_e^k)} [N(\sigma_e^{k-1}) + N'(\sigma_e^{k-1}) * (\sigma_e^k - \sigma_e^{k-1})],$$

7. passo; Se $\sum_{e=1}^E |(\sigma_e^k - \sigma_e^{k-1})| < 10^{-4}$, Sair ao passo 9

8. passo; Caso contrario

Calcular $N(\sigma_e^{k+1})$ e $N'(\sigma_e^{k+1})$ das Eqs. 4.45 e 4.51, logo retornar ao passo 5, para obter um novo valor de k

9. passo $\sigma_e = \sigma_e^{k+1}$, fim

Algoritmos similares são obtidos para $s \rightarrow 1, r > 0$ e $s \rightarrow 1, r \rightarrow 1$.

CAPÍTULO 5

RESULTADOS

Neste capítulo apresentam-se resultados dos casos:

1. Estimativa do coeficiente de absorção para um meio unidimensional
2. Estimativa do coeficiente de absorção e espalhamento para um meio bidimensional
3. Reconstrução de imagens, empregando o algoritmo *ART*
4. Reconstrução de imagens, empregando a Regularização de Tikhonov

Procurando um número que nos indique a qualidade nas estimativas das incógnitas de cada caso teste, definimos o erro percentual total (EP) dado por;

$$Ep_e = \left| \frac{\sigma_{e,exato} - \sigma_{e,calculado}}{\sigma_{e,exato}} \right| \% \qquad EP = \frac{\sum_{e=1}^E Ep_e}{E} \% \qquad (5.1)$$

onde o sub - índice e indica um determinado elemento discretizado, E o número total de elementos em que foi discretizado o meio.

É necessário lembrar que para cada caso teste se faz uma comparação entre os resultados obtidos com diferentes algoritmos, considerando como critério o erro percentual (EP). Os tempos computacionais gastos ou as iterações empregadas, nos algoritmos solução do problema inverso, são também apresentados. Para o meio unidimensional os tempos computacionais foram pequenos, menores em 1 s.

Como dados experimentais reais não estavam disponíveis, foram empregados dados exatos e dados experimentais sintéticos. Os dados exatos são aqueles obtidos da solução do problema direto e os dados experimentais sintéticos são obtidos ao adicionar ruído randômico aos dados exatos (valores entre 0 e 1), obtidos da expressão;

$$q_{i\text{exp}} = q_{i\text{exato}} + \nu_i \qquad \nu_i = \pm \text{percentagem} * \text{numero rand}(i) * q_{i\text{exato}} \qquad (5.2)$$

onde $q_{i\text{exato}}$ é o fluxo obtido com valores estabelecidos a priori para as “incógnitas” do problema inverso, ν_i é a percentagem máxima de ruído randômico, então com diferentes ν_i podem ser simulados diversos níveis de ruído no dado experimental.

Todas as grandezas aqui apresentadas são adimensionais.

5.1 Problema unidimensional

5.1.1 Tipo e localização dos detectores

Os detectores aqui considerados medem o valor da intensidade da radiação, ou do fluxo angular de partículas não carregadas, integradas no domínio angular, ou seja, a grandeza medida não apresenta dependência com a variável angular (vide Fig. 5.1).

Nos seguintes exemplos espera-se recuperar os coeficientes de absorção de um meio participante empregando o método de máxima entropia generalizada baseada em três famílias de distâncias de Bregman descritas no capítulo 2.

Em todos os exemplos considera-se: a intensidade da fonte é igual a 1 e o coeficiente de espalhamento conhecido. Espera-se estimar os coeficientes de absorção, a partir das medidas dos detectores que medem os fluxos transmitidos $q'_{i\max}$ e q'_1 , e os detectores que medem os fluxos refletidos q^r_1 e $q^r_{i\max}$. A Fig. 5.1a apresenta a geometria do problema.

As intensidades medidas pelos detectores são calculadas por:

$$q'_{i\max} = \sum_{\substack{l=1 \\ \mu_l > 0}}^{Lo} w_l \cdot Ie_{l,i\max} \cdot \mu_l \quad q^r_1 = \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} -w_l \cdot Iw_{l,1} \cdot \mu_l \quad (5.3)$$

$$q'_1 = \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} -w_l \cdot Iw_{l,1} \cdot \mu_l \quad q^r_{i\max} = \sum_{\substack{l=1 \\ \mu_l > 0}}^{Lo} w_l \cdot Ie_{l,i\max} \cdot \mu_l \quad (5.4)$$

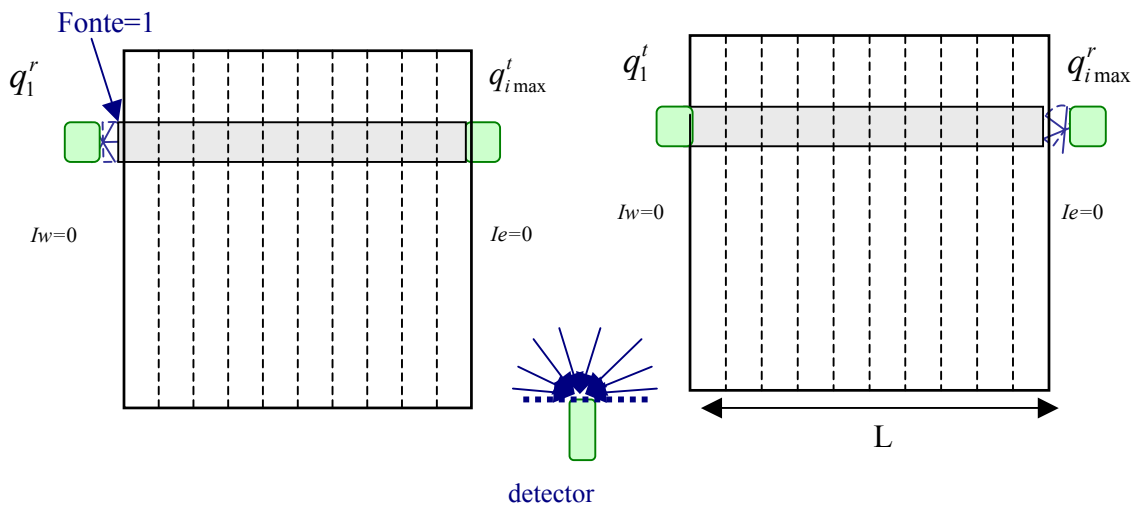


Fig. 5.1 - Geometria do problema

5.1.2 Validação do problema direto no meio unidimensional

A validação do algoritmo solução do problema direto foi feita por comparação dos nossos resultados com aqueles publicados (FIVELAND,1987; CHANDRASEKHAR,1960), para o seguinte problema:

Considere duas placas planas, isotérmicas, grandes, paralelas, separadas de uma distância L. Uma placa está a temperatura T_1 com emissividade $\varepsilon_1=1$ e a outra placa a temperatura $T_2=0$, com emissividade $\varepsilon_2=0$. O meio cinza difuso é absorvedor/emissor, tudo o que absorve, emite, conforme nossas equações (balanço de energia); não existe perda por absorção. Se o sistema esta em equilíbrio radiativo pede-se determinar o fluxo que atravessa a placa. O comprimento do meio é $L=1$.

Este problema foi resolvido para um meio com dispersão isotrópica. A equação empregada para determinar o fluxo que atravessa a placa foi:

$$q_y = \pi \cdot I_{bw} - q_{(1)}^r \quad (5.4)$$

onde $\pi \cdot I_{bw} = 1$ e está relacionada com T_1 . A Tabela 5.1 mostra os resultados exatos (MODEST, 1987) e os calculados, neste trabalho, com as quadraturas LSH S_{10} (FIVELAND, 1991), Gauss –Legendre (ABRAMOWITZ & STEGUN, 1970) com S_{10} e S_{20} ,

A Fig. 5.2 apresenta uma comparação dos q_y e dos tempos computacionais *versus* diferentes número de discretizações do domínio espacial, obtidos com as quadraturas: Gauss - Legendre S_{10} e S_{20} , e LSH S_{10} .

Neste trabalho se emprego a quadratura LSH S_{10} , com uma partição do domínio igual a 10, devido que ela oferece resultados similares à quadratura Gauss – Legendre S_{20} e o tempo computacional gasto do LSH S_{10} , é menos que a metade empregada no Gauss – Legendre S_{20} com uma partição do domínio igual a 20.

Tabela 5.1 Valores calculados de q_y

σ_t	Valores exatos	Gauss –Leg S_{10}	Gauss –Leg S_{20}	Este trabalho LSH S_{10}	Harmônicos esféricos P_1
0	1.000	1.0000	1.000	1.000	1.000
0.1	0.9157	0.9131	0.9145	0.9138	0.9302
0.5	0.7040	0.6982	0.7026	0.7026	0.7273
1.0	0.5532	0.5469	0.5517	0.5529	0.5714

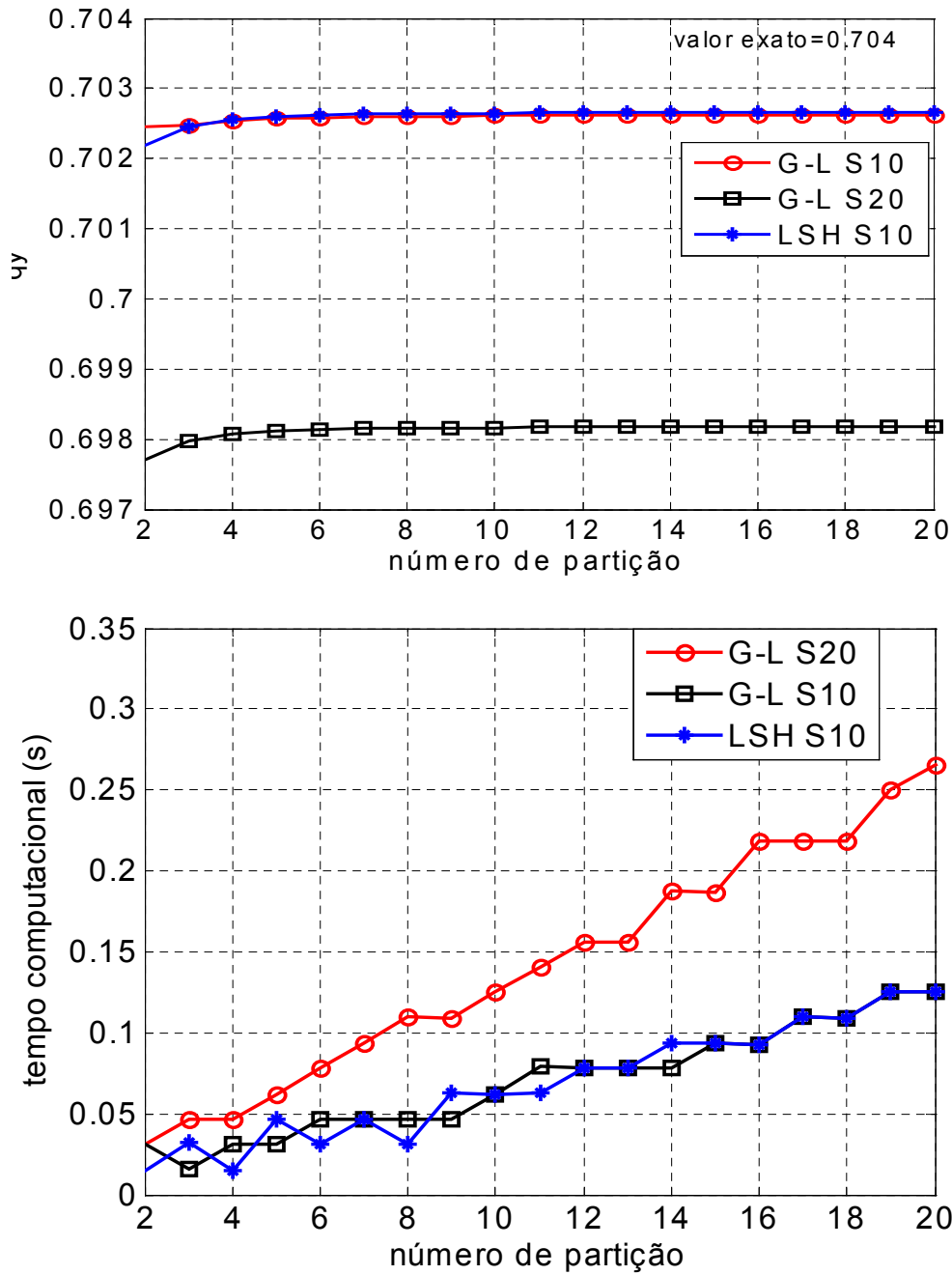


Fig. 5.2- Comparação dos q_y dos tempo computacional obtidos com as quadraturas: Gauss - Legendre S_{10} e S_{20} , e LSH S_1 para diferentes partições do domínio.

5.1.3 Estimativa do coeficiente de absorção empregando a função $\eta_{sr}(z_i) = \frac{(z_i^s - z_i^r)}{s - r}$ na construção da distância de Bregman. Método de máxima entropia generalizada

5.1.3.1 Estimativa do coeficiente de absorção

Nas Figs. 5.3, 5.4, 5.5 e 5.6 apresentam-se diversas curvas obtidas com os erros percentuais, EP, gerados com diferentes η_{sr} (valores de r e s), ao tratar de estimar os σ_a em um meio homogêneo, com propriedades dadas na tabela 5.2. A tabela 5.3, apresenta os valores iniciais σ_{ao} , valores mínimos lb e máximos ub onde as incógnitas vão ser procuradas.

Destas Figs., observa-se:

- 1.- Figs. 5.3, 5.4, 5.5 e 5.6 as linhas para $s = 0$ e $s \rightarrow 1$ se superpõem entre $1 < r < 8$,
- 2.- Figs. 5.3, 5.4, 5.5 e 5.6, a linha $s \rightarrow 1$ se une à linha $s=1$, em $r=1$
- 3.- Figs. 5.4, 5.5 e 5.6, a curva para $s \rightarrow 1$ apresenta menor EP e a linha $s = 1$ o maior EP comparadas com as outras curvas. Nos três casos σ_{ao} é menor que o dado a ser estimado
- 4.- Na Fig. 5.5, a curva $s = 1$ apresenta menor EP e a linha $s \rightarrow 1$ o maior EP comparadas com as outras curvas. Neste caso σ_{ao} é maior que o dado a ser estimado

Levados destas observações se fez um estudo dos dados iniciais e sua relação com a curva $s \rightarrow 1$ e $s = 1$, com o objetivo de encontrar os r ótimos que permitam um menor EP na estimativa de σ_a .

Tabela 5.2. Propriedades do meio em estudo

	σ_a	σ_s	Intensidade da fonte	Número de partição do meio
Fig. 5.3	0.9	0.5	1.0	10
Fig. 5.4	0.3	0.5	1.0	10
Fig. 5.5	0.9	0.5	1.0	10
Fig. 5.6	0.3	0.5	1.0	10

Tabela 5.3- Dados iniciais empregados para a estimativa dos σ_a

	σ_{ao}	σ_{os}	lb (valor mínimo)	ub (valor máximo)
Fig. 5.3a e 5.4	0.0001	0.5	0.0001	3.0
Fig. 5.4 e 5.5	0.5	0.5	0.0001	3.0

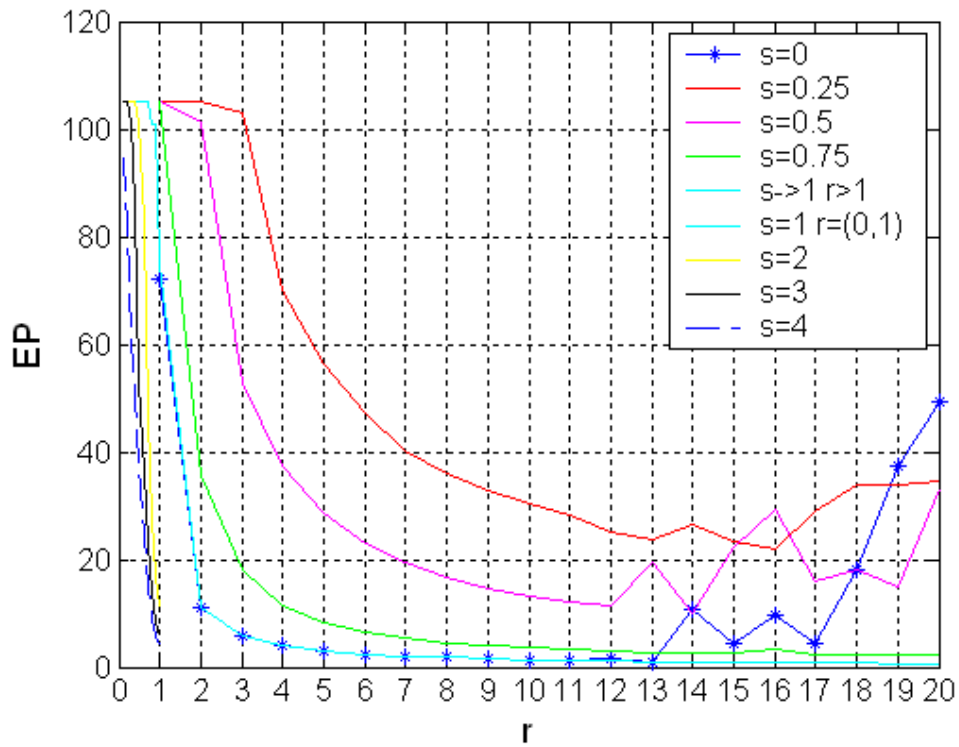


Fig. 5.3. Erros percentuais (EP) na estimativa de σ_a . Teste $\sigma_{ao} = 0.0001, \sigma_a = 0.8$

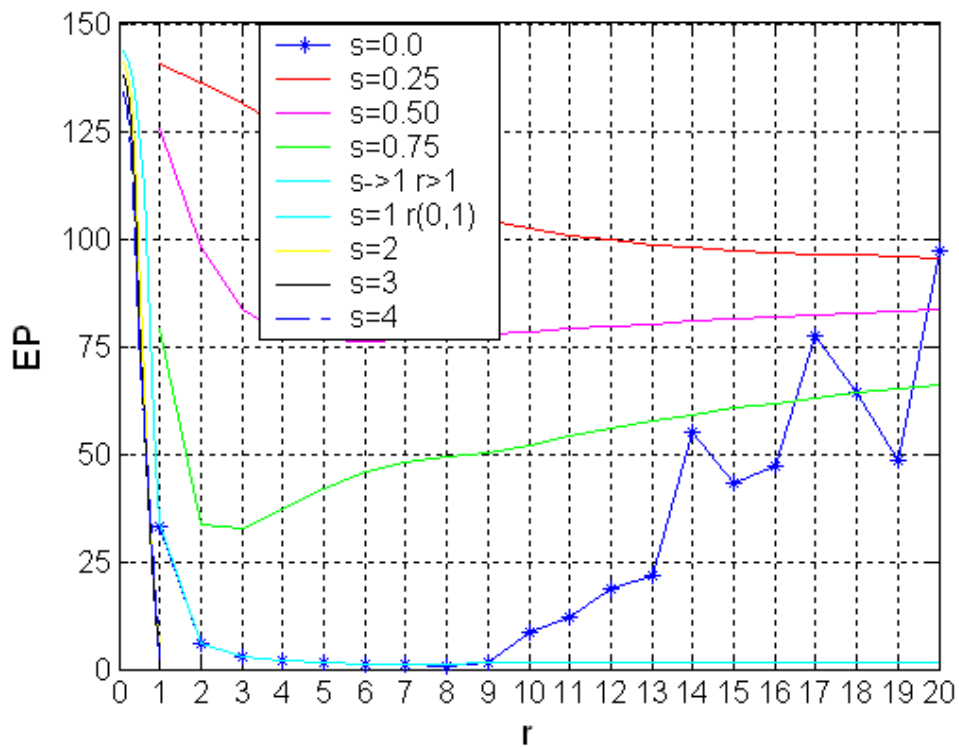


Fig. 5.4. Erros percentuais(EP) na estimativa de σ_a . Teste $\sigma_{ao} = 0.0001, \sigma_a = 0.3$

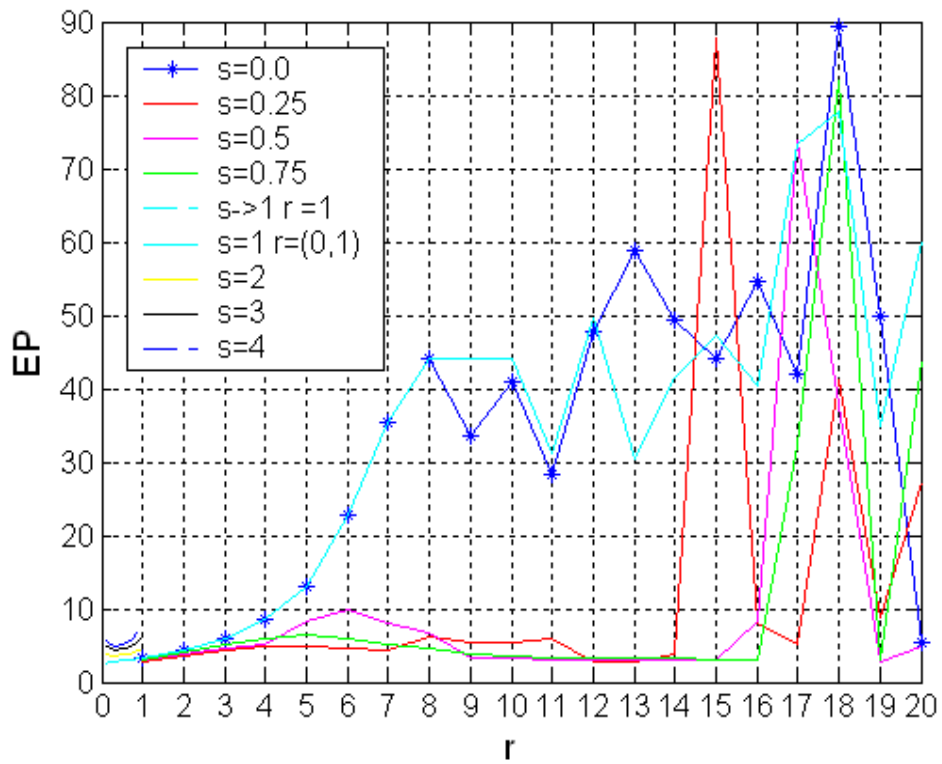


Fig. 5.5. Erros percentuais (EP) na estimativa de σ_a . Teste $\sigma_{ao} = 0.5$, $\sigma_a = 0.3$

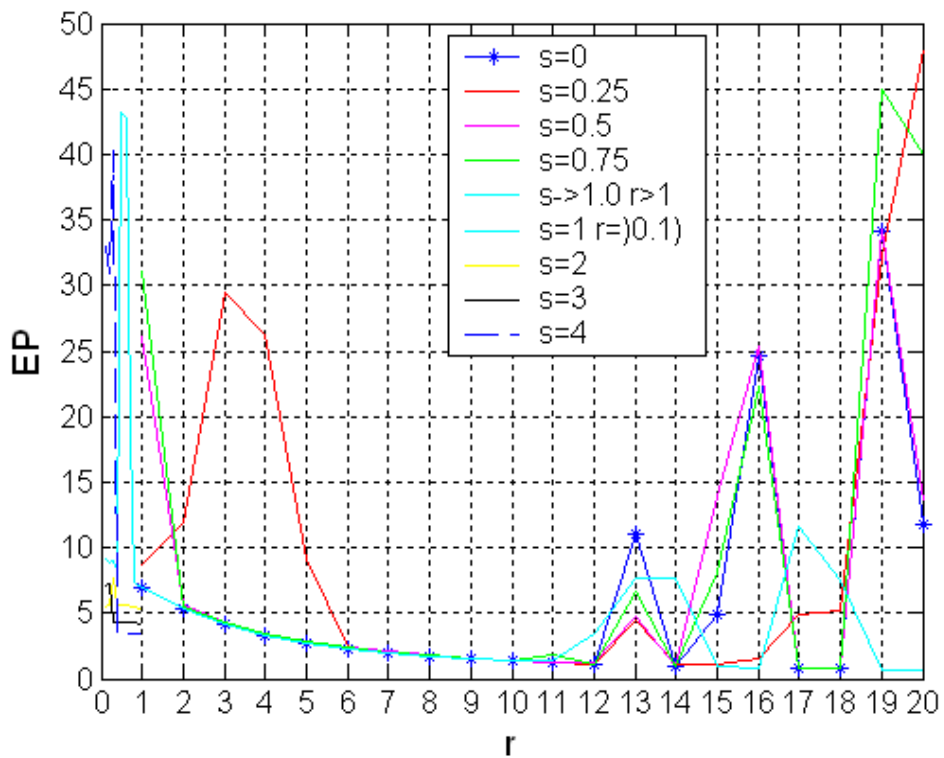


Fig. 5.6 Erros percentuais (EP) na estimativa de σ_a . Teste $\sigma_{oa} = 0.5$, $\sigma_a = 0.9$

5.1.3.2 Os valores iniciais σ_{ao}

As Figs. 5.7, 5.8, 5.9, 5.10, 5.11 e 5.12, apresentam os EP obtidos com diferentes σ_{ao} segundo a tabela 5.4, para $s \rightarrow 1$ e variando r no intervalo $[0,20]$.

Cada Fig., apresentam 10 curvas, cada curva é um caso para um σ_a diferente a ser estimado no meio. O valor de σ_a varia entre $[0.1$ e $1.0]$, vide legenda de cada Fig..

Tabela 5.4- Dados iniciais empregados para a estimativa dos σ_a

Fig. 5.7	$\sigma_{oa} = 0.0001$
Fig. 5.8	$\sigma_{oa} = 0.001$
Fig. 5.9	$\sigma_{oa} = 0.01$
Fig. 5.10	$\sigma_{oa} = 0.1$
Fig. 5.11	$\sigma_{oa} = 0.5$
Fig. 5.12	$\sigma_{oa} = 1.0$
Em todos os casos,	$\sigma_{os} = 0.5$
Em todos os casos,	$lb = 0.0001$
Em todos os casos,	$ub = 3.0$

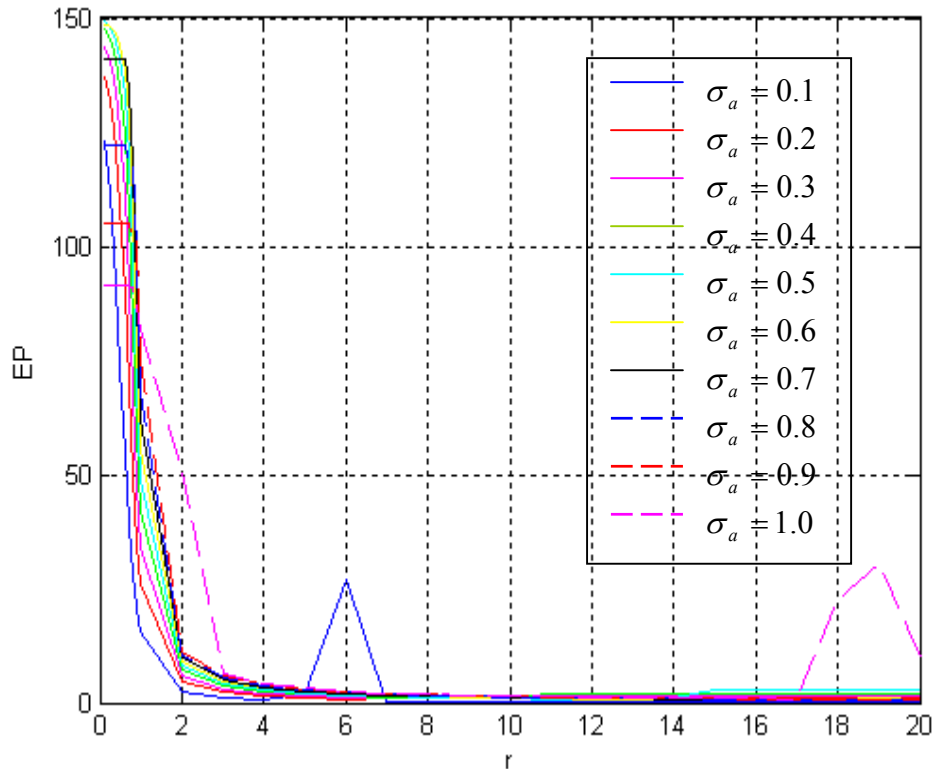


Fig. 5.7 - EP obtidos considerando $\sigma_{oa} = 0.0001$

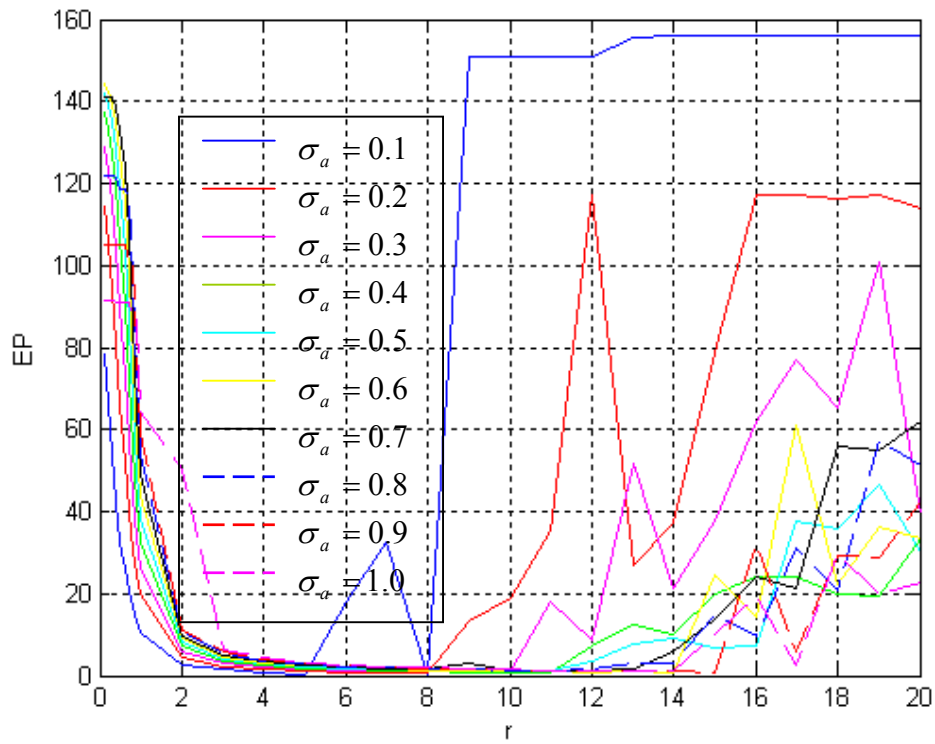


Fig. 5.8 EP obtidos considerando $\sigma_{oa} = 0.001$

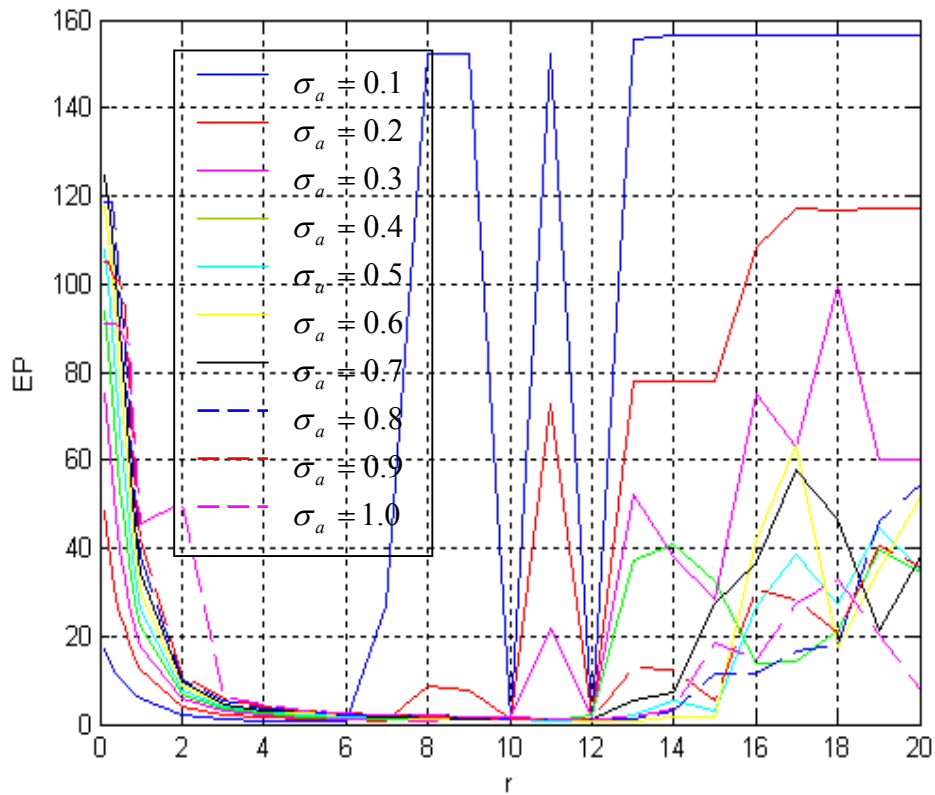


Fig. 5.9 - EP obtidos considerando $\sigma_{oa} = 0.01$

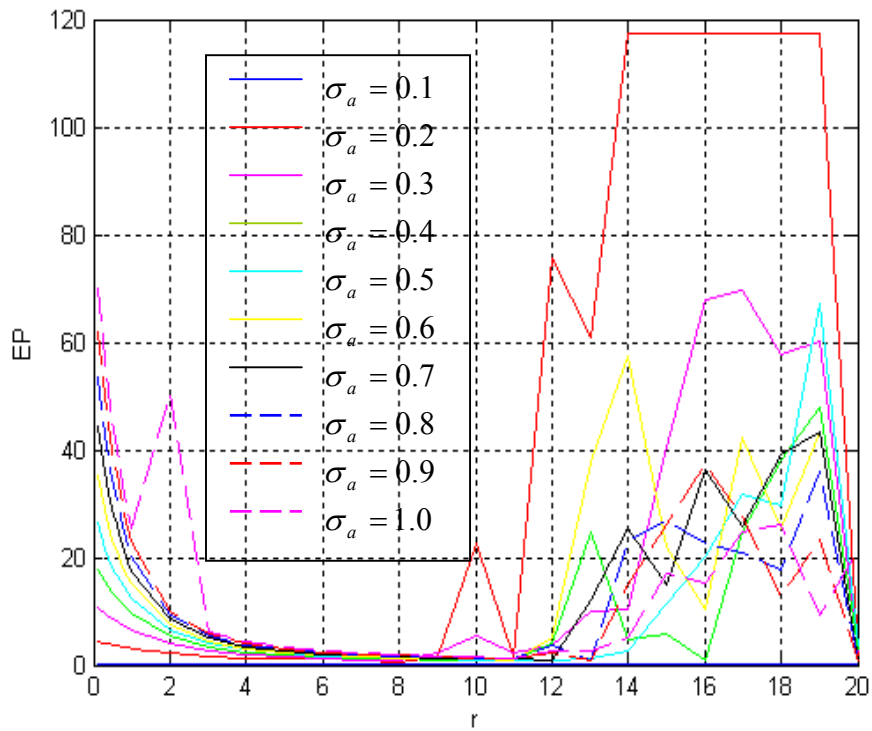


Fig. 5.10 - EP obtidos considerando $\sigma_{oa} = 0.1$

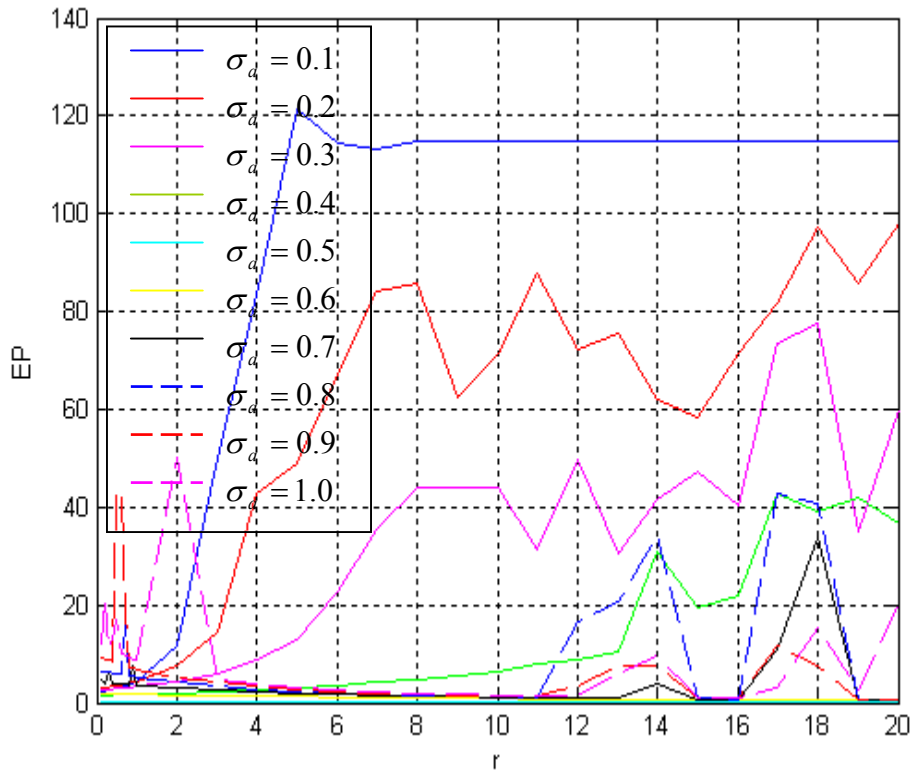


Fig. 5.11 - EP obtidos considerando $\sigma_{oa} = 0.5$

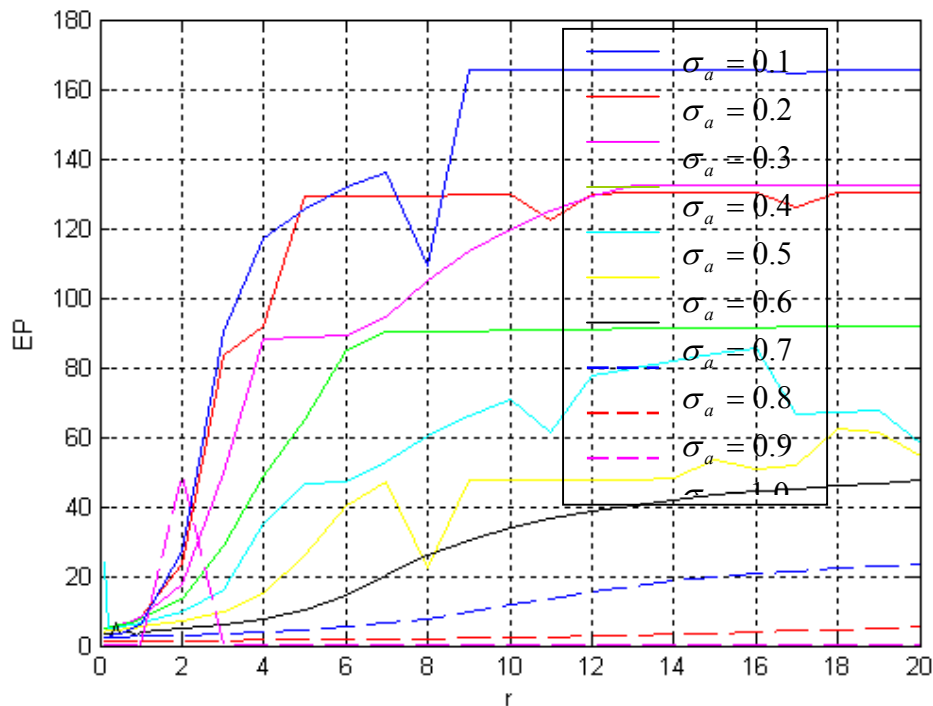


Fig. 5.12 - EP obtidos considerando $\sigma_{ao} = 1.0$

Das Figs. 5.7 ao 5.12, confirma-se que existe uma relação entre os dados iniciais e o EP obtido na estimativa de cada σ_a com diferentes r . A tabela 5.5 apresenta os valores de r e de σ_{ao} que ofereceram menor EP. É preciso comentar o seguinte; tentou-se trabalhar com σ_{ao} maiores que 1.0 porém sem sucesso, pois para algumas estimativas de σ_a o computador travava. Um resumo da tabela 5.5, seria a seguinte recomendação: Para a estimativa de σ_a se $\sigma_{ao} \ll \sigma_a$ sugere -se empregar r maiores a 2; e se $\sigma_{ao} > \sigma_a$ sugere-se empregar r pequenos.

Tabela 5.5- Dados iniciais e r ótimos para a estimativa dos σ_a

Valores iniciais	Valores de r
$\sigma_{ao} = 0.0001$	$3 < r < 5$ e $7 < r < 17$
$\sigma_{ao} = 0.001$	$4 < r < 5$
$\sigma_{ao} = 0.01$	$4 < r < 6$
$\sigma_{ao} = 0.1$	$5 < r < 9$
$\sigma_{ao} = 0.5$	1
$\sigma_{ao} = 1.0$	0.5

Com esta recomendação nós construímos o seguinte algoritmo para a estimativa do coeficiente de absorção com dados exatos, que vai ser resolvido considerando dois passos.

5.1.3.3 Algoritmo 1. Meio homogêneo

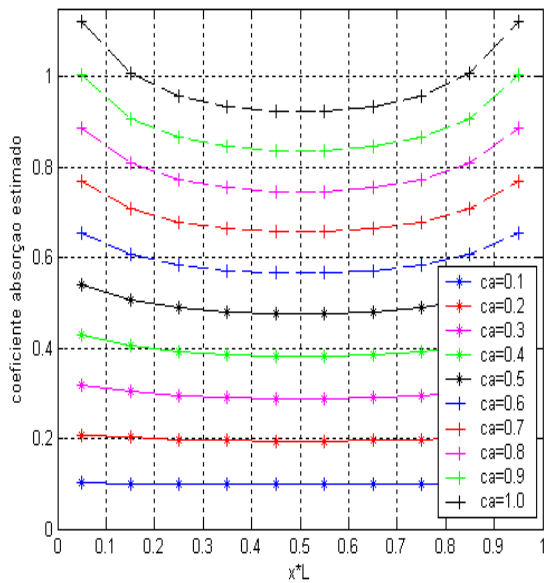
Passo 1.- *Minimizar* $D(\sigma_a, \sigma_{ao})$ sujeito a $F(\sigma_a)$, com $s=1, r=3$ e $\sigma_{ao} = 0.0001$

Passo 2. Seja σ_a os valores obtidos do passo 1, com todos eles calculamos o valor médio $\overline{\sigma_a}$. e

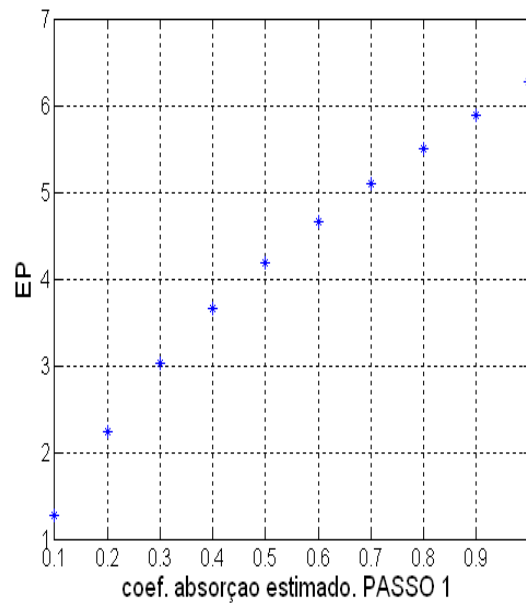
Passo 3.- *Minimizar* $D(\sigma_a, \sigma_{ao})$ sujeito a $F(\sigma_a)$, com $s=1, r=0.2$ e $\sigma_{ao} = \overline{\sigma_a}$

Exemplo 1. Meio homogêneo. Dados exatos

Neste exemplo empregando o algoritmo 1, intenta-se estimar σ_a para dez meios homogêneos, em todos os casos σ_s é conhecido e iguais a 0.5. As Figs. 5.13a e 5.13b apresentam os σ_a e EP obtidos com o passo 1. As Figs. 5.14a e 5.14b apresentam os σ_a e EP obtidos com o passo 1- 3 . Nas Figs. 5.13a e 5.14a apresentam-se dez curvas, cada curva apresenta os σ_a estimado no longo do meio. O valor procurado em cada curvo é dado na legenda. Observa-se nas Figs. 5.13b e 5.14b uma redução do EP.



a) Estimativa de $\sigma_a = ca$ na legenda



b) Estimativa do EP para cada σ_a

Fig. 5.13 – Estimativa σ_a com o passo 1. Exemplo 1 dados exatos

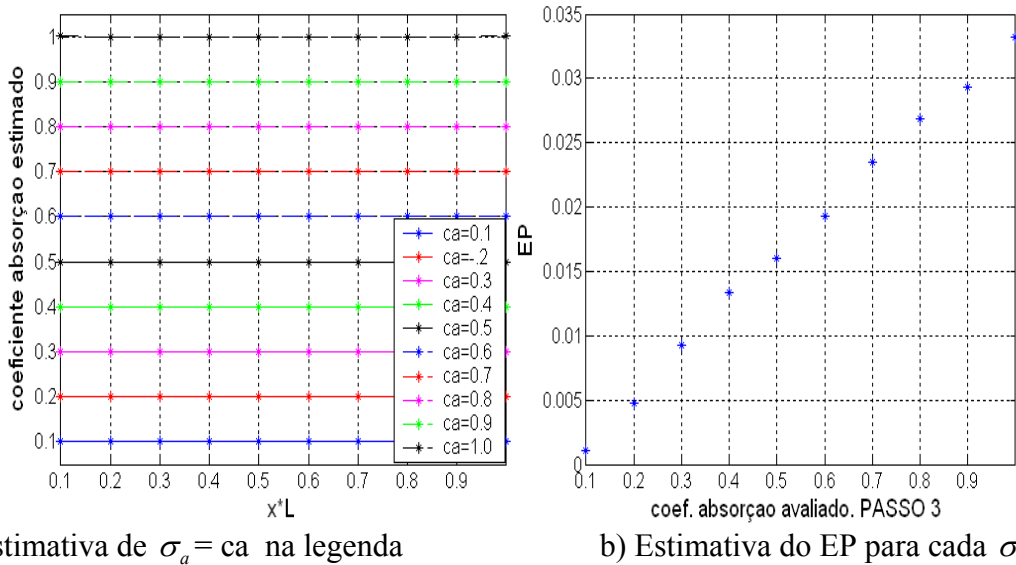


Fig. 5.14 – Estimativa σ_a com o passo 3. Exemplo 1, dados exatos

Exemplo 2. Meio homogêneo com ruído randômico

Considere-se um meio com $\sigma_a = 0.9$ e $\sigma_s = 0.5$. Tratando de simular medidas experimentais dos detectores (10 casos testes) como apresenta a Fig. 5.1, adiciona-se um ruído randômico até 3% à medida exata, para logo poder estimar o coeficiente de absorção.

Na Fig. 5.15 são apresentadas o coeficiente de absorção médio ($\bar{\sigma}_a$) e o erro percentual (EP) calculados considerando as seguintes discretizações: na Fig.5.15a uma discretização espacial ((DE) igual a 10, na Fig. 5.15b considera-se D.E=4, na Fig. 5.15c considera-se D.E=6, na Fig. 5.15d considera-se D.E=12, na Fig. 5.15e considera-se D.E=16, na Fig. 5.15f considera-se D.E=20.

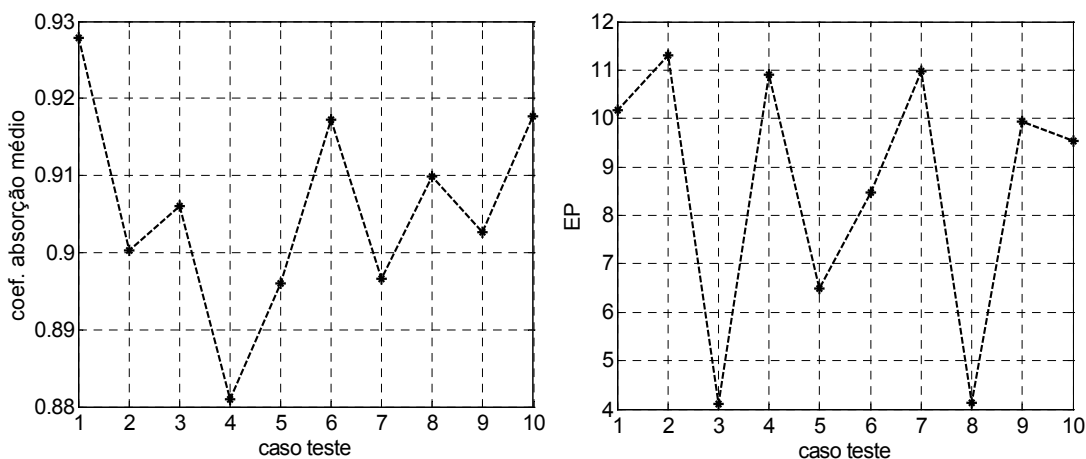


Fig. 5.15a – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=10.

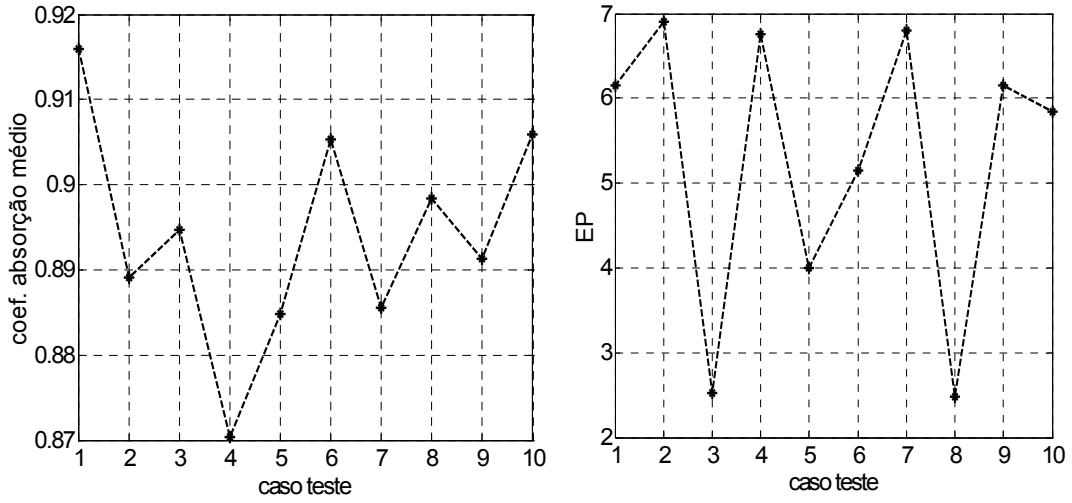


Fig. 5.15b – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=4.

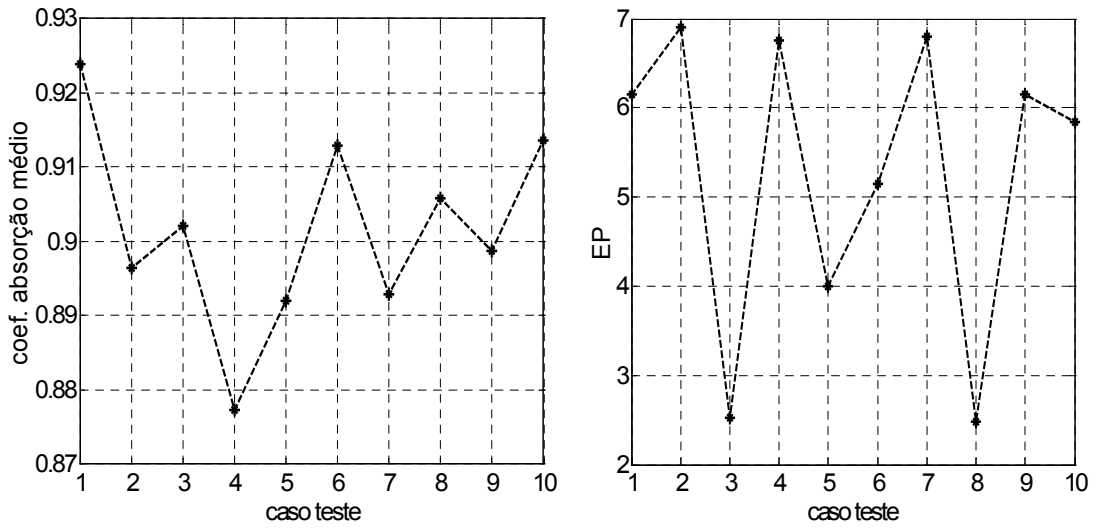


Fig. 5.15c – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=6.

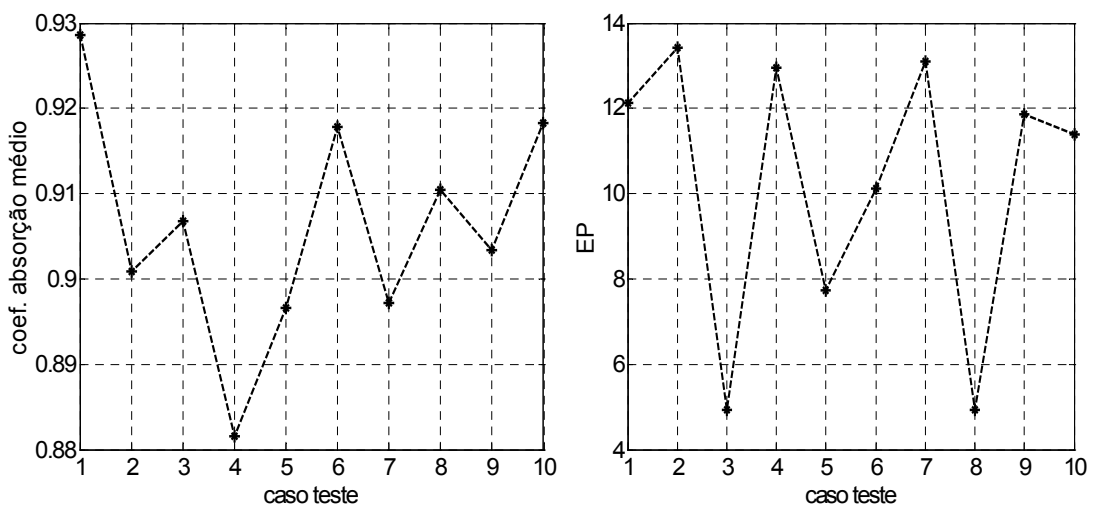


Fig. 5.15d – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=12

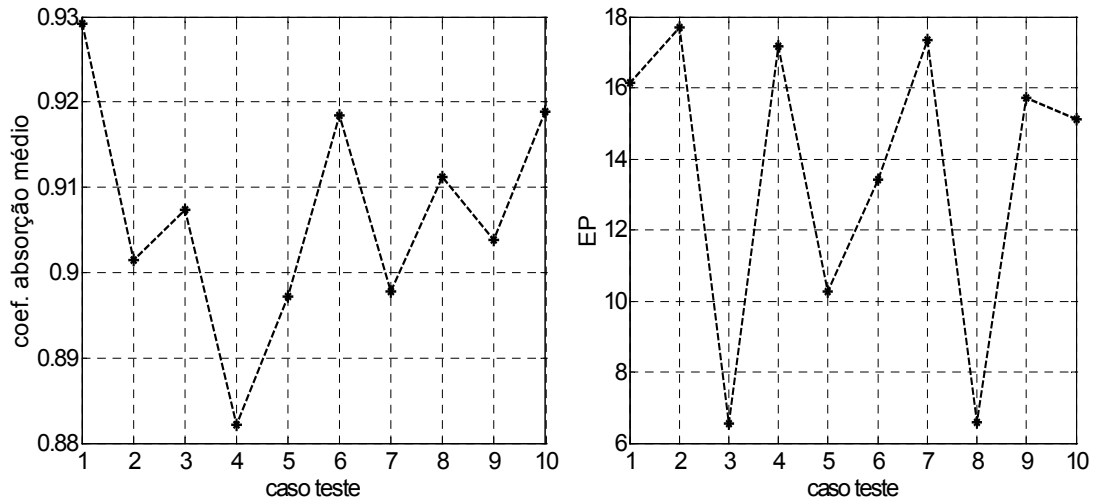


Fig. 5.15e – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=16

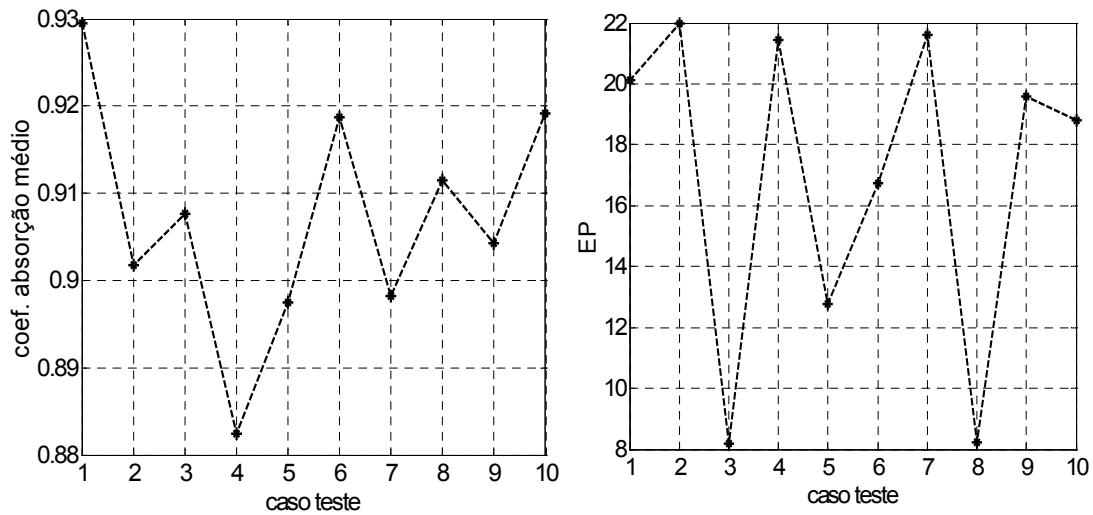


Fig. 5.15f – Estimativa de σ_a . Exemplo 2, meio homogêneo, com DE=20

5.1.3.4 Algoritmo 2. Estimativa de σ_a para um meio formado por duas placas

Consideremos um meio formado por duas placas paralelas. Sendo conhecida as espessuras das placas, é possível determinar os coeficientes de absorção que as formam com o seguinte algoritmo:

Passo 1.- *Minimizar* $D(\sigma_a, \sigma_{oa})$ sujeito a $F(\sigma_a)$, com $s=1$, $r=3$ e $\sigma_{oa} = 0.0001$

Passo 2.- Considerando as espessuras das placas dividimos em intervalos as discretizações que formam o meio em estudo. Com os σ_a obtidos do passo anterior, calculamos as médias para cada intervalo. Os novos valores iniciais são os $\overline{\sigma_a}$ calculados para cada intervalo.

Passo 3-. *Minimizar* $D(\sigma_a, \sigma_{oa})$ sujeito a $F(\sigma_a)$, com $s=1$, $r=0.2$. e considerando para cada intervalo $\sigma_{oa} = \overline{\sigma_a}$ calculado do passo 2.

Passo 4 – Repetir o passo 2 e 3 , cinco vezes.

Este algoritmo foi testado com sucesso para todas as combinações possíveis na espessura das placas e com diferentes σ_a , no intervalo (0.1, 0.2,...,1.0).

Exemplo 3- Meio de duas placas. Dados exatos

Considera-se um meio formado segundo a Fig. 5.16. O problema consiste em determinar os coeficientes de absorção. As Figs. 5.17 e 5.18 apresentam os coeficientes de absorção estimados e o erro percentual obtidos em cada repetição dos passos 2 e 3.

Exemplo 4. Meio de duas placas Dados com ruído randômico.

As Figs. 5.19, 5.20 e 5.21 apresentam os erros percentuais de 10 casos-testes, onde foi adicionado ruído randômico até 1.5% do valor exato nos medidas dos detectores, para um meio com a mesma geometria da Fig. 5.16. A placa 1 tem o coeficiente de absorção igual a 0.3 e a placa 2 com coeficiente de absorção igual a 0.7.

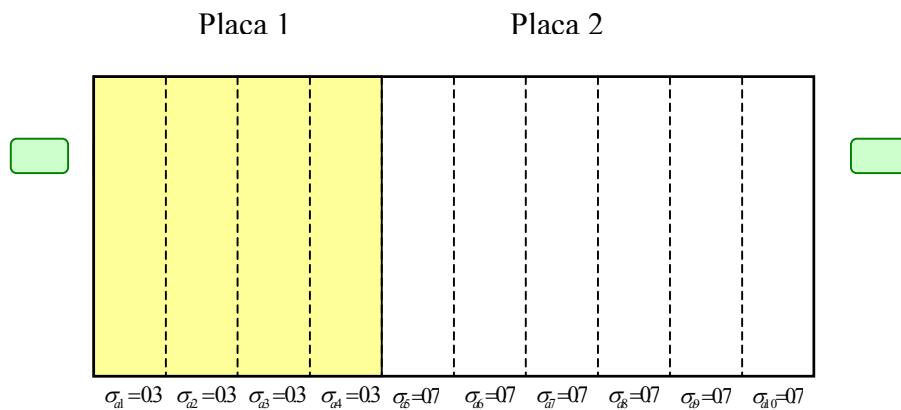


Fig. 5.16- Geometria do problema heterogêneo. Exemplo3

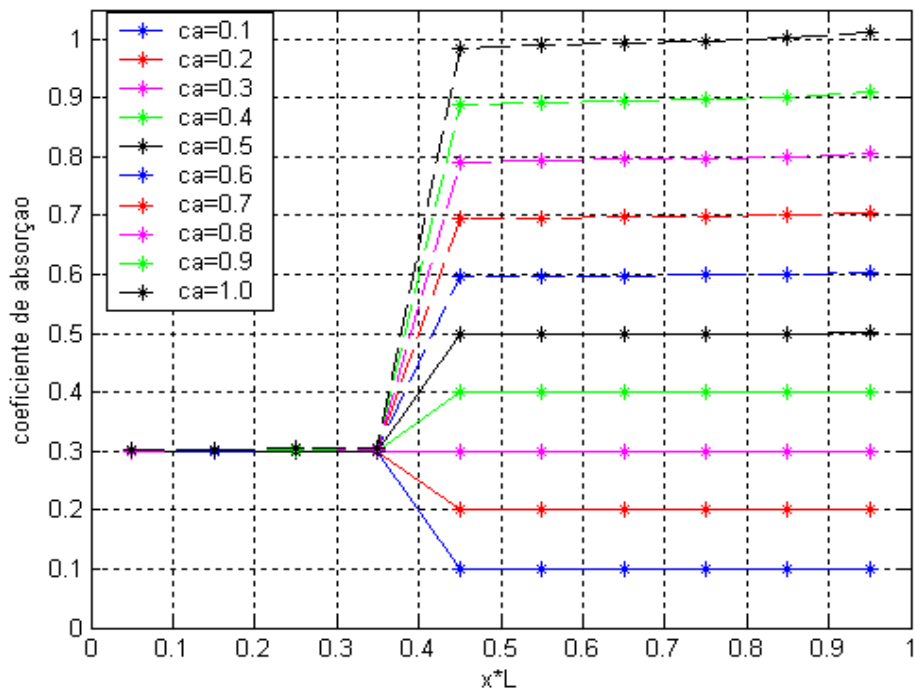


Fig. 5.17- Exemplo 3 meio formado por 2 placas. Dados exatos.

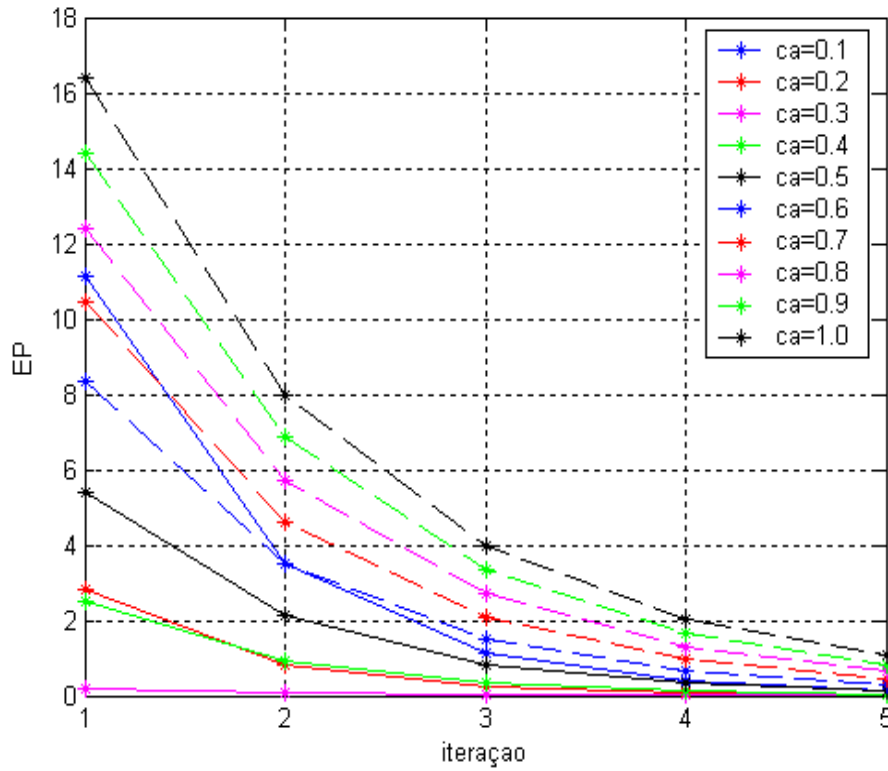


Fig. 5. 18 -Exemplo 3, EP per cada iteração

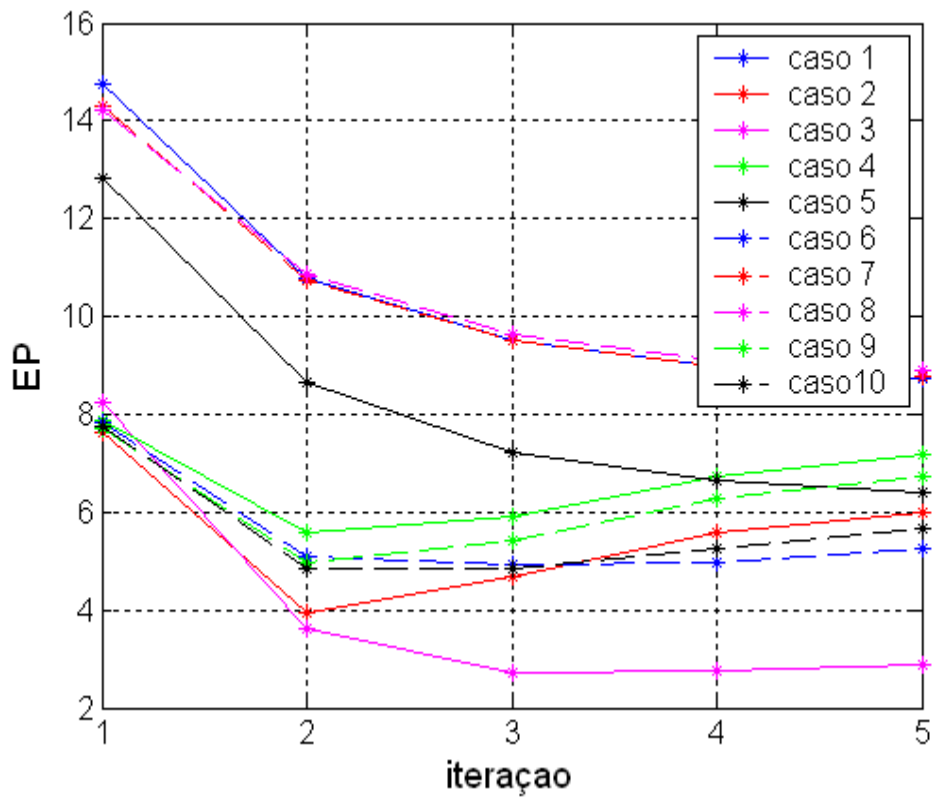


Fig. 5.19 – Estimativa dos σ_a . Exemplo 4, dados com ruído randômico ate 1.5%.

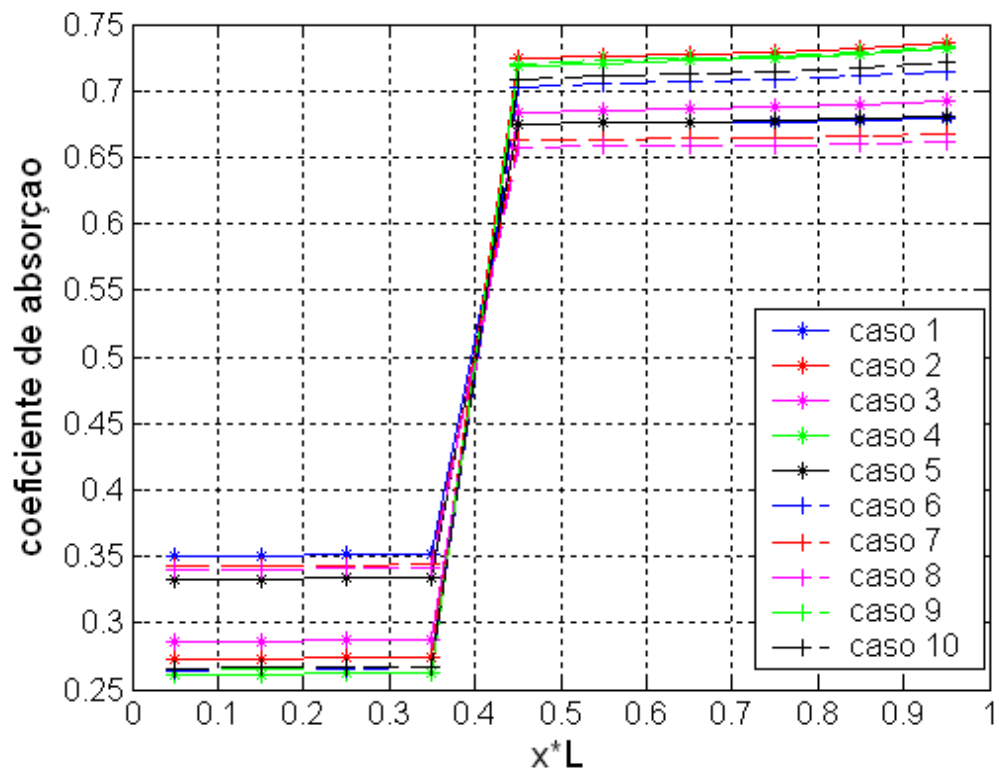


Fig. 5.20 – EP versus iteracao. Exemplo.4, dado com ruído ate 1.5%.

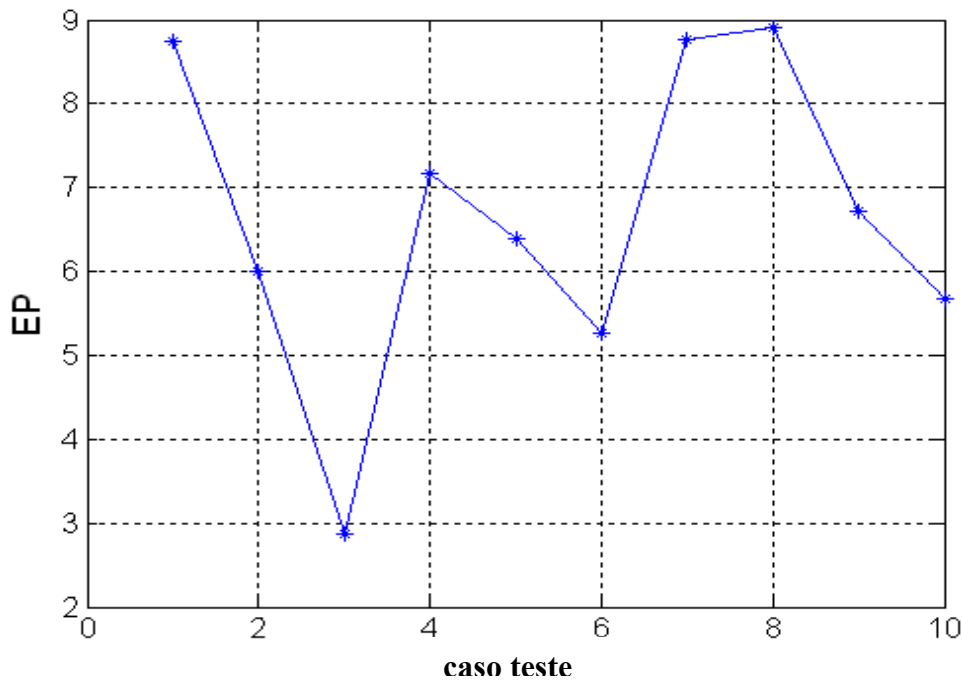


Fig. 5.21 - O EP obtido em cada caso teste. Exemplo 4

5.1.3.5.- Algoritmo 3 Meio formado por dois materiais. Estimativa de espessuras

Tendo como conhecido os valores do coeficiente de absorção e espalhamentos das duas placas que formam um meio, se vai estimar as espessuras das placas. O algoritmo empregado neste problema é apresentado a seguir:

Passo 1.-Definimos os parâmetros $\sigma_{a\min}$, $\sigma_{a\max}$, como os coeficientes de absorção das placas que formam o meio. Se $\sigma_{a\min} < \sigma_{a\max}$, e definimos $\sigma_{amed} = (\sigma_{a\max} + \sigma_{a\min})/2$. definimos $no=10$ (discretização do meio). Logo:

Minimizar $D(\sigma_a, \sigma_{ao})$ sujeito a $F(\sigma_a)$, com $s=1$, $r=1$ e $\sigma_{ao} = \sigma_{ao\max}$

Passo 2.- Com os valores de σ_a estimados do passo 1, calculamos a media $\overline{\sigma_a}$

Passo 3.- *Minimizar* $D(\sigma_a, \sigma_{oa})$ sujeito a $F(\sigma_a)$, com $s=1$, $r=0.2$. Para cada elemento discretizado i , se faz:

Se $\sigma_{ai} < \overline{\sigma_a}$; $\sigma_{aoi} = \sigma_{a\min}$; caso contrario $\sigma_{aoi} = \sigma_{a\max}$

Passo 4.- $B = 0$, $h = 0$

Enquanto $B = 0$, faça:

$h = h + 1$

$\sigma_a^h = \sigma_a$,

Para $i = 1$ até no

Se $\sigma_{ai} < \sigma_{oamed}$, $\sigma_{ai} = \sigma_{oa\min}$; caso contrario $\sigma_{oi} = \sigma_{oa\max}$, fim se

Fim para

Logo:

Para $i = 1$ até $no - 1$

Se $|\sigma_{oamed} - \sigma_{ai}| > |\sigma_{oamed} - \sigma_{ai+1}| \cdot 1.25$ então $\sigma_{oi+1} = \sigma_{oi}$, fim se

Se $|\sigma_{oamed} - \sigma_{ai}| \cdot 1.25 < |\sigma_{oamed} - \sigma_{ai+1}|$ então $\sigma_{oi} = \sigma_{oi+1}$, fim se

Fim para

Minimizar $D(\sigma_a, \sigma_{oa})$ sujeito a $F(\sigma_a)$, com $s=1$, $r=0.2$.

Se $\sum_{i=1}^{no} |\sigma_{ai} - \sigma_{ai}^h| < 10^{-4}$ então $B = 1$

Fim

Exemplo 5 – A espessura de duas placas juntas

Neste exemplo emprega-se uma geometria similar à Fig. 5.16. Os coeficientes de absorção das placas são $\sigma_{a\min} = 0.3$, $\sigma_{a\max} = 0.7$, então os valores iniciais para o PASSO 1 é $\sigma_{ao} = 0.5$. As Figs. 5.22 e 5.23 apresentam a geometria estimada do meio e o E_p obtido para cada iteração.

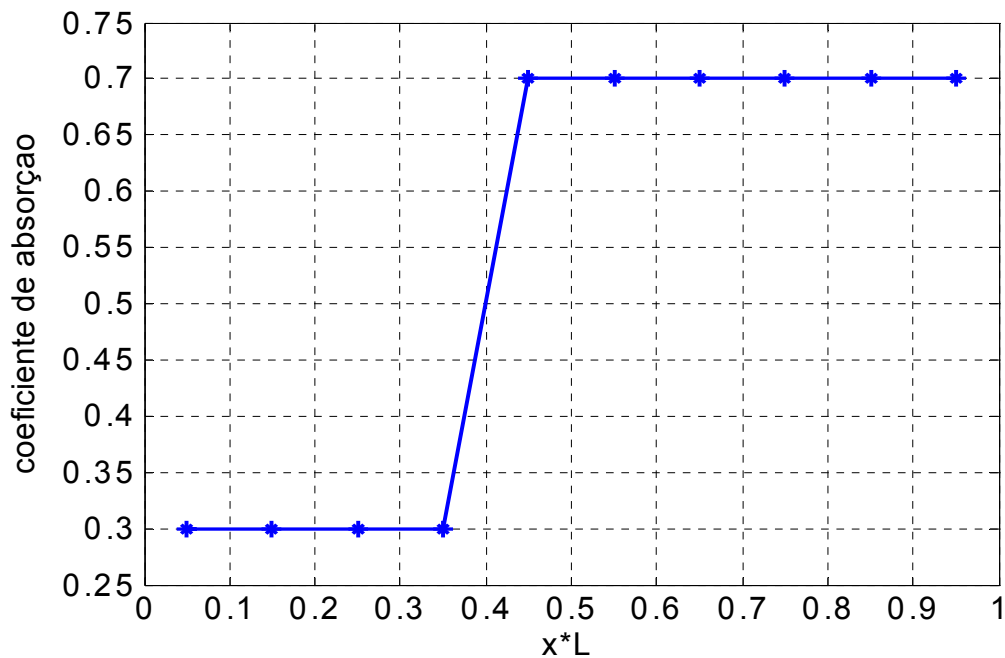


Fig. 5.22.- Estimativa da geometria no exemplo 5.

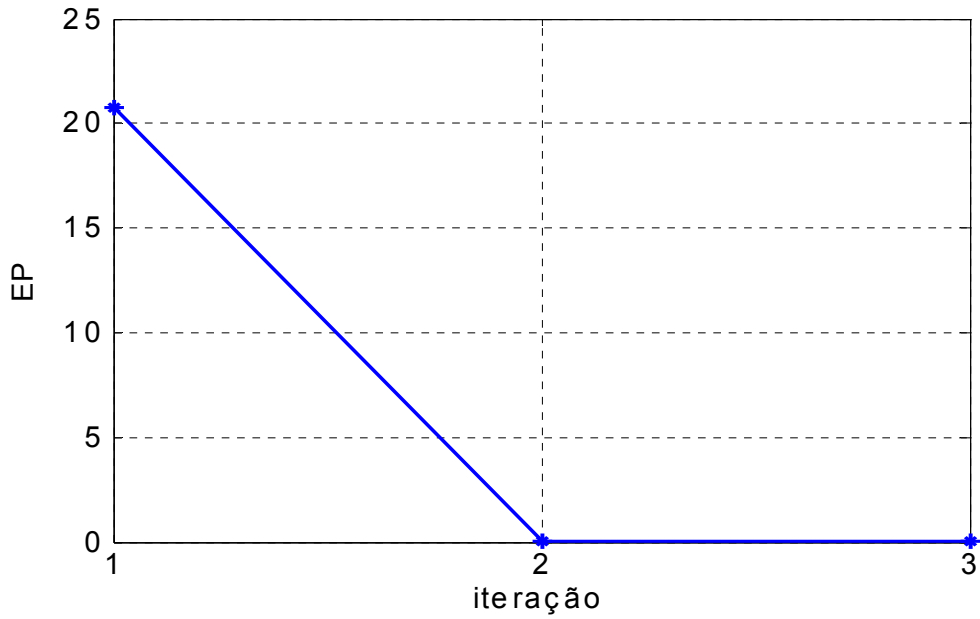


Fig. 5.23.- Estimativa do EP em cada iteração do exemplo 5.

Exemplo 6 – Estimativa da espessura de duas placas. Dado com ruído ate 1%.

Neste exemplo emprega-se uma geometria dada pela Fig. 5.22, para 10 casos –testes onde é adicionado ruído ate 1%. A Fig. 5.24 apresenta as geometrias estimadas. A Fig. 5.25 apresenta o EP calculado em cada iteração.

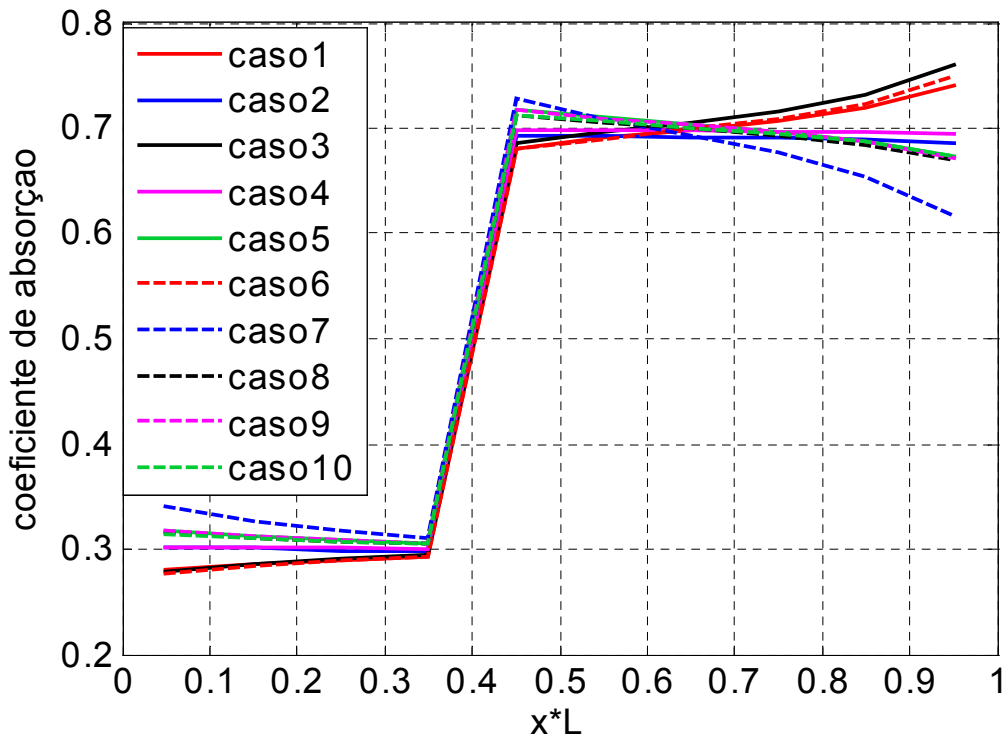


Fig. 5.24- Geometria do meio estimado. Exemplo 6.

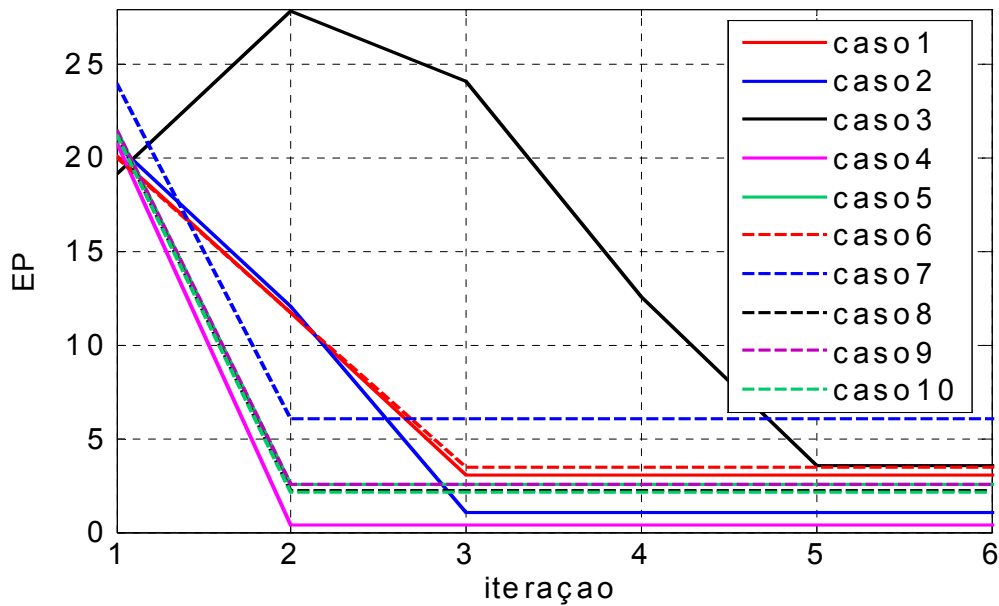


Fig. 5.25- O EP por cada iteração. Exemplo 6.

5.1.3.6.- Estimativa do coeficiente de espalhamento

Exemplo 7.- Meio homogêneo, empregando o algoritmo1.

Considerando σ_a conhecido e igual a 0.5, se vai estimar o coeficiente de espalhamento, em 10 meios com diferentes σ_s entre 0.1 e 1.0 (vide legenda da Fig. 5.26). A figura 5.26 apresenta, os valores estimados de σ_s e o erro percentual obtidos com o passo 1 definido para meios homogêneos. O passo 3 torna o EP menor.

Em poucos caso- teste logrou-se estimar o σ_s para um meio heterogêneo.

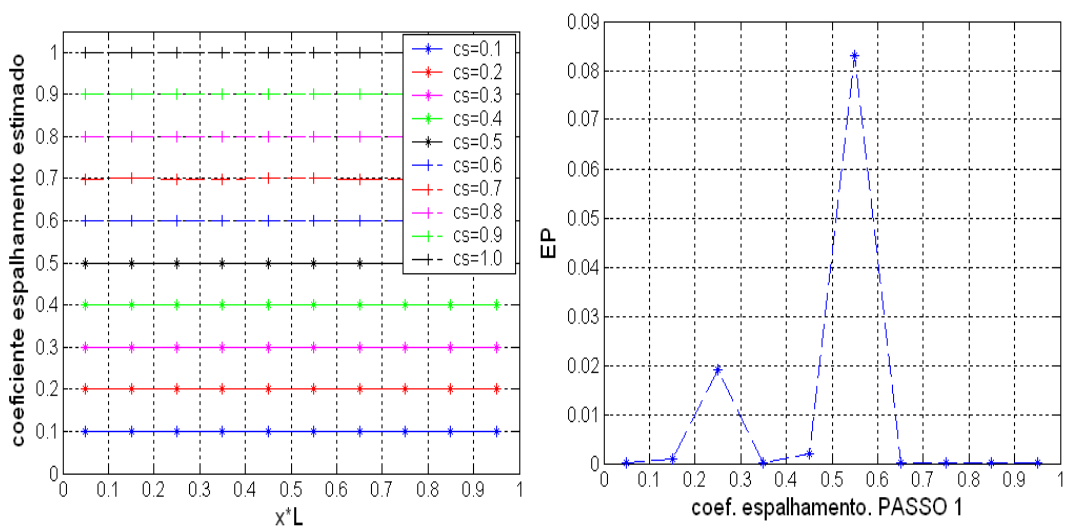


Fig. 5.26.- Estimativa σ_s como o passo 1. Exemplo 7.

5.1.4 Estimativa do coeficiente de absorção empregando a função $\eta_r(z) =$

$$\frac{1}{r-1} \ln \sum_{j=1}^i z_j^r \text{ na construção da distância de Bregman.}$$

5.1.3.1 Meio homogêneo

Na Fig. 5.27-, apresentamos os EP gerados ao solucionar o problema inverso empregando o funcional $\eta_r(z)$, variando o parâmetro r no intervalo: $0 < r < 1$. Cada Fig. apresenta dez curvas que indicam um σ_a diferente a ser estimado, segundo a legenda. Os valores iniciais σ_{ao} empregados foram para a Fig. 5.27a - $\sigma_{ao}=0.0001$, a Fig. 5.27b - $\sigma_{ao}=0.001$, a Fig. 5.27c - $\sigma_{ao}=0.01$, a Fig. 5.27d - $\sigma_{ao}=0.1$, a Fig. 5.27e - $\sigma_{ao}=0.5$. Para $\sigma_{ao} > 0.5$, o computador não conseguiu estimar para todos os σ_a .

De todas estas Figs. observa-se que os menores EP são obtidos com $\sigma_{ao}=0.5$.

observa-se que o funcional $\eta_r(z) = \frac{1}{r-1} \ln \sum_{j=1}^i z_j^r$ não oferece melhores estimativas que

$$\text{o funcional } \eta_{sr}(z_i) = \frac{(z_i^s - z_i^r)}{s-r}.$$

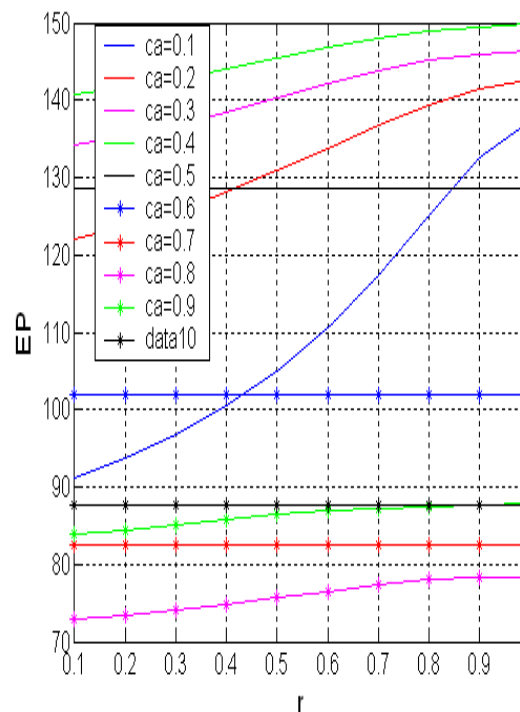
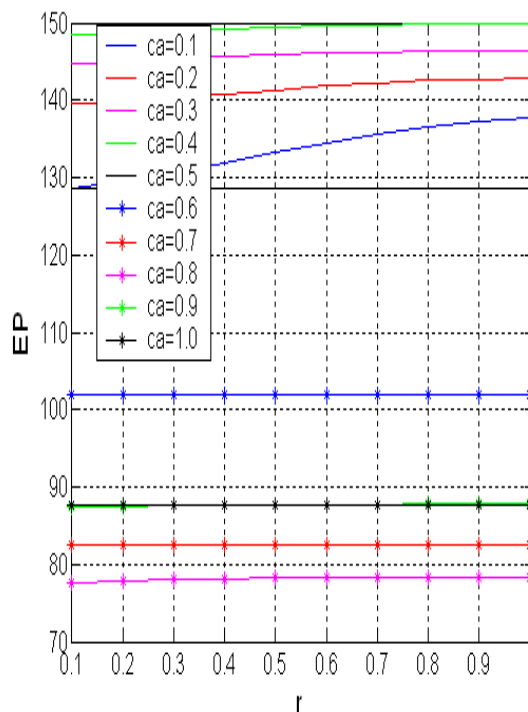


Fig. - 5.27a. EP versus r , com $\sigma_{ao}=0.0001$ Fig. 5.27b - EP versus r , com $\sigma_{ao}=0.001$

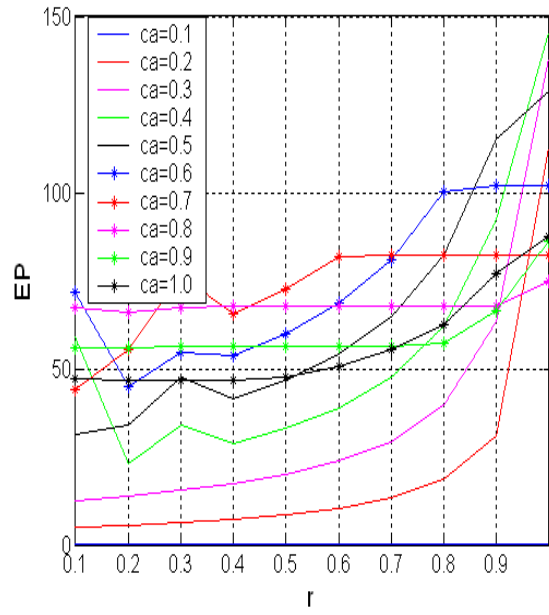
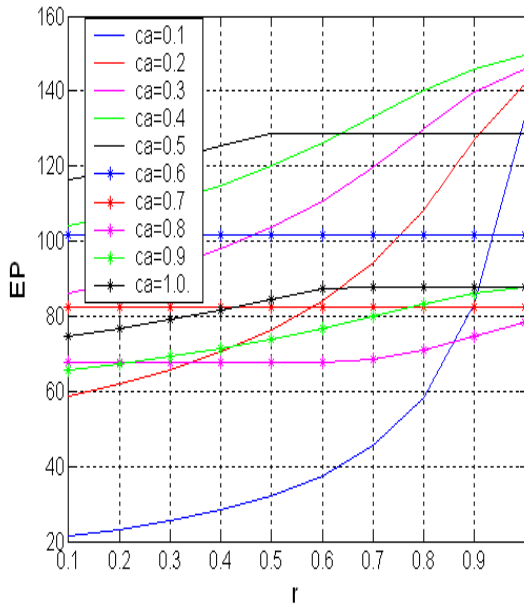


Fig. 5.27c - EP versus r , com $\sigma_{ao}=0.01$

Fig. 5.27d - EP versus r , com $\sigma_{ao}=0.1$

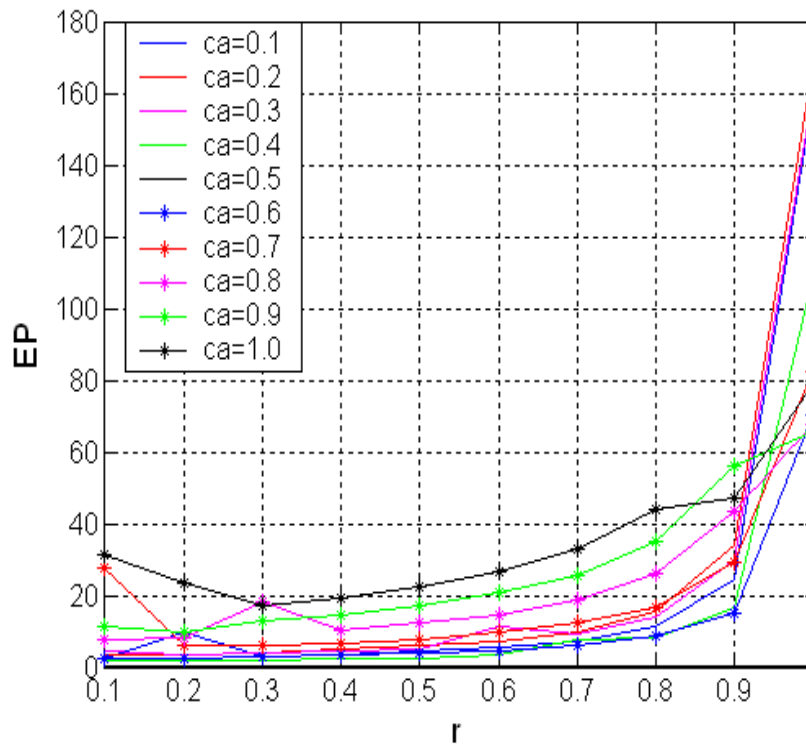


Fig. 5.27e. EP versus r , com $\sigma_{ao}=0.5$

5.1.5 Estimativa do coeficiente de absorção empregando a função $\eta_B(z) = -\ln z_i$ na construção da distância de Bregman.

5.1.5.1 Meio homogêneo

Na Fig. 5.28, apresentamos os EP gerado ao solucionar o problema inverso empregando o funcional $\eta_B(z)$. A Fig. apresenta oito curvas, que se diferenciam por ter sido utilizados diferente dados iniciais σ_{ao} . A legenda na Fig. 5.28, indica quais foram os valores utilizados. Em cada curva apresentam-se os EP estimados para dez meios com diferentes σ_a , como indica o eixo x. Observe-se que os EP menores são obtidos quando o dado inicial está próximo do dado a ser estimado. Para $\sigma_{ao}=2$ ou $\sigma_{ao} > 3$, o EP se incrementa a mais de 10^4 para alguns valores de σ_a a serem estimados, sendo esta a razão por que não foram considerados na Fig. 5.28

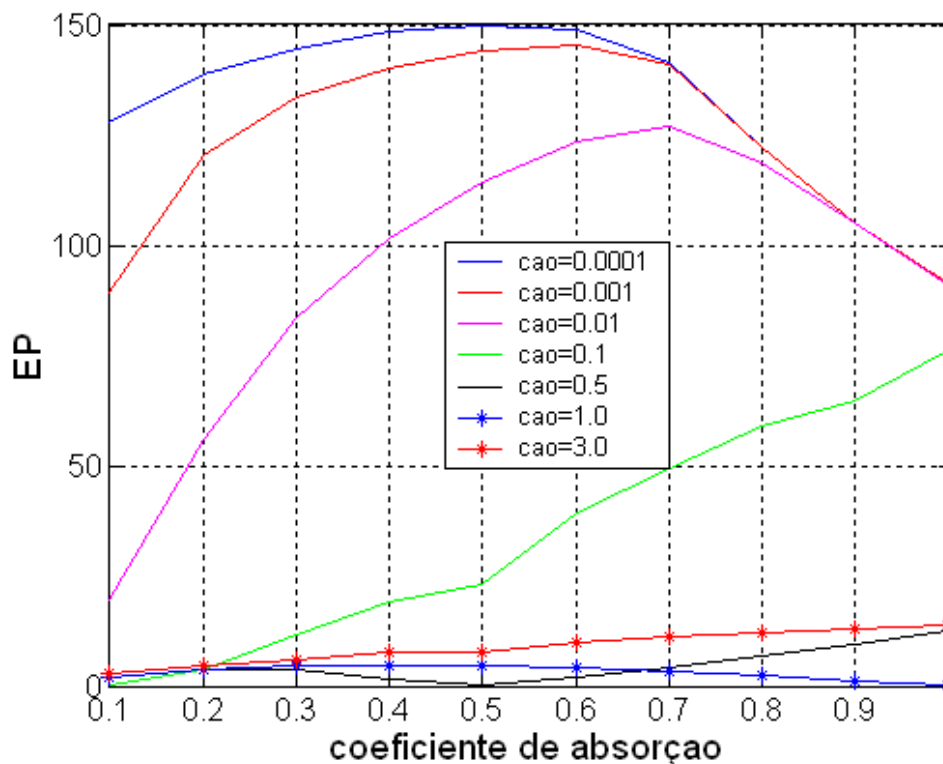


Fig.. 5.28. EP gerados com o funcional $\eta_B(z)$, com diferentes σ_a

5.2. Problema Bidimensional.

5.2.1 Validação do Problema direto para um meio bidimensional

A validação do problema direto para um meio bidimensional, empregada nesta tese, foi apresentada na minha tese de mestrado, (BERROCAL, 2001).

5.2.2 Localização dos detectores para a estimativa do coeficiente de absorção e espalhamento

Para a estimativa dos coeficientes de absorção e espalhamento em um meio bidimensional, simulamos medidas experimentais de duas geometrias, relacionadas à localização das fontes e detectores no meio de estudo.

5.2.2.1 Geometria 1

A Fig. 5.29 apresenta a Geometria 1, observa-se que as medidas devem ser obtidas uma a uma, uma fonte e dois detectores D^t e D^r , como no caso unidimensional. O detector D^t mede o fluxo transmitido q^t e o detector D^r mede o fluxo refletido na fronteira q^r . Segundo a localização do detector os fluxos q^t e q^r são calculados das equações:

$$q_{(i \max, j)}^t = \sum_{\substack{l=1 \\ \mu_l > 0}}^{Lo} w_l \cdot Ie_{l(i \max, j)} \cdot \mu_l \quad q_{(1, j)}^r = \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} - w_l \cdot Iw_{l(1, j)} \cdot \mu_l \quad (5.5)$$

$$q_{(i, j \max)}^t = \sum_{\substack{l=1 \\ \xi_l > 0}}^{Lo} w_l \cdot In_{l(i \max, j)} \cdot \xi_l \quad q_{(i, 1)}^r = \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} - w_l \cdot Is_{l(1, j)} \cdot \xi_l \quad (5.6)$$

$$q_{(1, j)}^t = \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} - w_l \cdot Iw_{l(1, j)} \cdot \mu_l \quad q_{(i \max, j)}^r = \sum_{\substack{l=1 \\ \mu_l > 0}}^{Lo} w_l \cdot Ie_{l(i \max, j)} \cdot \mu_l \quad (5.7)$$

$$q_{(i, 1)}^t = \sum_{\substack{l=1 \\ \xi_l < 0}}^{Lo} - w_l \cdot Is_{l(i \max, j)} \cdot \xi_l \quad q_{(i, j \max)}^r = \sum_{\substack{l=1 \\ \xi_l > 0}}^{Lo} w_l \cdot In_{l(i \max, j)} \cdot \xi_l \quad (5.8)$$

Os valores empregados para μ_l são os mesmos que no caso unidimensional, LSH S_{10} dada na tabela 3.1, considera-se $\xi_l = \mu_l$.

Esta geometria permite a estimativa do coeficiente de absorção com maior aproximação que a estimativa do coeficiente de espalhamento. Pois a perda da radiação por absorção se dá em cada elemento de área que forma o caminho da radiação fonte-detector, em quanto à perda e contribuição da radiação por espalhamento se dá entre os elementos que formam o caminho do meio e os elementos vizinhos..

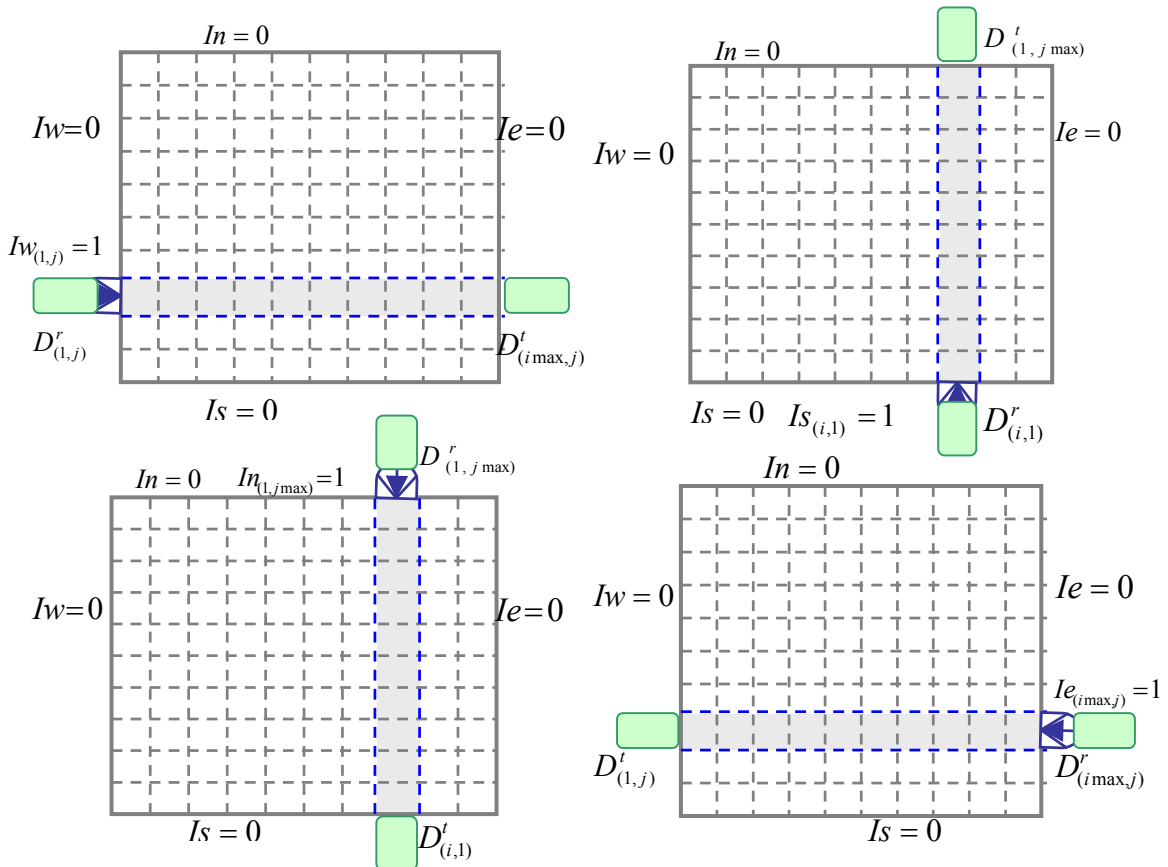


Fig. 5.29 - Posição das fontes e detectores para a estimativa dos σ_a e σ_s

5.2.2.2 Geometria 2

A Fig. 5.30 apresenta a geometria 2, observa-se só detectores D^r que medem fluxo refletido do meio. Não é possível medir o fluxo transmitido devido ao aparecimento do efeito raio na solução do problema direto por ordenada discreta. Não se considera radiação externa que entre no meio que não das fontes. Com esta geometria só é possível fazer estimativa do coeficiente de espalhamento,

Segundo a localização da fonte e do detector nas faces do meio, vide Fig. 5.39, tem-se que o fluxo refletido medido é dado:

$$\begin{aligned}
q_{(i,1)}^r &= \sum_{\substack{l=1 \\ \xi_l < 0}}^{Lo} w_l \cdot Is_{l(i,1)} \cdot |\xi_l| & q_{(i,j_{\max})}^r &= \sum_{\substack{l=1 \\ \xi_l > 0}}^{Lo} w_l \cdot In_{l(i,j_{\max})} \cdot |\xi_l| \\
q_{(1,j)}^r &= \sum_{\substack{l=1 \\ \mu_l < 0}}^{Lo} w_l \cdot Iw_{l(1,j)} \cdot |\mu_l| & q_{(i_{\max},j)}^r &= \sum_{\substack{l=1 \\ \mu_l > 0}}^{Lo} w_l \cdot Ie_{l(i_{\max},j)} \cdot |\mu_l| \quad (5.9)
\end{aligned}$$

com $i = 1, 2, \dots, i_{\max}$; $j = 1, 2, \dots, j_{\max}$; os valores para $\xi_l = \mu_l$ são obtidos da tabela 3.2.

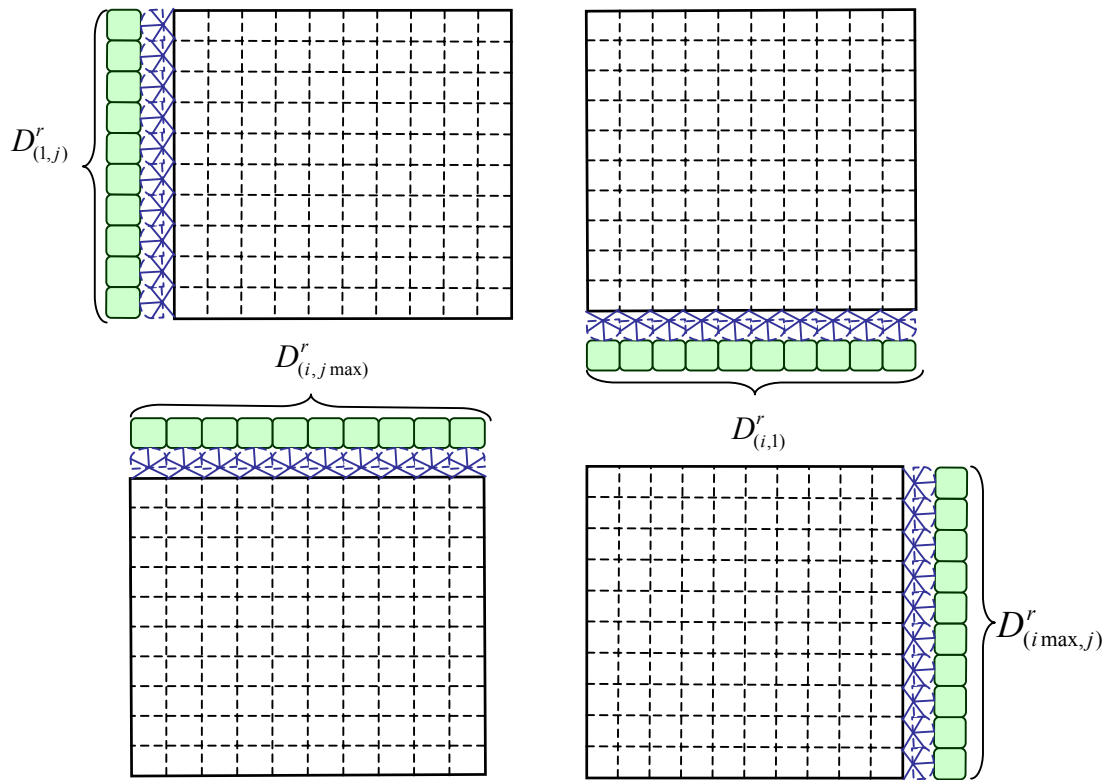


Fig. 5.30 - Posição das fontes e detectores para a estimativa dos σ_s

Outras considerações: na maioria dos casos testes aqui apresentados considera-se um médio particionado com uma malha quadrada de 10*10. A intensidade das fontes é 1.

5.2.3.- Estimativa do coeficiente de absorção e espalhamento, empregando $\eta_{1r}(z_i) =$

$\frac{(z_i - z_i^r)}{1-r}$ na construção da distância de Bregman.

Considere-se um meio discretizado por uma malha quadrada de 10 divisões por lado, com 100 elementos de áreas, 100 coeficientes de absorção e 100 coeficientes de espalhamento, um total de 200 incógnitas a estimar.

Na primeira geometria tem-se as medidas experimentais de 80 detectores, o que dá um sistema formado de 80 equações e 200 incógnitas a ser resolvido. Para isso, minimizamos a distância de Bregman $D_{1r}(\sigma, \sigma_0)$, construída com o funcional $\eta_{1r}(z_i) = \frac{(z_i - z_i^r)}{1-r}$, sujeito ao funcional erro $F(\sigma)$.

Trata-se de encontrar o melhor r que ofereça o menor EP, minimizamos com diferentes valores de r no intervalo de [0.2, 20]. Os coeficientes de absorção estimados com o melhor r são guardados no arquivo r1.tmp.

Ajudados com os dados do arquivo r1.tmp na segunda geometria tem-se um sistema formado por 40 equações e 100 coeficientes de espalhamentos a serem estimados. Neste sistema também se vai procurar o melhor r que ofereça o menor EP,. Os coeficientes de espalhamento estimados com o melhor r são guardados no arquivo r2.tmp.

Empregando como dados iniciais os valores dos arquivos r1.tmp e r2.tmp para σ_a e σ_s respectivamente, voltamos à primeira geometria. Tendo por conhecimento que existe uma relação entre os dados iniciais e o melhor valor de r que resolve um sistema, voltamos a procurar o melhor r que ofereça o menor EP. Os σ_a e σ_s estimados com o melhor r são guardados no arquivo r3.tmp.

Nos 3 casos testes seguintes emprega-se e comenta-se o procedimento descrito acima para a estimativa de σ_a e σ_s .

5.2.3.1 Meios homogêneo

Apresentamos a seguir 3 casos testes de meios homogêneos com σ_a e σ_s igual a 0.1, 0.5 e 0.8 respectivamente.

Considerando as medidas dos detectores obtidas da geometria 1. Na Fig. 5.31 apresentam-se as curvas obtidas dos erros percentuais EP gerados com diferentes r ao tratar de estimar σ_a para cada meio homogêneo. Na Fig. 5.32 apresentam-se os EP ao

tratar de estimar os σ_s com diferentes r . Na Figs. 5.33, apresenta-se o tempo computacional gasto na estimativa de σ_a e σ_s para cada r e em cada caso teste respectivo. A tabela 5. 6 apresentam os valores iniciais σ_{ao} e σ_{so} , os valores mínimos lb e máximos ub onde as incógnitas foram procuradas.

Para $\sigma_{ao} > 0.0001$, não foi possível construir a curva dos EP *versus* r , já que para vários valores de r o computador ficava calculando por varias horas, considero-se 6 horas como o tempo computacional máximo de espera para algum resultados.

Tabela 5.6, Valores iniciais, mínimos e máximos para a estimativa de σ_a e σ_s

	σ_{ao}	σ_{os}	Valor mínimo lb		Valor máximo ub	
			σ_a	σ_s	σ_a	σ_s
Fig. 5.40, $\sigma_a=0.1$ e $\sigma_s=0.1$	0.0001	0.0001	0.0001	0.0001	3.0	1.0
Fig. 5.41, $\sigma_a=0.5$ e $\sigma_s=0.5$	0.0001	0.0001	0.0001	0.0001	3.0	1.0
Fig. 5.42, $\sigma_a=0.8$ e $\sigma_s=0.8$	0.0001	0.0001	0.0001	0.0001	3.0	1.0

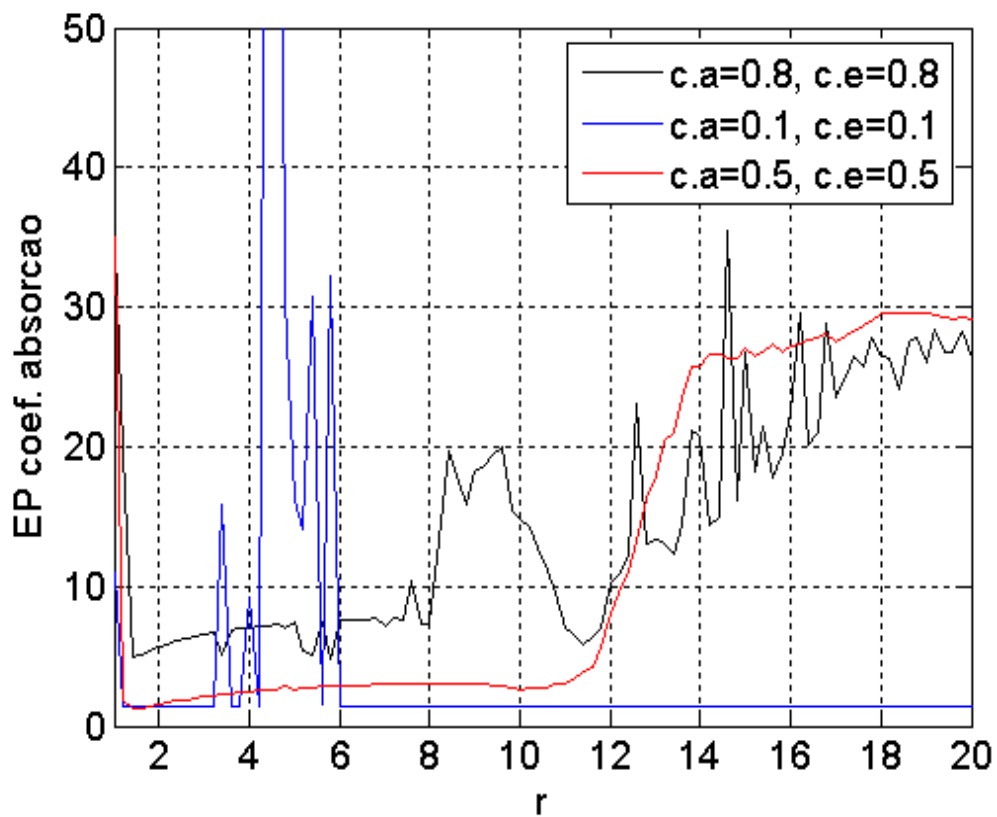


Fig. 5.31.- Erro percentual (EP) na estimativa de σ_a *versus* r

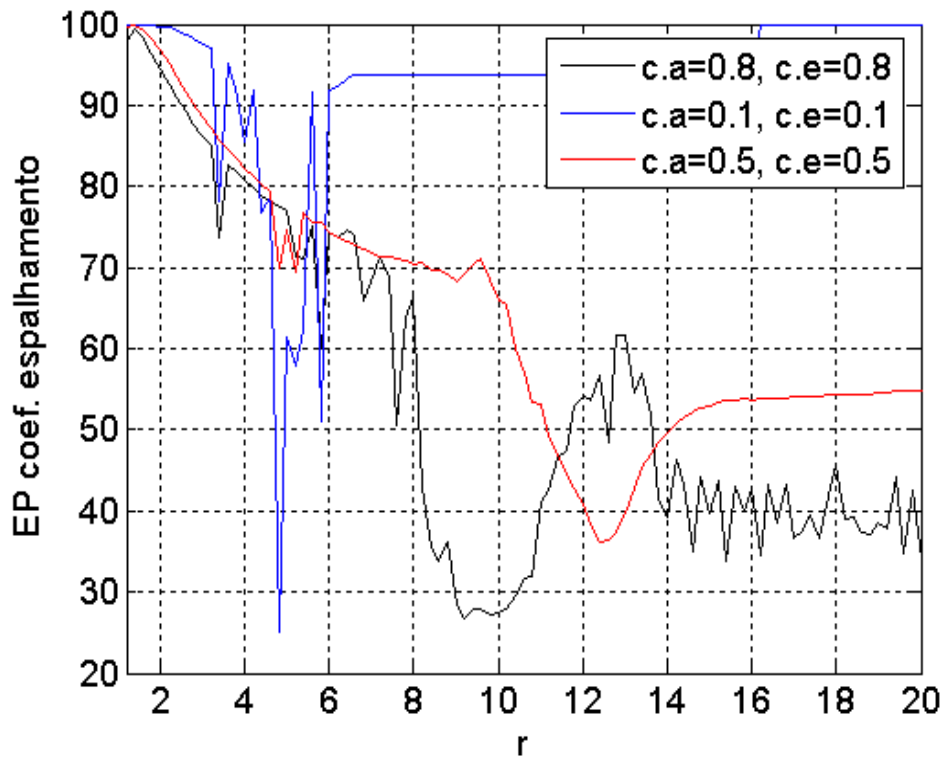


Fig. 5.32.- EP na estimativa de σ_s versus r

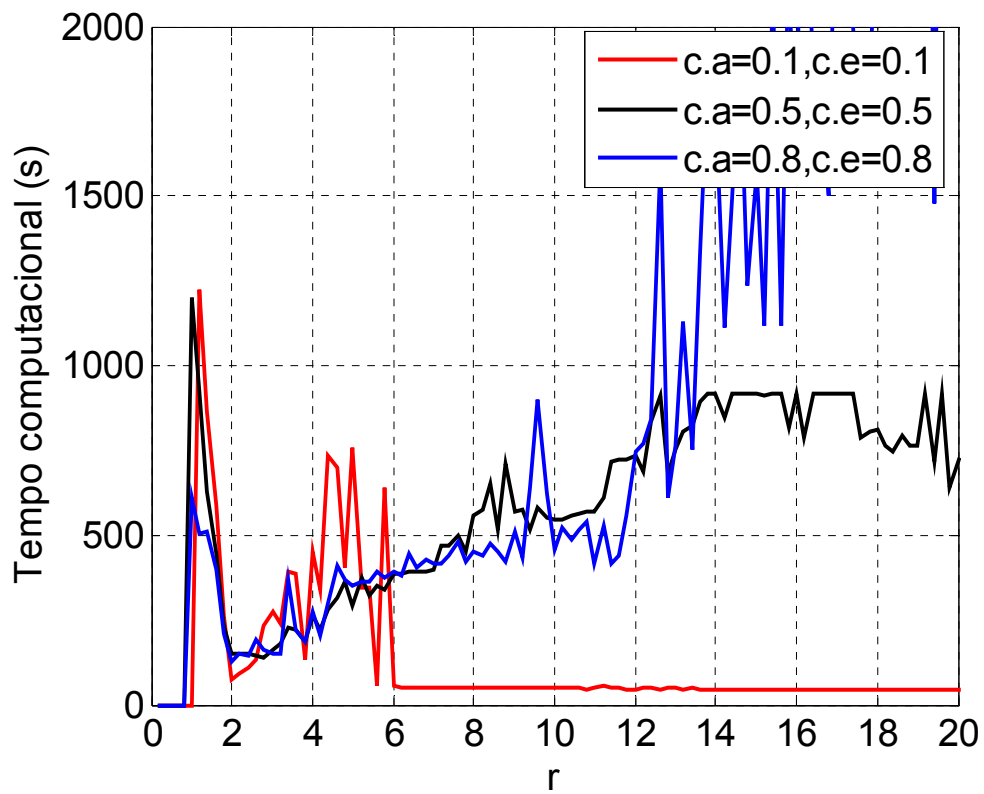


Fig. 5.33.- Tempo computacional gasto versus r

Nas Figs. 5.31, 5.32, e 5.33 observa-se que o menor erro percentual pode ser considerado para r no intervalo de [2-3].

Para $r < 1$ não foi possível obter resultados, o tempo computacional excedia a 6 horas.

Como os erros percentuais para a estimativa de espalhamento são altos, estes valores não vão ser considerados como dados iniciais do programa que resolve a geometria 2.

Para os três casos teste, os valores estimados dos σ_a com $r=2$ na geometria 1 são empregado para a estimativa do coeficiente de espalhamento. Na geometria 2, considerando como dados iniciais $\sigma_s=0.0001$ e procuramos o melhor r que nos dê o menor erro percentual.

As Figs. 5.34 e 5.35 apresentam os EP e os tempos computacionais, respectivamente, *versus* r , para a estimativa de σ_s , nos três casos testes dadas na tabela 5.6.

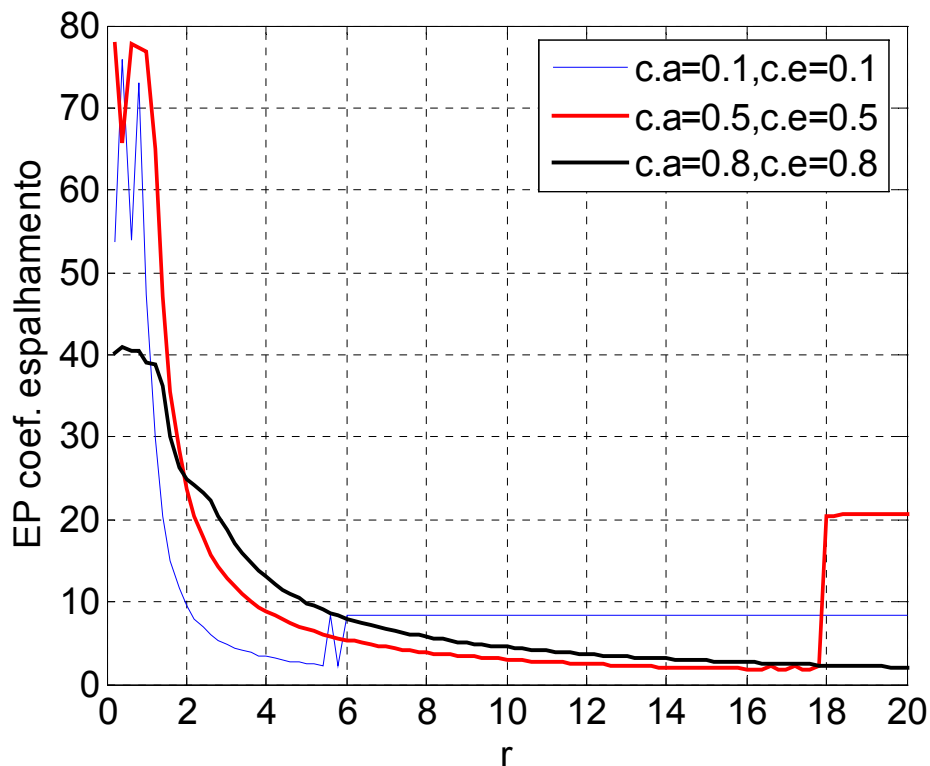


Fig. 5.34 Erro percentual (Ep) na estimativa de σ_s *versus* r . Geometria 2

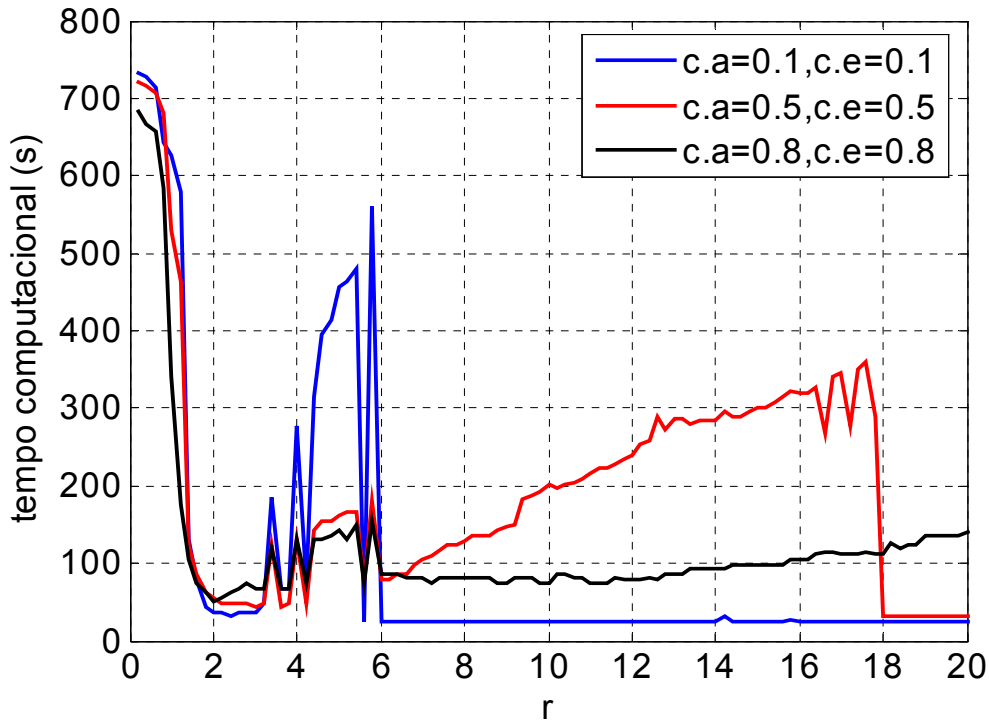


Fig. 5.35 Tempo computacional na estimativa de σ_s versus r . Caso teste 2. Meio Homogêneo. Geometria 2

Da Fig. 5.34, observa-se nos três casos testes que com $r=15$, o erro percentual na estimativa de σ_a é menor que 10. Note que para o primeiro e o terceiro caso teste, $r=15$, não é o melhor valor que dá o menor erro percentual.

Os valores de σ_s obtido com $r=15$ são guardados no arquivo r2.tmp. Um estudo mais detalhado foi feito depois de terminar estes três casos testes.

Logo, com a geometria 1 se vai retornar a estimar σ_a considerando agora como dados iniciais de σ_{a0} e σ_{s0} os valores dos arquivos de r1.tmp e r2.tmp respectivamente.

As Figs. 5.36, 5.37 e 5.38 apresentam os novos σ_a estimados para diferentes valores de r . Note que os erros percentuais na estimativa de σ_a são menores que 1.5 e não variam muito com diferentes r , enquanto σ_s não varia muito em relação aos dados iniciais, dados por r2.tmp.

Pode-se considerar $r=9$ como o valor ótimo, que apresenta menor erro percentual na estimativa de σ_a nos três casos testes. Os dados estimados são guardados no arquivo r3.tmp.

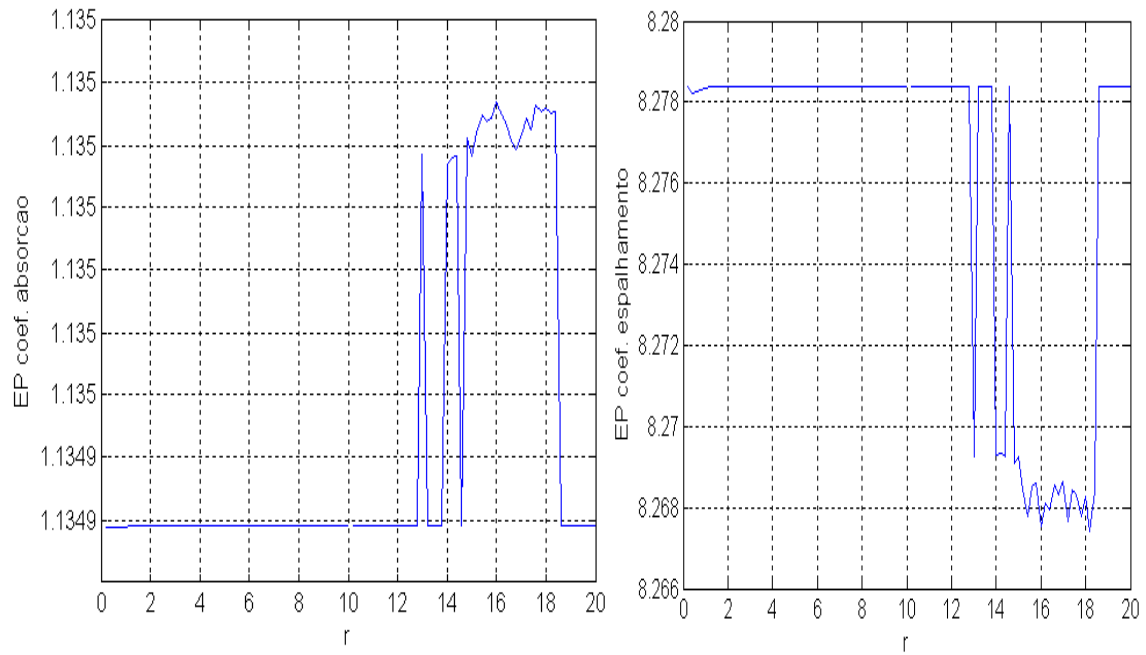


Fig. 5.36. Ep na estimativa de σ_a e σ_s versus r Caso teste 1. Meio Homogêneo $\sigma_a=0.1$ e $\sigma_s=0.1$. Geometria 1

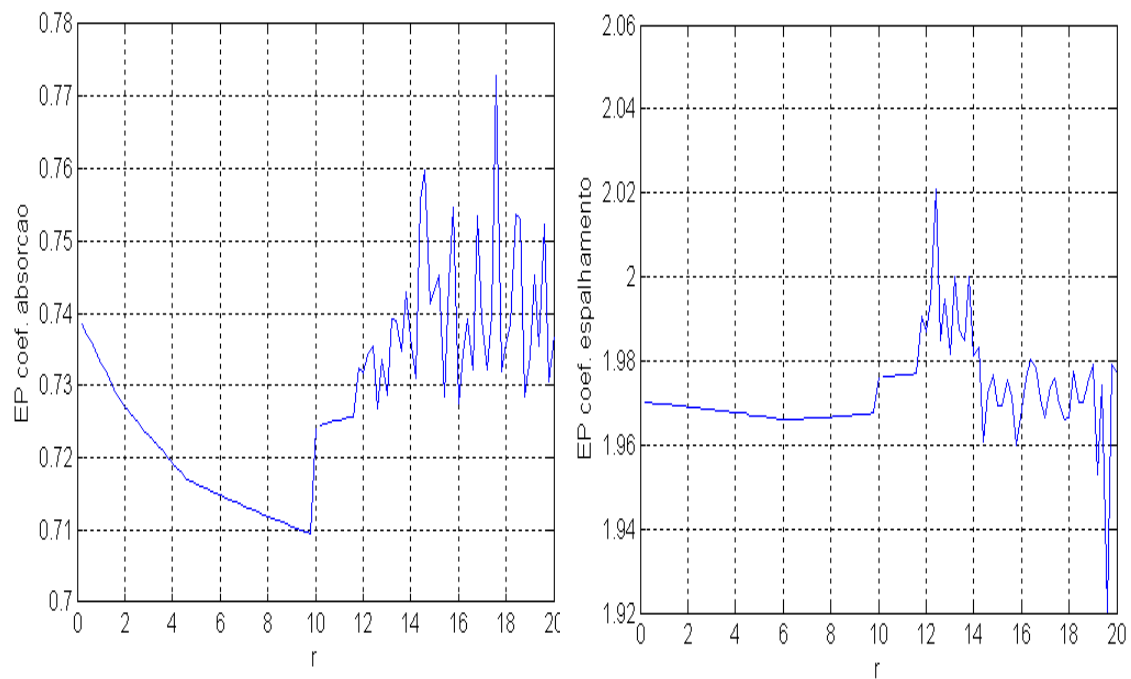


Fig. 5.37. Ep na estimativa de σ_a e σ_s versus r Caso teste 2. Meio Homogêneo $\sigma_a=0.5$ e $\sigma_s=0.5$. Geometria 1

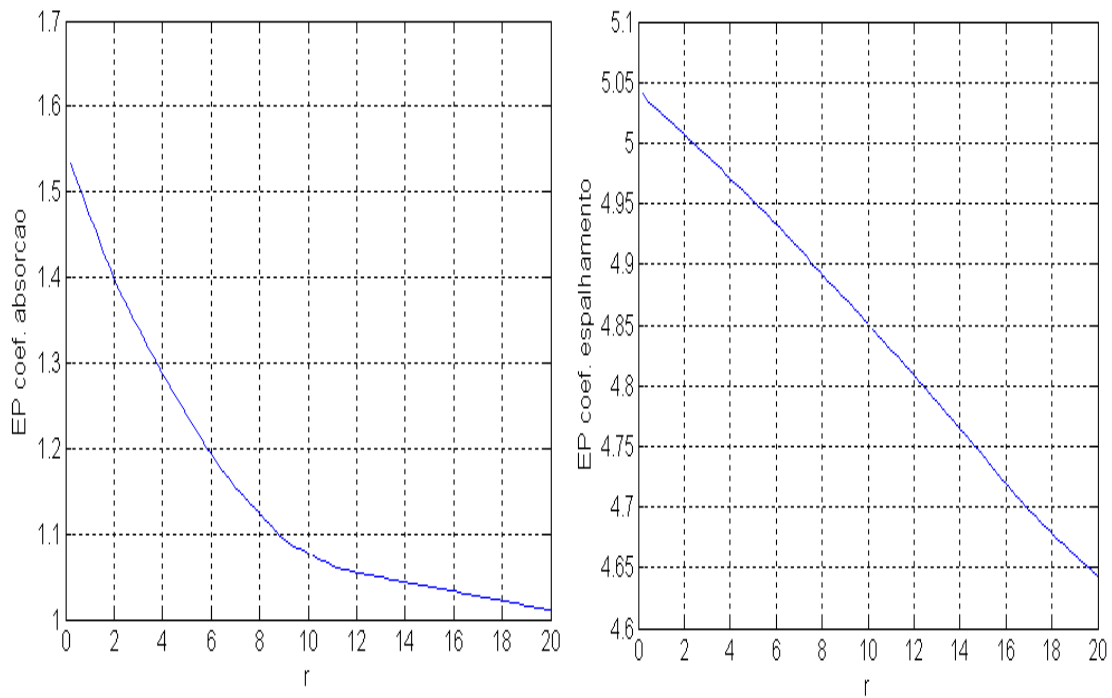


Fig. 5.38. Ep na estimativa de σ_a e σ_s versus r Caso teste 3. Meio Homogêneo $\sigma_a=0.8$ e $\sigma_s=0.8$. Geometria 1.

Tratando de melhorar as estimativas do coeficiente de espalhamento, fez-se um estudo sobre a variação dos r ótimos segundo o valor de σ_s a ser estimado para os pontos $[0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]$. Os r ótimos destes pontos são apresentados por asteriscos na Fig. 5.39. Nesta Fig., a curva é obtida fazendo um ajuste sobre os pontos com seus respectivos r ótimos.

Algoritmo proposto para a estimativa de σ_a e σ_s

- 1.- Considerando: o problema direto da geometria 1, 200 incógnitas, $\sigma_{ao} = \sigma_{so} = 0.001$ Minimizar $D_{1r}(\sigma, \sigma_0)$ sujeito $F(\sigma)$ com $r=2.0$.
- 2.- Considerando: o problema direto da geometria 2, 100 incógnitas, σ_a conhecido do passo 1. Minimizar $D_{1r}(\sigma, \sigma_0)$ sujeito $F(\sigma)$ com $r=15.0$.
- 3.- Considerando: o problema direto da geometria 1, 200 incógnitas, σ_{ao} e σ_{so} conhecido do passo 1 e 2. Minimizar $D_{1r}(\sigma, \sigma_0)$ sujeito $F(\sigma)$ com $r=9.0$.

4.- Considerando: o problema direto da geometria 2, 100 incógnitas, $\sigma_{so} = 0.001$, σ_a conhecido do passo 3, Minimizar $D_{1,r}(\sigma, \sigma_0)$ sujeito $F(\sigma)$ com um r calculado da curva apresentada 5.51, a partir de um $\bar{\sigma}_s$ médio obtido dos σ_s estimados no passo 2.

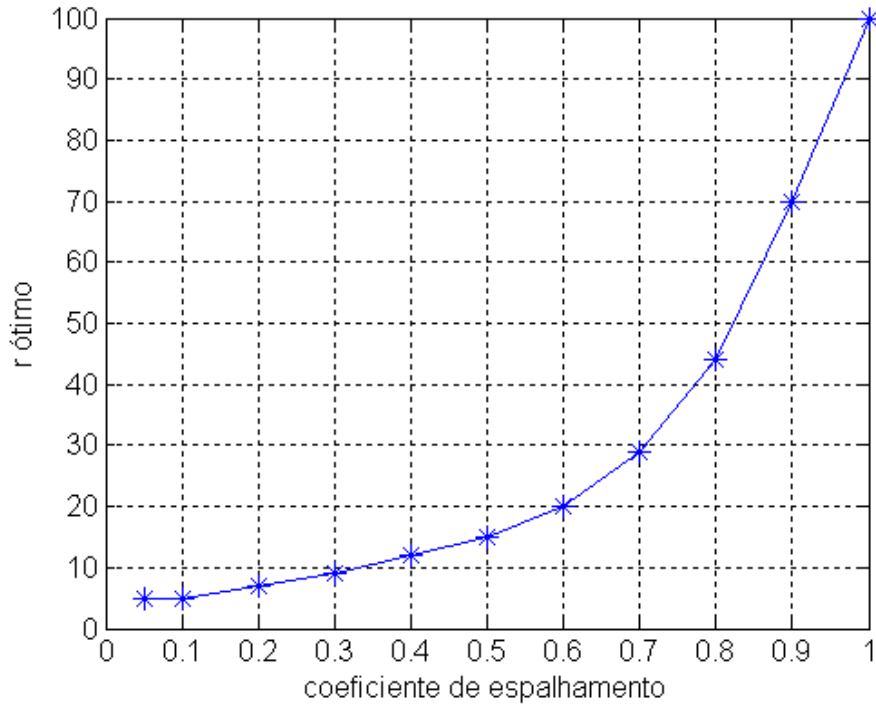


Fig. 5.39. Os r ótimos para a estimativa do coeficiente de espalhamento.

Exemplo 1.- Meio Homogêneo com $\sigma_a = 0.7$ e $\sigma_s = 0.7$, meio particionado em 10x10

A Fig. 5.40 apresenta os σ_a e σ_s estimados e a Tabela 5.7 apresenta os erros percentuais obtidos em cada passo. O tempo computacional gasto foi de 15 minutos.

Tabela 5.7. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 1

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	3.7747827e+000	
EP Passo 2		3.0775669e+000
EP Passo 3	1.1877521e+000	
EP Passo 4		1.3936111e+000

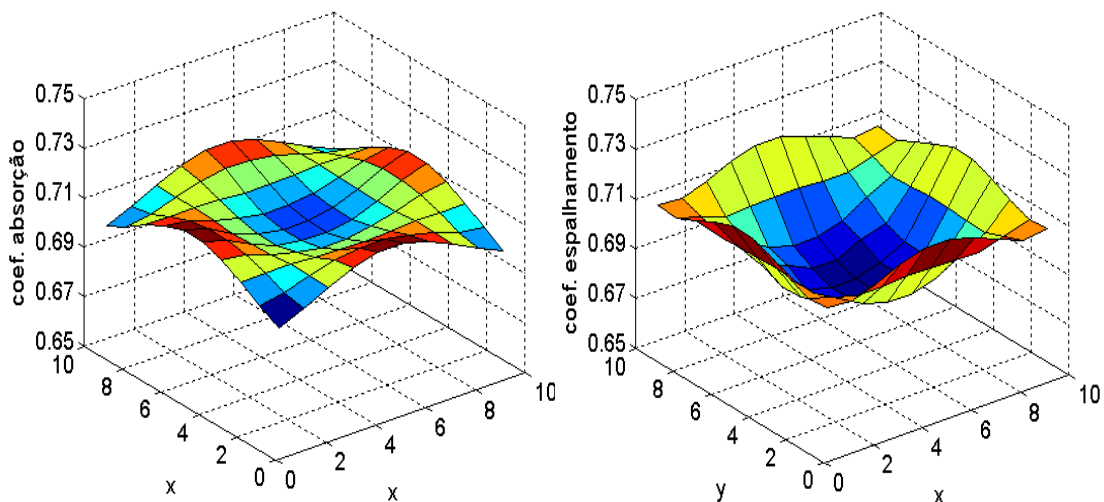


Fig. 5.40. Estimativa de $\sigma_a = 0.7$ e $\sigma_s = 0.7$. Exemplo 1

Exemplo 2.- Meio Homogêneo com $\sigma_a = 0.35$ e $\sigma_s = 0.35$, meio particionado em 10x10. A Fig. 5.41 apresenta os σ_a e σ_s estimados e a tabela 5.8 apresenta os erros percentuais obtidos em cada passo. O tempo computacional gasto foi de 5 minutos.

Tabela 5.8. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 2.

	σ_a	σ_s
Valores iniciais	$\sigma_{a0} = 0.0001$	$\sigma_{s0} = 0.0001$
EP Passo 1	1.21600951124759	
EP Passo 2		15.71650461017580
EP Passo 3	0.69722759669040	
EP Passo 4		2.40007109728560

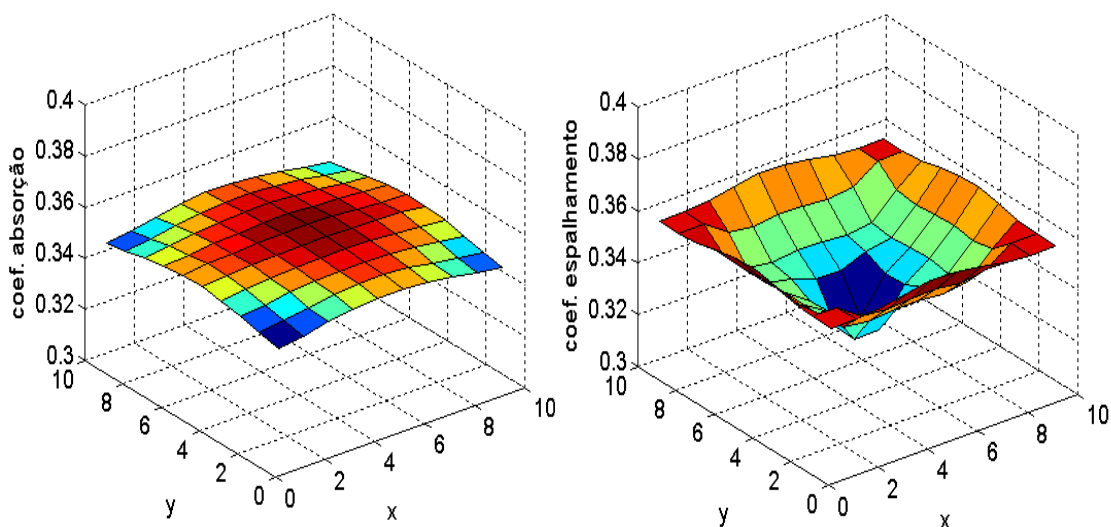


Fig. 5.41.- Estimativa de $\sigma_a = 0.35$ e $\sigma_s = 0.35$. Exemplo 2

Exemplo 3.- Meio Homogêneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, meio particionado em 8x8

A Fig. 5.54 apresenta os σ_a e σ_s estimados e a tabela 5.9 apresenta os erros percentuais obtidos em cada passo. O tempo computacional gasto foi de 15 minutos.

Tabela 5.9. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 3

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	1.51816177694081	
EP Passo 2		2.15862702772479
EP Passo 3	0.70400606135955	
EP Passo 4		1.90465418530397

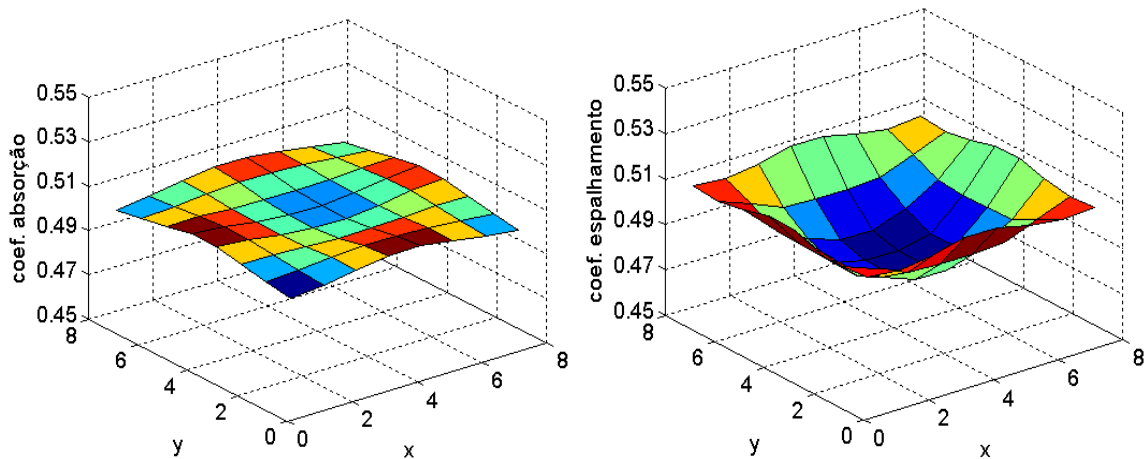


Fig. 5.42.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 3

Exemplo 4.- Meio Homogêneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, meio particionado 12*12

A Fig. 5.43 apresenta os σ_a e σ_s estimados e a tabela 5.10 apresenta os erros percentuais obtidos em cada passo. O tempo computacional gasto foi de 30 minutos.

Tabela 5.10. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 4

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	1.53528975939561	
EP Passo 2		2.19699902341753
EP Passo 3	0.70943877516242	
EP Passo 4		1.94802120751292

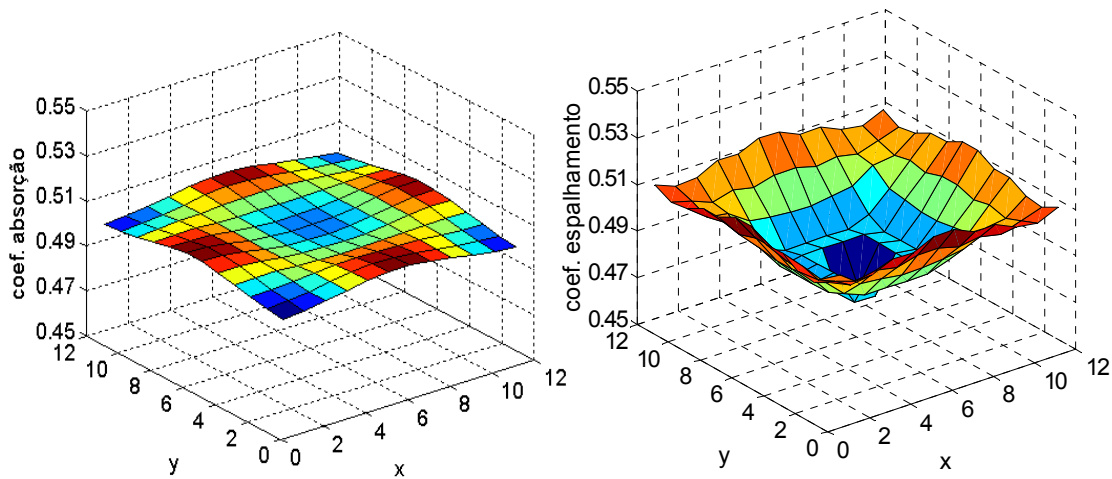


Fig. 5.43.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 4

Exemplo 5.- Meio Homogêneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, Meio particionado 16*16

A Fig. 5.44 apresenta os σ_a e σ_s estimados e a tabela 5.11 apresentam os erros percentuais obtidos em cada passo. O tempo computacional gasto foi de 90 minutos.

Tabela 5.11. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 5

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	1.54183428649248	
EP Passo 2		2.13883082697250
EP Passo 3	0.70604473085297	
EP Passo 4		1.99580669168826

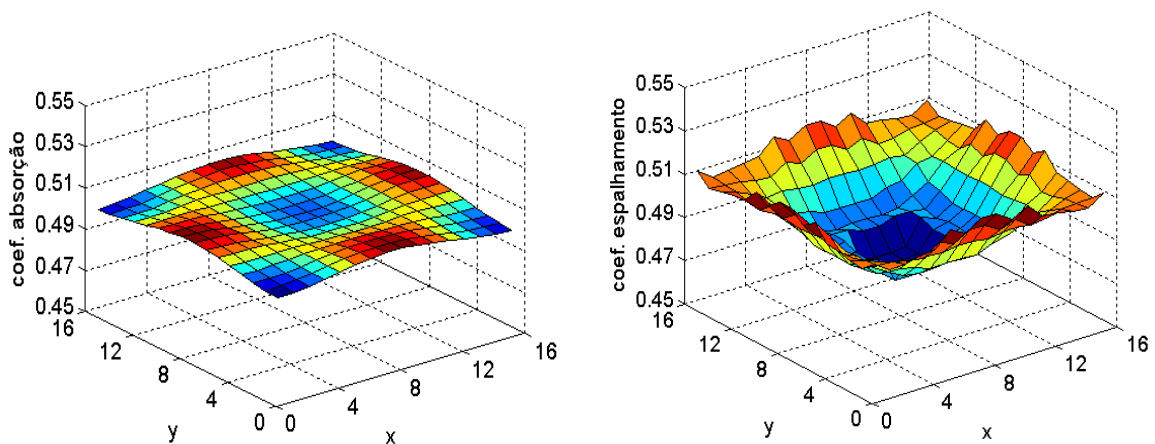


Fig. 5.44.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 5

Exemplo 6.- Meio Homogêneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, Meio particionado 20*20

A Fig. 5.45 apresenta os σ_a e σ_s estimados e a tabela 5.12 apresenta os erros percentuais obtidos em cada passo. O tempo computacional foi de 126 minutos

Tabela 5.12. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 6

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	1.5444735e+000	
EP Passo 2		2.1913965e+000
EP Passo 3	0.91607841	
EP Passo 4		2.0015647895

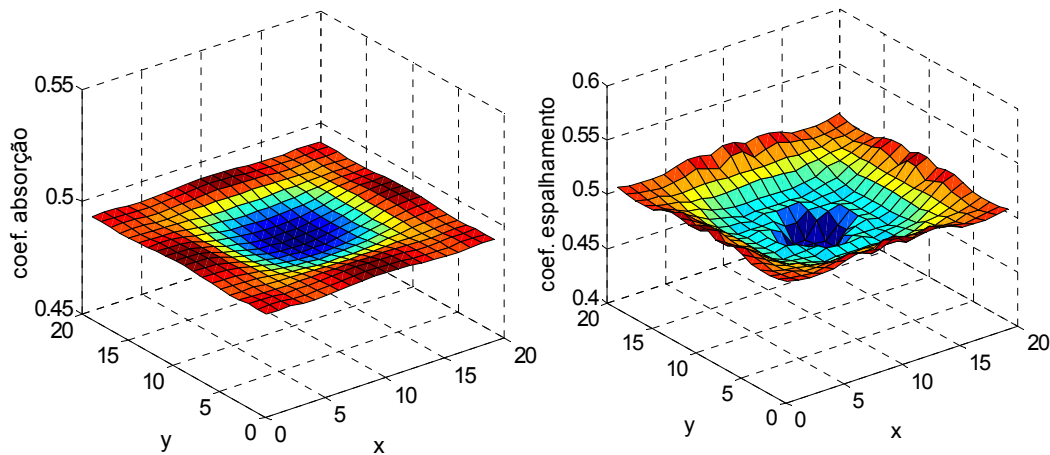


Fig. 5.45.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 6

Exemplo 7.- Meio Homogêneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, com ruído randômico at 0.5%. Meio particionado 10*10.

A Fig. 5.46 apresenta os σ_a e σ_s estimados e a tabela 5.13 apresenta os erros percentuais obtidos em cada passo. O tempo computacional foi de 47 minutos.

Tabela 5.13. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 7

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	2.77627432327319	
EP Passo 2		3.05813627286066
EP Passo 3	2.61689197194661	
EP Passo 4		2.97019499506470

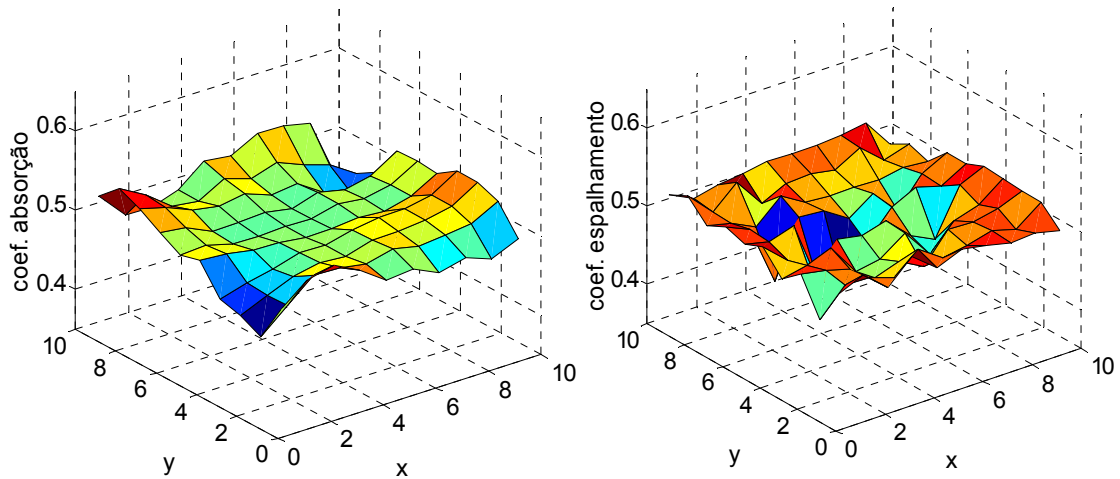


Fig. 5.46.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 7

Exemplo 8.- Meio Homôgeneo com $\sigma_a = 0.5$ e $\sigma_s = 0.5$, com ruído randômico até 1%. Meio particionado 10*10.

A Fig. 5.47 apresenta os σ_a e σ_s estimados e a tabela 5.14 apresenta os erros percentuais obtidos em cada passo. O tempo computacional foi de 71 minutos.

Tabela 5.14. EP obtidos para a estimativa de σ_a e σ_s . Exemplo 8

	σ_a	σ_s
Valores iniciais	$\sigma_{ao} = 0.0001$	$\sigma_{os} = 0.0001$
EP Passo 1	4.571575631605	
EP Passo 2		6.5528196216
EP Passo 3	5.0945423742	
EP Passo 4		7.169384186

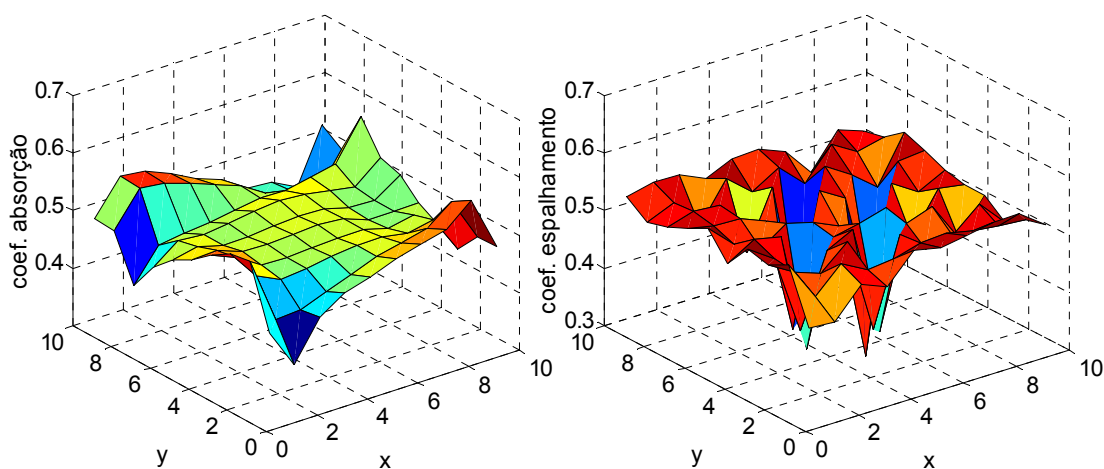


Fig. 5.47.- Estimativa de $\sigma_a = 0.5$ e $\sigma_s = 0.5$. Exemplo 7

5.2.3.2 Meios heterogêneos

Para meios heterogêneos vão-se empregar, os dois primeiros passos dados no algoritmo para meio homogêneos

Passo 1.- Considerando: o problema direto da geometria 1, 200 incógnitas, $\sigma_{ao} = \sigma_{so} = 0.0001$ Minimizar $D_{1r}(\sigma, \sigma_0)$ sujeito à $F(\sigma)$ com $r = 2.0$.

Passo 2.- Considerando: o problema direto da geometria 2, 100 incógnitas, σ_a conhecido do passo 1. Minimizar $D_{1r}(\sigma, \sigma_0)$ sujeito $F(\sigma)$ com $r=15.0$.

Exemplo 1.- Considere um meio formado por 2 materiais segundo a Fig. 5.48, com diferentes σ_a e igual σ_s . A tabela 5.15 apresenta as propriedades do meio e os erros percentuais obtido em cada passo. Na Fig. 5.49 apresentam-se, os valores estimados de σ_a e σ_s depois do passo 2. O tempo computacional foi de 15 minutos.

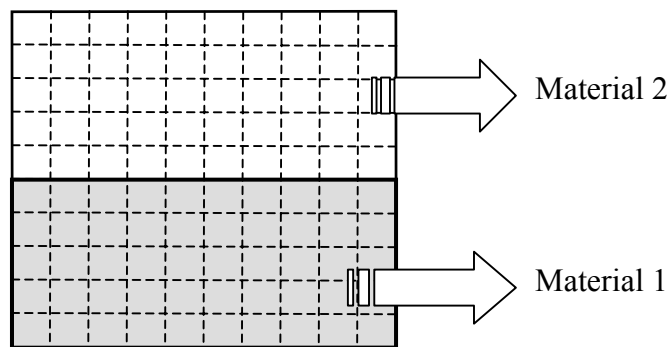


Fig. 5.48.- Localização dos materiais

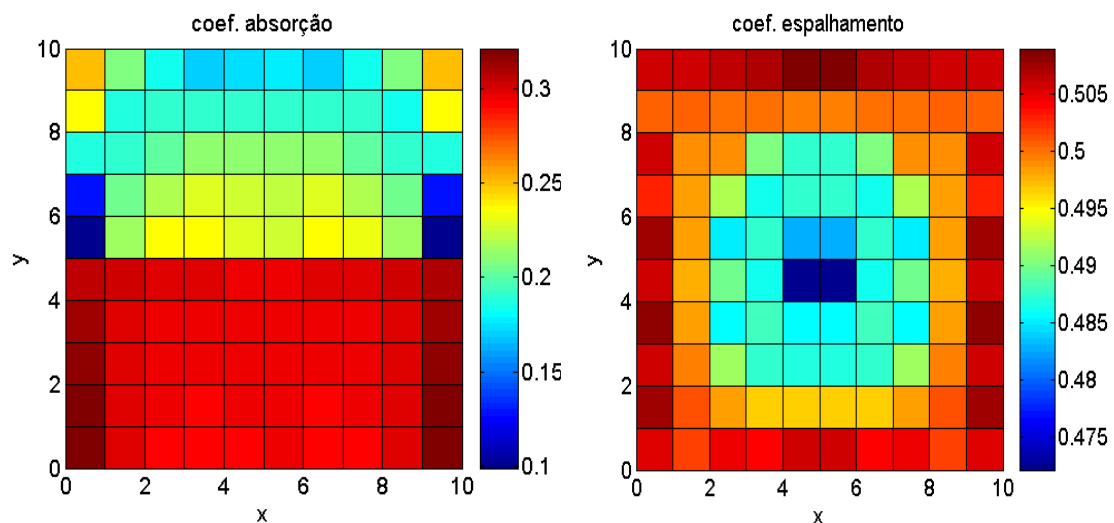


Fig. 5.49.- Estimativa do coeficiente de σ_a e σ_s . Exemplo1

Tabela 5.15. Propriedades dos materiais e os EP de σ_a e σ_s . Exemplo 1

	σ_a	σ_s
Material 1	0.3	0.5
Material 2	0.2	0.5
Valores iniciais	$\sigma_{ao}=0.0001$	$\sigma_{os}=0.0001$
EP Passo1	1.0910281	
EP Passo 2		1.4023716

Exemplo 2.- Considere um meio formado por 2 materiais segundo a Fig. 5.48, com igual σ_a e diferentes σ_s . A tabela 5.16 apresenta as propriedades do meio e os erros percentuais obtidos em cada passo. Na Fig. 5.50 apresentam-se os valores estimados de σ_a e σ_s depois do passo 2. O tempo computacional foi de 14 minutos.

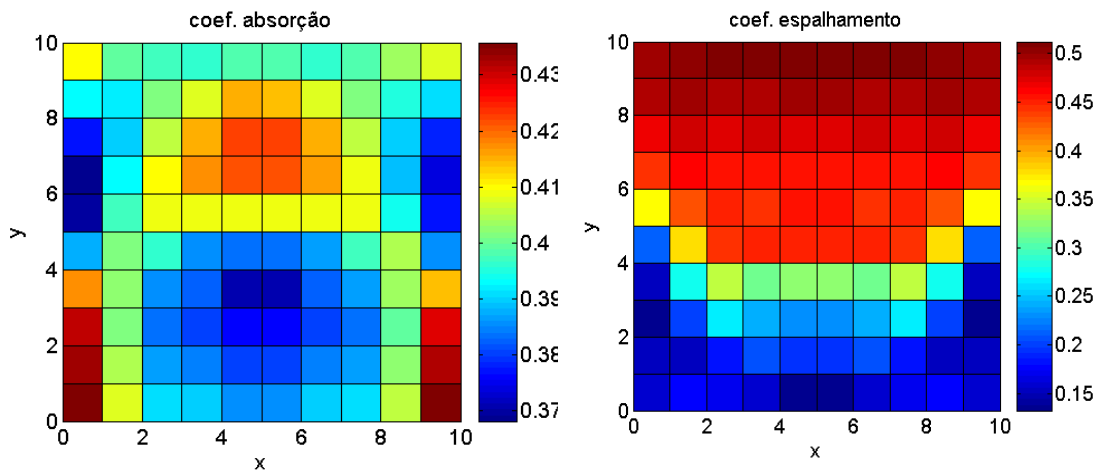


Fig. 5.50.- Estimativa do coeficiente de σ_a e σ_s . Exemplo 2

Tabela 5.16. Propriedades dos materiais e os EP de σ_a e σ_s . Exemplo 2

	σ_a	σ_s
Material 1	0.4	0.2
Material 2	0.4	0.5
Valores iniciais	$\sigma_{ao}=0.0001$	$\sigma_{os}=0.0001$
EP Passo1	3.625605078	
EP Passo 2		22.8492401

Exemplo 3.- Considere um meio formado por 2 materiais segundo a Fig. 5.51, com diferentes σ_a e σ_s . A tabela 5.17 apresenta as propriedades do meio e os erros

percentuais obtido em cada passo. Na Fig. 5.60 apresentam-se os valores estimados de σ_a e σ_s depois do passo 2. Tempo computacional foi de 15 minutos.

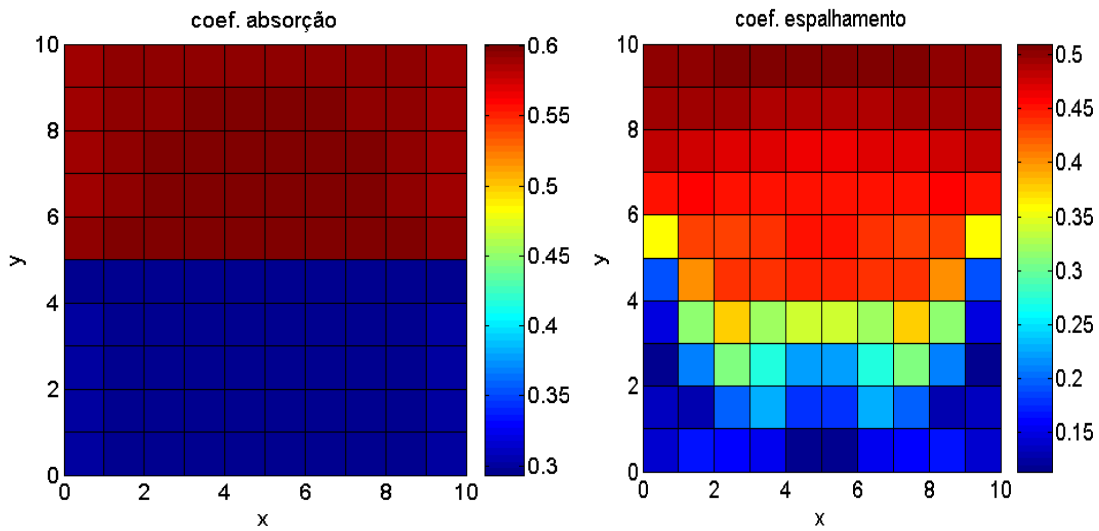


Fig. 5.51.- Estimativa do coeficiente de σ_a e σ_s . Exemplo 2

Tabela 5.17. Propriedades dos materiais e os EP de σ_a e σ_s . Exemplo 2

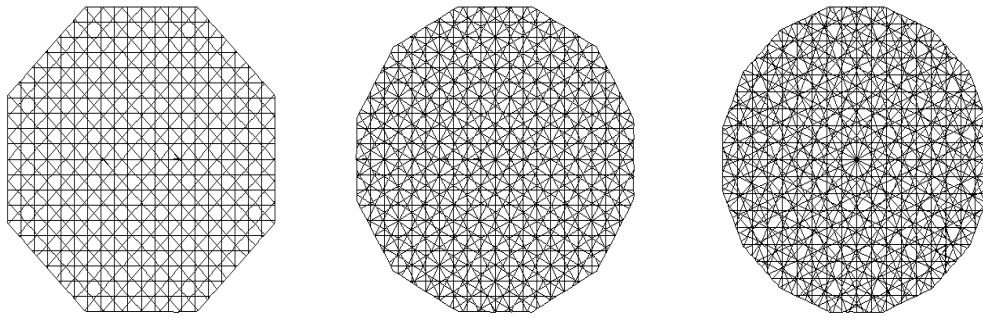
	σ_a	σ_s
Material 1	0.3	0.2
Material 2	0.6	0.5
Valores iniciais	$\sigma_{ao}=0.0001$	$\sigma_{os}=0.0001$
EP Passo1	1.13355649581256	
EP Passo 2		26.31271990884450

Destes 3 exemplos, observa-se que, a estimativa de σ_s em meios heterogêneos foi possível nos casos em que os σ_a eram diferentes entre os materiais que formam o meio.

5.3.- Reconstrução de imagem para tomografia de feixes paralelos

5.3.1 Reconstrução de imagens com o algoritmo *ART* empregando o funcional $\eta_{s,r}$

Na Fig. 5.52 apresentam-se as partições do meio que foram empregadas nos casos testes. A tabela 5.19 apresenta o número de elementos de área gerados em cada partição e o numero de detectores a ser empregados na reconstrução.



(a) 4 vistas e 20 faixas.

(b) 6 vistas e 20 faixas.

(c) 8 vistas e 18 faixas.

Partição 1

Partição 2

Partição 3

Fig. 5.52 – Construção da partição de domínio empregada nos casos testes.

Tabela 5.19. Partições do meio

	Número de vistas	Número de faixas	Número de elementos de área	Número de detectores
Partição 1	4	20	1552	80
Partição 2	6	20	2940	120
Partição 3	8	18	4832	144

5.3.1.1 Exemplo 1. Meio homogêneo

Para este exemplo considera-se um meio homogêneo com $\sigma_a = 0.056$. As figuras. 5.53 e 5.54 apresentam os erros percentuais e os tempos computacionais, respectivamente, na estimativa de σ_a com diferentes r . Em cada figura, cada curva apresenta a partição empregada. Curvas similares às mostradas na Fig. 5.53 foram obtidas para valores iniciais, σ_{a0} com valores 0.00001, 0.0001, 0.001, 0.01, 0.1 e 1.0.

Das Figs. 5.53 e 5.54, observa-se uma boa estimativa independente do valor inicial da partição empregada. A melhor estimativa é obtida para a partição 1 e 3, com $r=2$. O tempo computacional gasto foi menor há 30 segundos.

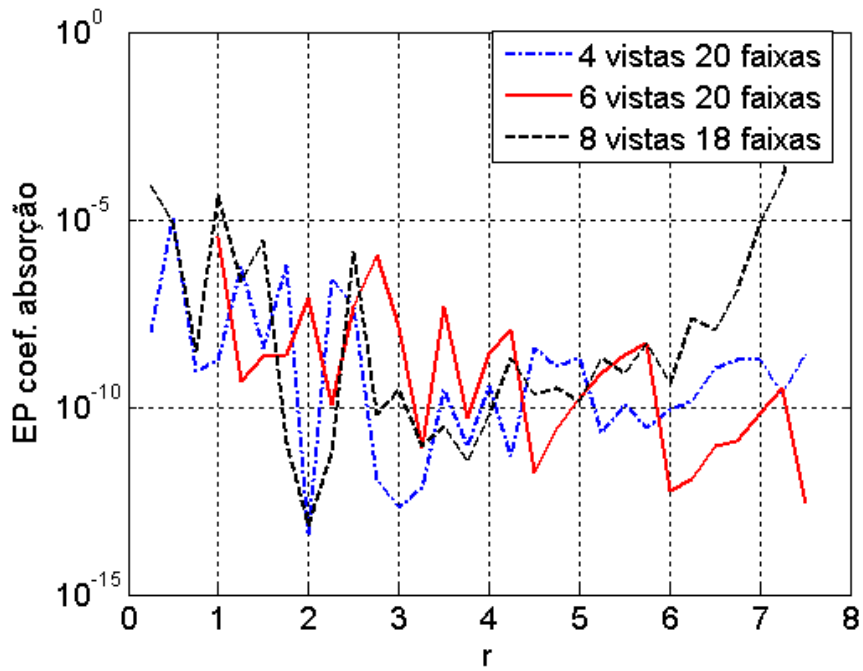


Fig. 5.53.- Erro percentual *versus* r . Exemplo 1

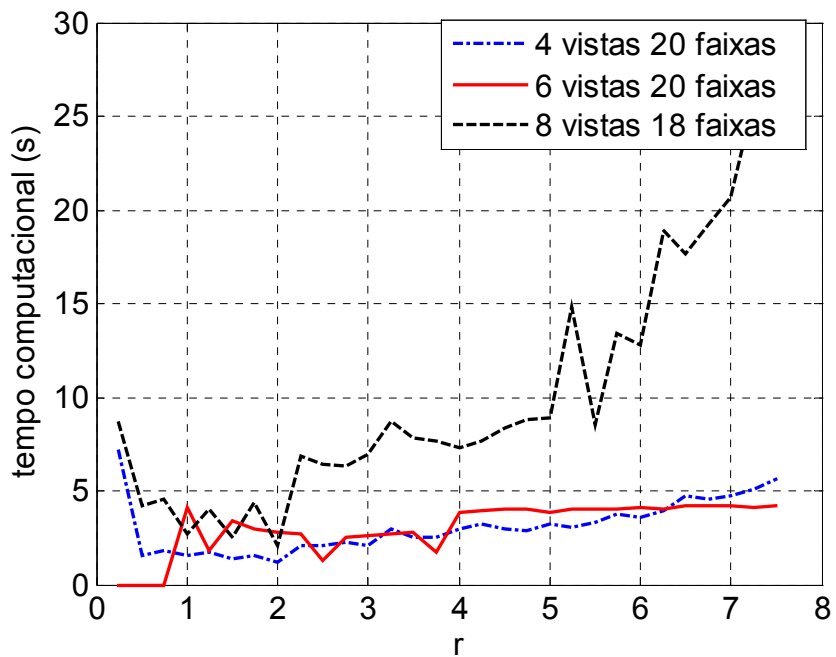


Fig. 5.54.- Tempo computacional (s) *versus* r. Exemplo 1

A Figs. 5.55 apresentam imagens reconstruídas com $r=2$ com um meio particionado com 6 vistas 20 faixas, $\sigma_{ao} = 0.001$. A Fig.5.55 (a) apresenta a imagem original e a Fig. 5.55 (b) a imagem reconstruída, a imagem 5.55 (c) a diferença entre o σ_a original e o calculado.

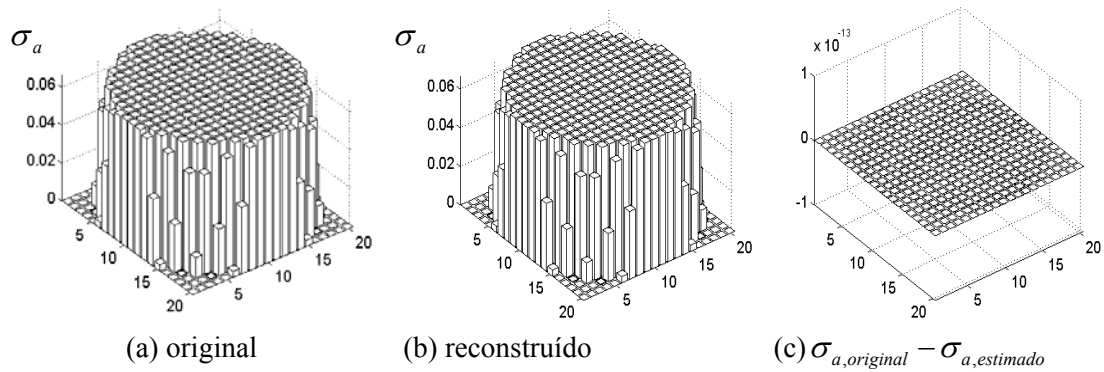


Fig. 5.55.- Reconstrução de um meio homogêneo. Exemplo 1

5.3.1.2 Exemplo 2. Meio heterogêneo com uma deformação

Neste item, para os casos testes aqui apresentados as propriedades do meio são $\sigma_a=0.0028$ e na deformação $\sigma_a=0.035$. A imagem a ser reconstruída é similar à figura 5.59a. Procurando encontrar o melhor valor inicial, as Figs. 5.56, 5.57 e 5.58 apresentam os erros percentuais obtidos com diferentes r e para diferentes valores iniciais. Na Fig. 5.56 emprega-se a partição 1, na Fig. 5.57 a partição 2 e na Fig. 5.58 a partição 3, todas elas descritas na tabela 5.19.

Das Figs. 5.56, 5.57 e 5.58 observa-se que o menor erro percentual é obtido com a partição 3, de 8 vistas e 18 faixas. Os melhores valores iniciais estão na faixa de $[0.001 - 0.1]$ quando $r=0.25$, para a partição 1 e $r=0.1$ para a partição 2 e 3.

A Fig. 5.59 apresenta as imagens reconstruídas com $\sigma_{ao} = 0.001$, $r=0.1$ em um meio particionado de 8 vistas e 18 faixas. A Fig. 5.59(a) apresenta a imagem original, a Fig. 5.59 (b) a reconstruída, e a Fig. 5.59 (c) a diferença entre o σ_a original e o calculado.

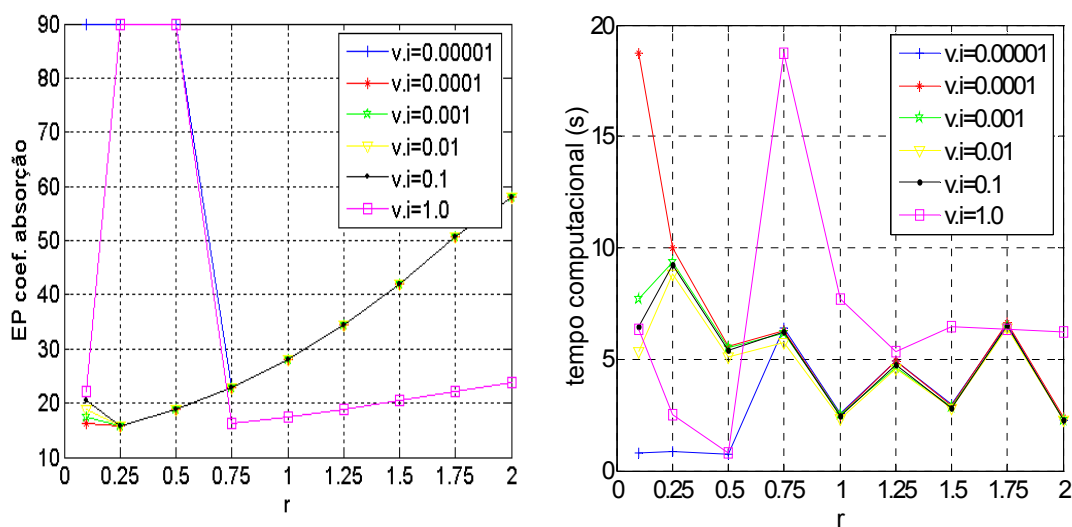


Fig. 5.56.- Erro percentual versus r. Partição 1: 4 vistas 20 faixas. Exemplo 2

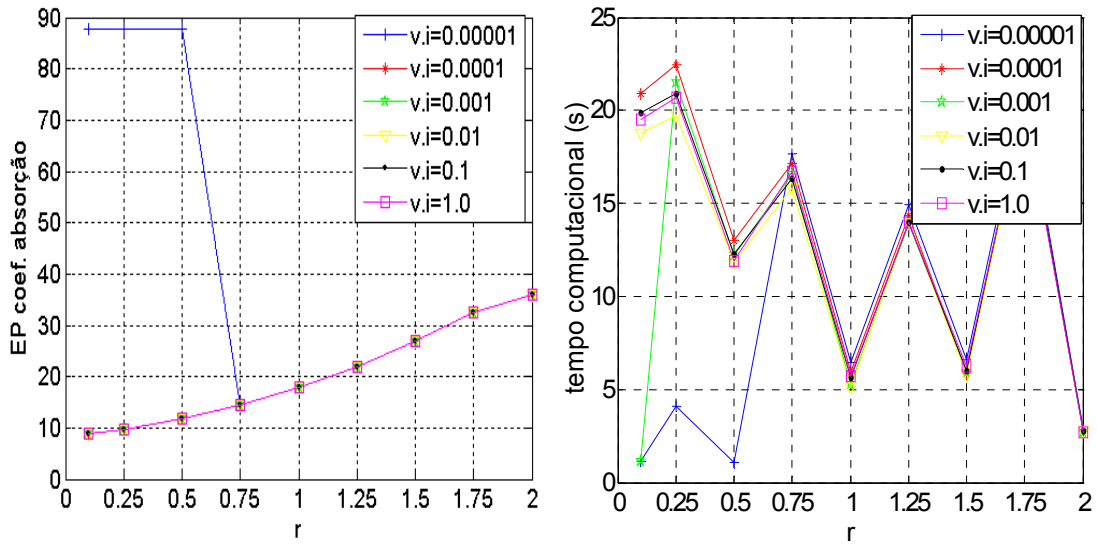


Fig. 5.57.- Erro percentual *versus* r. Partição 2: 6 vistas 20 faixas. Exemplo 2

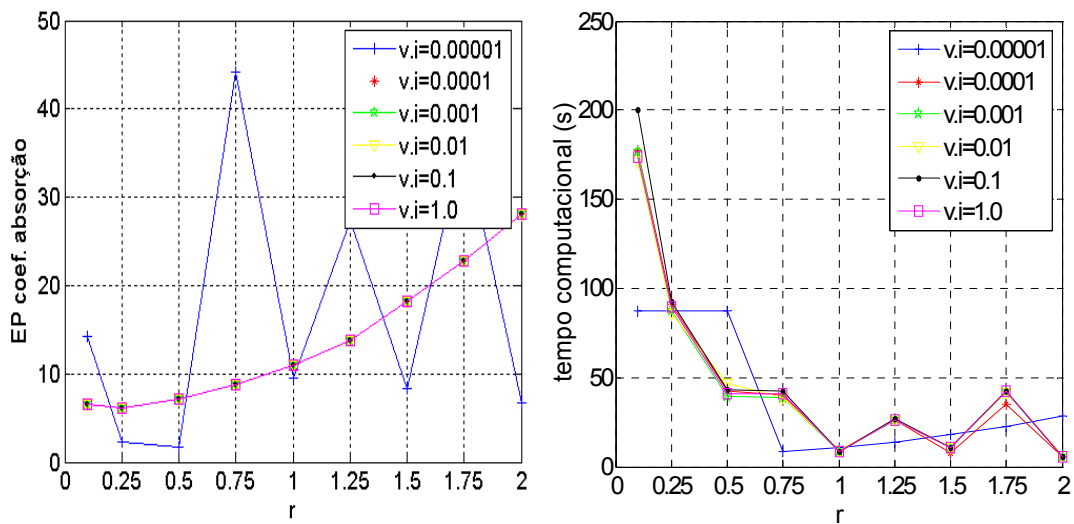


Fig. 5.58.- Erro percentual *versus* r. Partição 3: 8 vistas 18 faixas. Exemplo 2

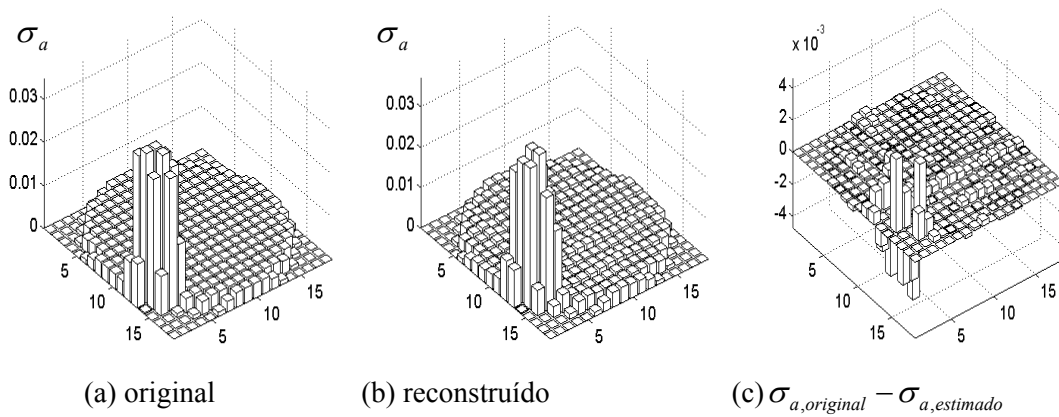


Fig. 5.59.- Erro percentual *versus* r. Partição 3: 8 vistas 18 faixas. Exemplo 2

5.3.1.3.- Exemplo 3.- Meio heterogêneo com duas deformações.

Para todos os casos testes, neste item, as propriedades do meio são $\sigma_a=0.0028$ e na deformação $\sigma_a=0.035$. A imagem a ser reconstruída é similar à figura apresentada no final deste item. Procurando o melhor valor inicial, as Figs. 5.60, 5.61 e 5.62 apresentam os erros percentuais obtidos com diferentes r e diferentes valores iniciais (σ_{ao}). Na Fig. 5.60 emprega-se a partição 1, na Fig. 5.61 a partição 2 e na Fig. 5.62 a partição 3, descritas na tabela 5.19.

Como no exemplo 1, observa-se que, o menor erro percentual é obtido com a partição 3, de 8 vistas e 18 faixas, os melhores σ_{ao} estão na faixa de $[0.001 - 0.1]$ quando $r=0.25$ para a partição 1, e $r=0.1$ para a partição 2 e 3.

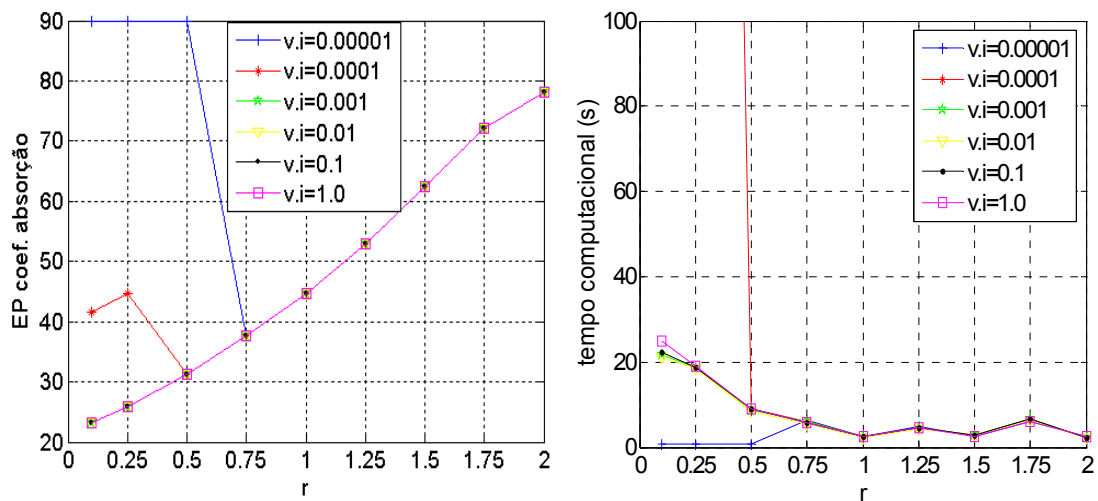


Fig. 5.60.- Erro percentual *versus* r. Partição 1: 4 vistas 20 faixas. Exemplo 3

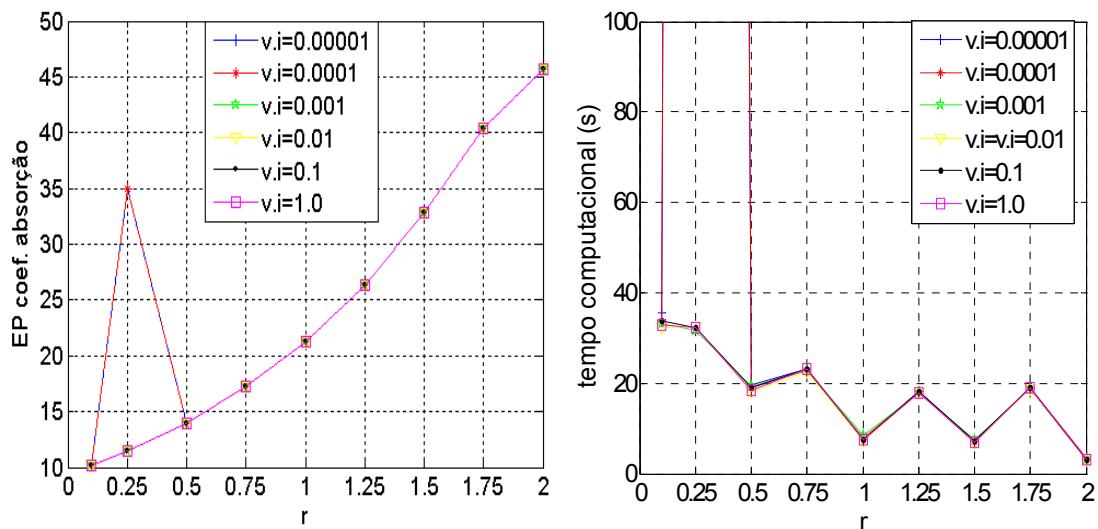


Fig. 5.61.- Erro percentual *versus* r. Partição 2: 6 vistas 20 faixas. Exemplo 3

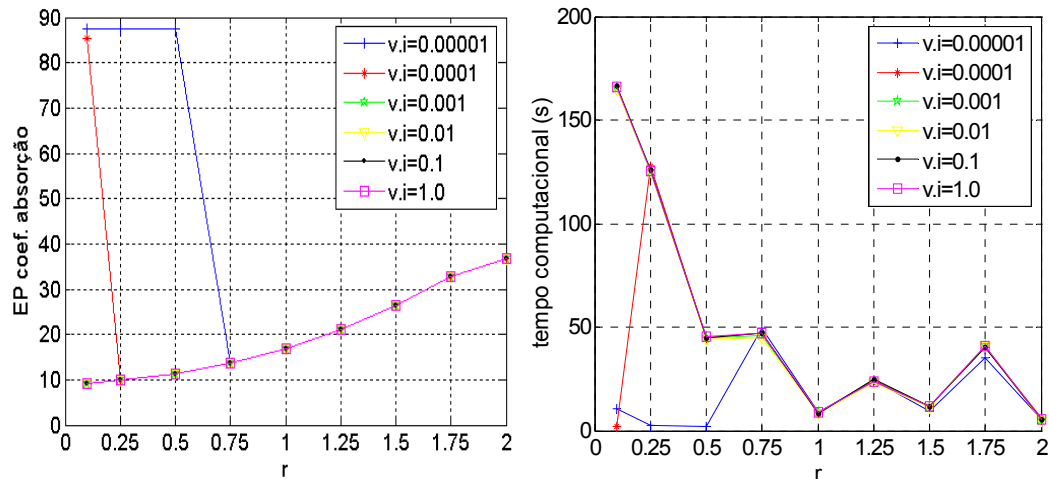


Fig. 5.62.- Erro percentual *versus* r. Partição 3: 8 vistas 18 faixas. Exemplo 3

A Fig. 5.63 apresenta a imagem reconstruída com $\sigma_{ao} = 0.001$, $r=0.1$, empregando a partição 3. A Fig. 5.63 (a) apresenta a imagem original, a Fig. 5.63 (b) a imagem reconstruída, a Fig. 5.63. (c) a diferença entre σ_a original e o calculado

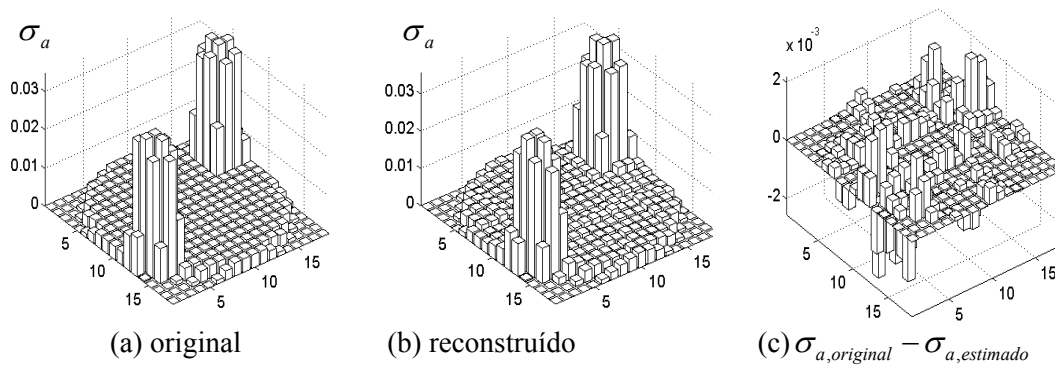


Fig. 5.63. - Erro percentual *versus* r. Partição 3: 8 vistas 18 faixas. Exemplo 3

5.3.1.4.- Exemplo 4.- Dados com ruído

A Fig. 5.64 apresenta o caso teste de uma imagem reconstruída de um meio homogêneo, a Fig. 5.65 uma imagem reconstruída de um meio com uma deformação e a Fig. 5.66 um meio com duas deformações. Nos três casos - testes acima mencionados a imagem (a) apresenta a imagem original, nas outras imagens foram empregados dados com ruído randômico para a imagem (b) até 1% e a imagem (c) até 10%.

Na tabela 5.20 apresentam-se as propriedades do meio, o erro percentual e o tempo computacional gasto, das três imagens reconstruídas do exemplo 4.

Nos três casos foi usada a partição 3, formada de 8 vista e 18 faixas, o valor de $r=0.2$ e os valores iniciais de $\sigma_{ao} = 0.001$.

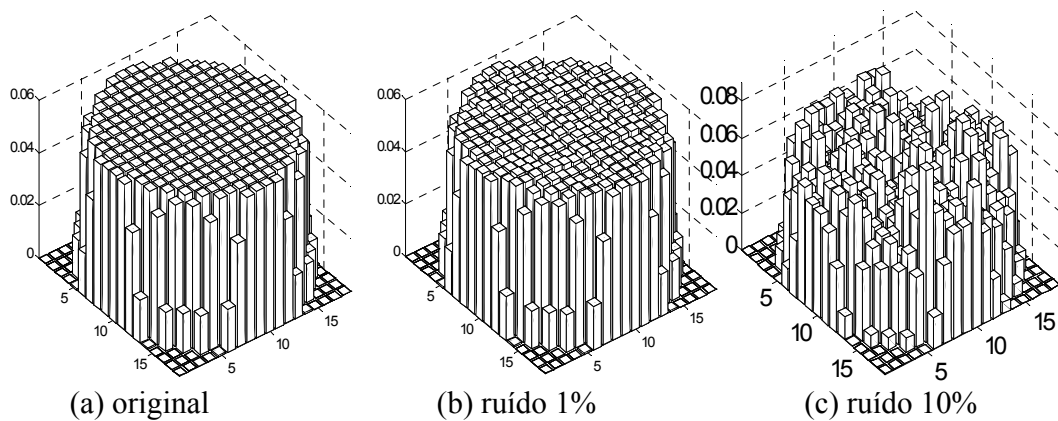


Figura 5.64.- Imagem reconstruída meio homogêneo, dados com ruído

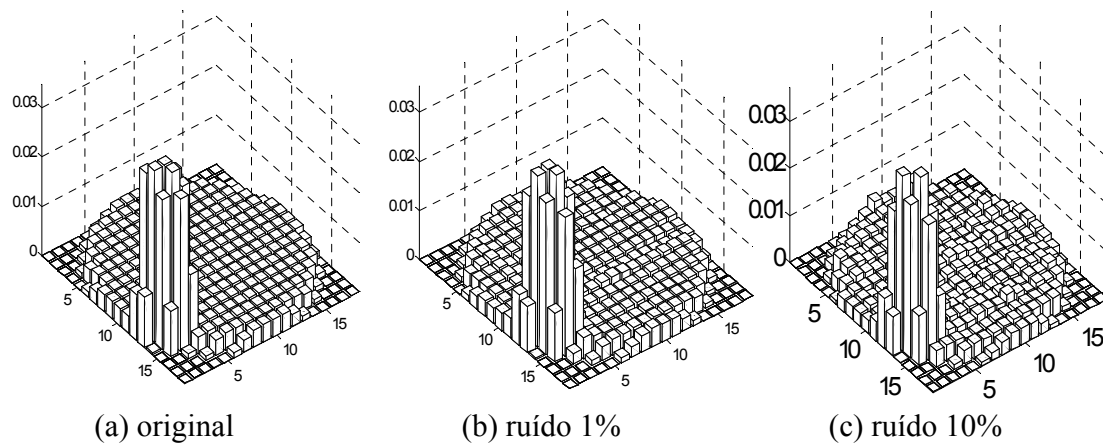


Figura 5.65.- Imagem reconstruída com uma deformação, dados com ruído

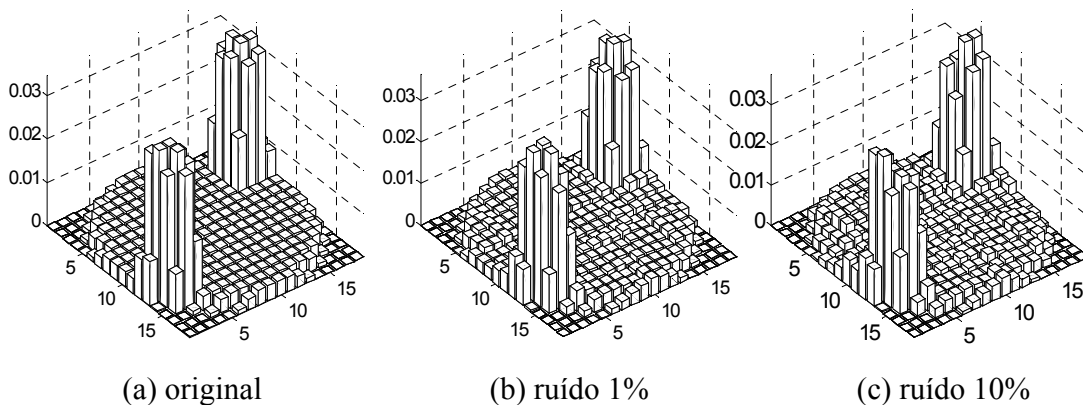


Figura 5.66.- Imagem reconstruída com duas deformações, dado com ruído

Tabela 5.20. Propriedades do meio. Exemplo 4

	Meio homogêneo	Meio Heterogêneo com uma deformação	Meio heterogêneo com duas deformações
Figuras	5.64	5.65	5.66
σ_a no meio	0.0056	0.0028	0.0028
σ_a nas deformações	0.0056	0.035	0.035
Valores iniciais	$\sigma_{ao} = 0.001$	$\sigma_{ao} = 0.001$	$\sigma_{ao} = 0.001$
E.P. sem ruído	0.000	5.8632	9.897
E.P. com ruído 1%	0.9427	6.1865	9.9630
E.P. com ruído 2%	1.77	6.27	10.26
E.P. com ruído 3%	2.98	6.83	10.39
E.P. com ruído 4%	4.74	6.91	11.07
E.P. com ruído 5%	5.62	8.00	11.74
E.P. com ruído 6%	7.29	9.17	12.71
E.P. com ruído 7%	9.06	10.8	13.83
E.P. com ruído 8%	11.25	11.33	15.02
E.P. com ruído 9%	13.52	12.63	16.49
E.P. com ruído 10%	15.08	12.83	18.42
Tempo computacional. Caso sem ruído	45 s	109 s	130 s

A Fig. 5.67 apresenta os erros percentuais de *versus* o % ruído randômico adicionado aos dados exatos, e a Fig. 5.8 apresenta os EP para 10 casos-teste com 1% de ruído randômico nos dados.

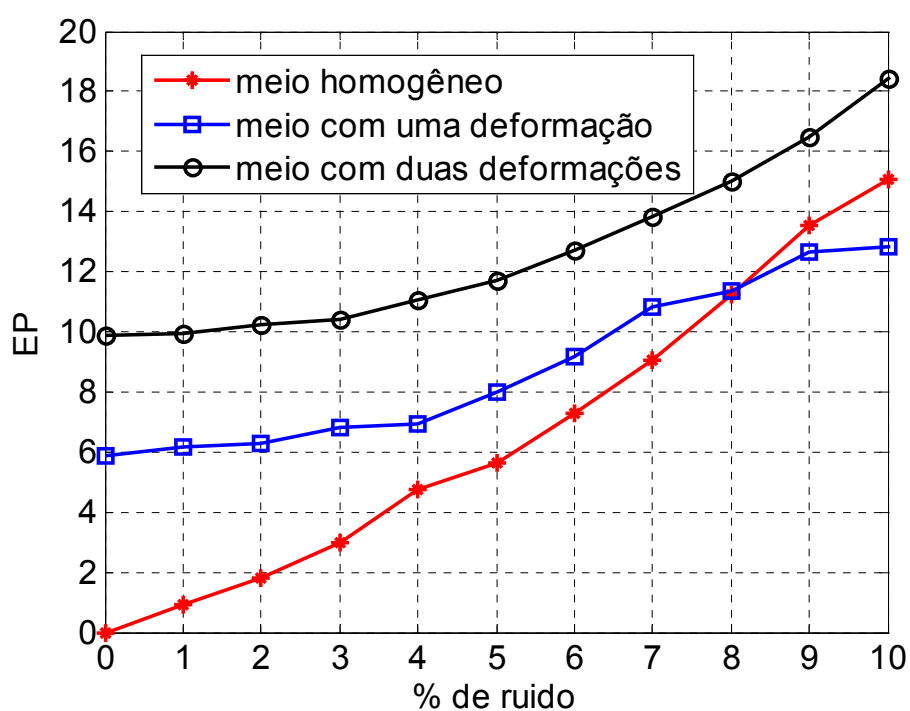


Fig. 5.67- EP *versus* ruído randômico (%) nos dados dos detetores.

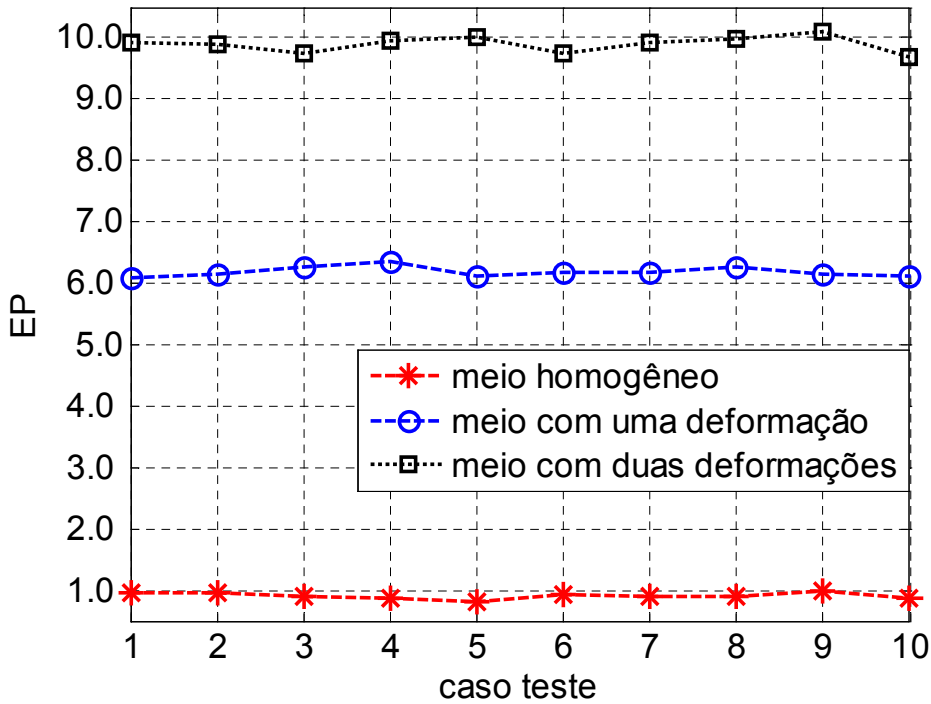


Fig. 5.68- EP para 10 casos-teste com ruído randômico de 1%.

5.3.1.5- Comparação entre o algoritmo *ART* construído com o funcional $\eta_{s,r}$ e o algoritmo q-ART

Consideremos dois meios,, com uma e duas deformações, com geometria similar aos exemplos dadas nas seções 5.3.1.2 e 5.3.1.3 respectivamente. As propriedades dos meios são $\sigma_a=0.0035$ e nas deformações $\sigma_a=0.035$.

A figura 5.64a apresenta os EP obtidos com o algoritmo *ART* empregando o funcional $\eta_{s,r} = \frac{(z_i^s - z_i^r)}{s - r}$.

A figura 5.64b apresenta os EP obtidos com o algoritmo q-*ART*. (CARITA,2004), que emprega a funcional

$$\eta_q = \frac{1}{q+1} \left(z_i \frac{(z_i^q - 1)}{q} \right)$$

Nas Figs. 5.69 e 5.70 empregou-se $\sigma_{ao}=0.001$.

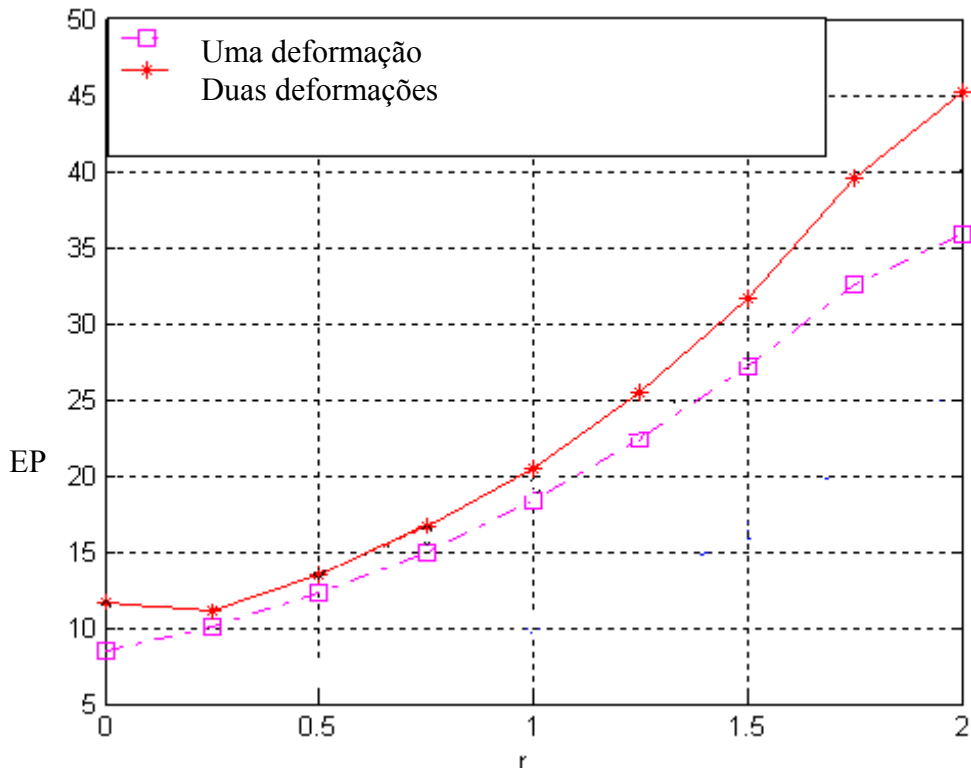


Fig. 5.69- EP versus r. empregando o algoritmo ART e a funcional $\eta_{s,r}$, $s=1$.

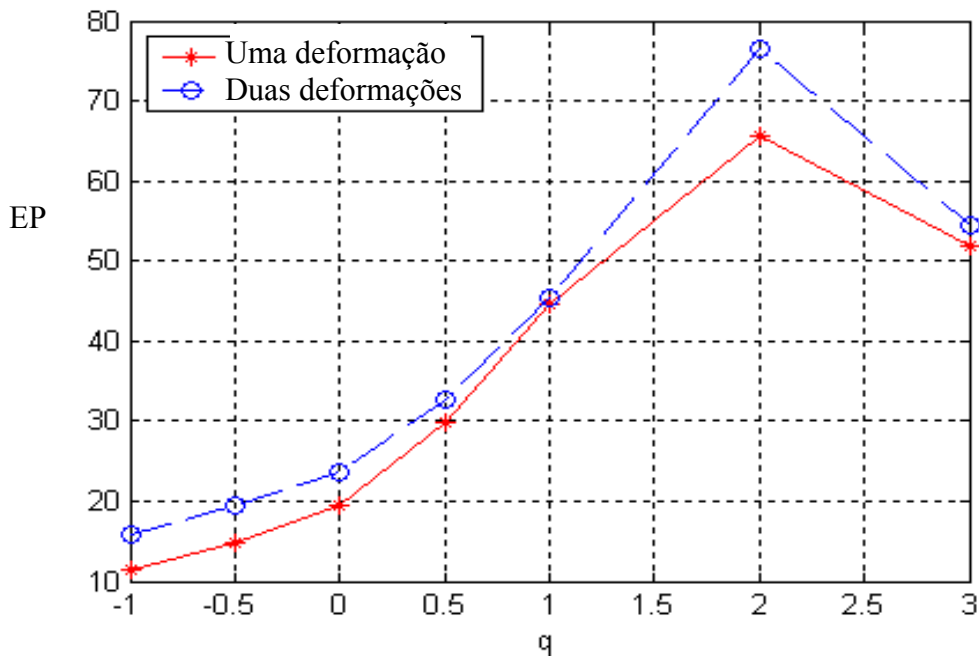


Fig. 5.70- EP versus q. empregando o algoritmo q -ART.

5.3.2 Reconstrução de imagens com a regularização de Tikhonov

Os três exemplos a seguir consideram uma malha quadrada a ser reconstruída. Na Fig. 5.71 observa-se uma malha regular de 10×10 , com 100 σ_a a ser estimado. A

medida dos detectores é gerada faixa a faixa de um sistema que gira até ângulos menores que 180 graus.

Nos casos-teses aqui apresentados o sistema girado é formada por 14 faixas que giram nos ângulos de 0, 15, 30, 60, 75, 90, 105, 120, 150, 165, o que gera 160 medidas de detectores a serem empregadas para a reconstrução da imagem.

Nos exemplos tem-se como interesse, procurar o melhor coeficiente de regularização “ α ” e o melhor “ r ” para ser empregado no Funcional de Tikhonov na reconstrução de imagem. Outro interesse é saber se existe alguma vantagem de empregar o funcional de Tikhonov construído com a distância de Bregman para a reconstrução de imagens, e assim comparar os resultados obtidos pelo método de Tikhonov como os métodos de Levenberg-Marquardt, Mínimos Quadrados e o LSQLIN (programa oferecido pelo o Matlab em seu toolbox de otimização).

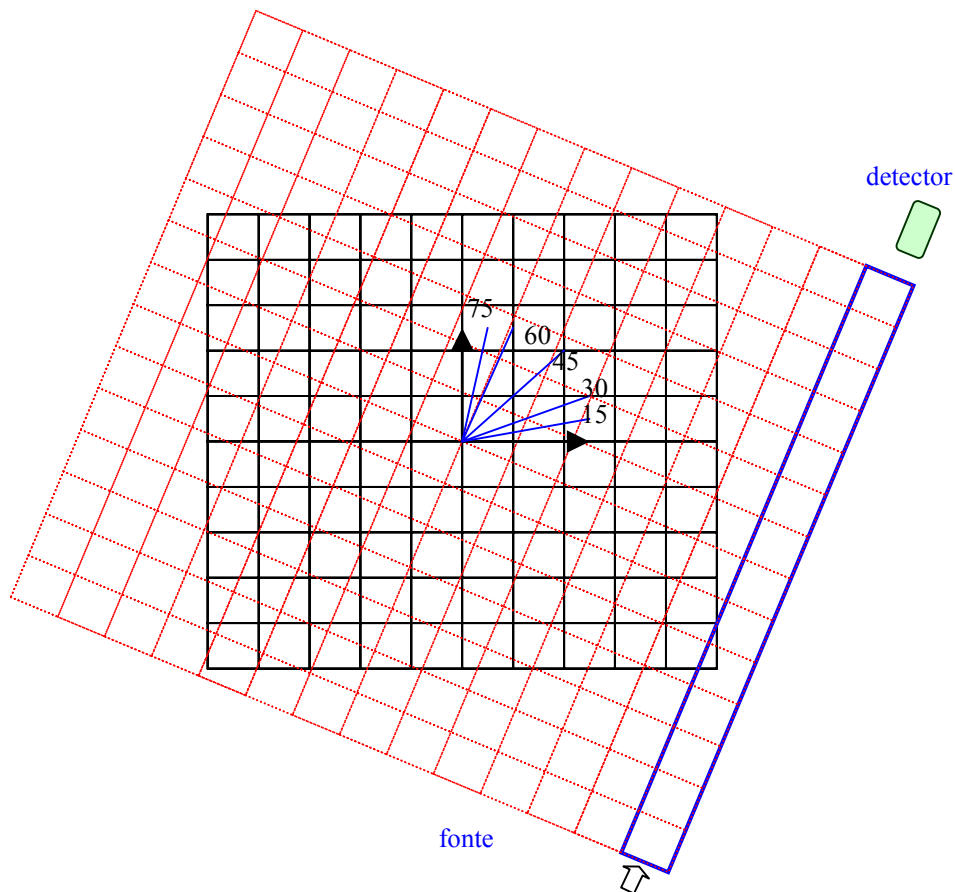


Fig. 5.71.- A malha quadrada e os sistemas girados

Neste trabalho se emprega $\sigma_{ao} = 0.0005$ como valor inicial nos casos testes. Para a solução do problema inverso empregando os métodos de Tikhonov, mínimos quadrados

e Levenberg Marquardt fez-se um programa para cada método em Matlab. O critério de convergência foi de $\sum_{e=1}^E |\sigma_e^k - \sigma_e^{k-1}| < 10^{-6}$, ou $\sum_{e=1}^E |\sigma_e^k - \sigma_e^{k-1}| < 10^{-4}$ nos casos com mais de 1000 iterações. O número máximo de iterações consideradas nos programas foi de 60000.

5.3.2.1 Meio homogêneo

Exemplo 1.-

Empregando dados exatos para a medida dos detectores, neste exemplo considera-se um meio homogêneo com $\sigma_a = 0.08$. A Fig. 5.72 apresenta os erros percentuais obtidos com o método de Levenberg - Marquardt (L.M) e o funcional de Tikhonov (R.T.), quando $r \rightarrow 1$ e $s \rightarrow 1$ (entropia de Shannon), e α varia entre 10^{-10} e 10^{-1} .

Exemplo 2.-

Considerando ruído nas medidas dos detectores de 1% , 2% e 3%, a Fig. 5.73 apresenta os erros percentuais obtidos com o método de Levenberg - Marquardt e o funcional de Tikhonov quando $r \rightarrow 1$ e $s \rightarrow 1$ (entropia de Shannon), *versus* α . Para α variando entre 10^{-10} e 10^{-1} . Note que, o valor ótimo de α varia com o % de ruído, a maior % de ruído maior valor de α . Neste trabalho considerou-se $\alpha = 0.1$, como o valor médio para dados com ruído entre 1 e 10 % de ruído, para a funcional de Tikhonov e para o método de Levenberg-Marquardt.

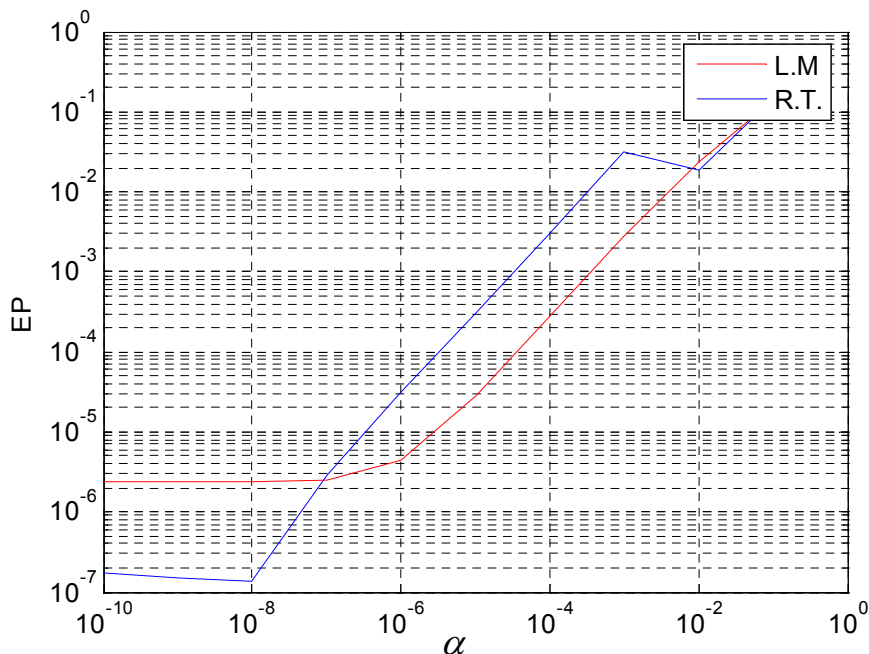


Fig. 5.72.- Erro percentual *versus* α , com dados exatos. Exemplo 1

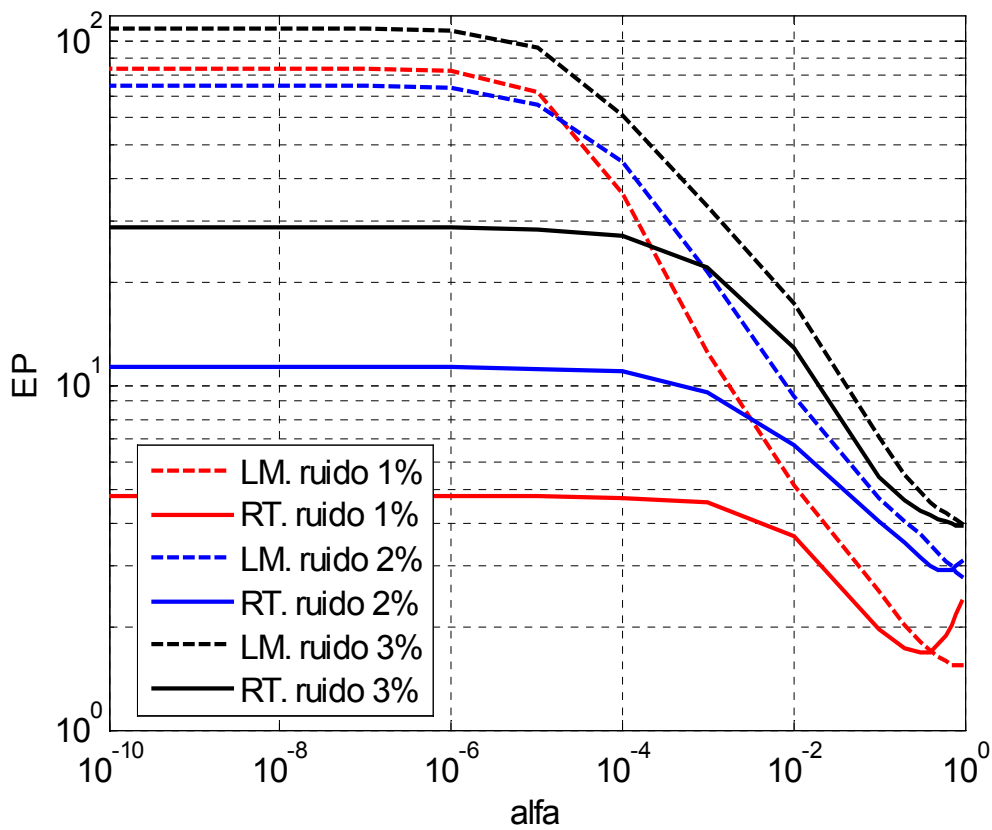


Fig. 5.73.- Erro percentual obtido *versus* α , dados com ruído de 1%, 2% e 3%.

A tabela 5.21 apresenta os erros percentuais obtidos com os métodos: mínimos quadrados, Levenberg - Marquardt ($\alpha=0.1$), o funcional de Tikhonov (quando $r \rightarrow 1$ e $s \rightarrow 1, \alpha=0.1$), LSQLIN

Tabela 5.21.- Os erros percentuais obtidos com diferentes métodos

	Min. Quadrado	Levenberg Marquardt	Método de Tikhonov		LSQLIN
			r=1	r=3	
E.P. sem ruído	0.0000010	0.1407	0.775		0.00015
E.P. ruído 1%	0.2102	2.4025	0.4513	0.254	11.8789
E.P. ruído 2%	0.4327	5.0555	0.9039	0.6055	20.8351
E.P. ruído 3%	0.5641	7.1273	1.2263	1.019	29.9086
E.P. ruído 4%	0.8445	8.9835	2.5108	2.50	39.2304
E.P. ruído 5%	1.0154	11.0196	3.0469	3.03	28.5858
E.P. ruído 6%	1.2157	13.0956	4.6633	3.964	44.8182
E.P. ruído 7%	1.4573	14.4343	5.9177	4.74	41.5656
E.P. ruído 8%	1.9681	15.0129	6.3928	5.324	39.9363
E.P. ruído 9%	1.6776	24.1578	8.8962	5.812	48.4287
E.P. ruído 10%	2.1069	24.9663	10.8003	6.834	47.4838

A Fig. 5.74 apresenta os erros percentuais e as iterações obtidas com o funcional de Tikhonov para valores de r entre 0.1 até 5, empregando dados com ruído de 1%. A curva em vermelho foi obtida com $\alpha=0.1$, a azul com $\alpha=0.01$, e a de cor de vinho com $\alpha=0.001$.

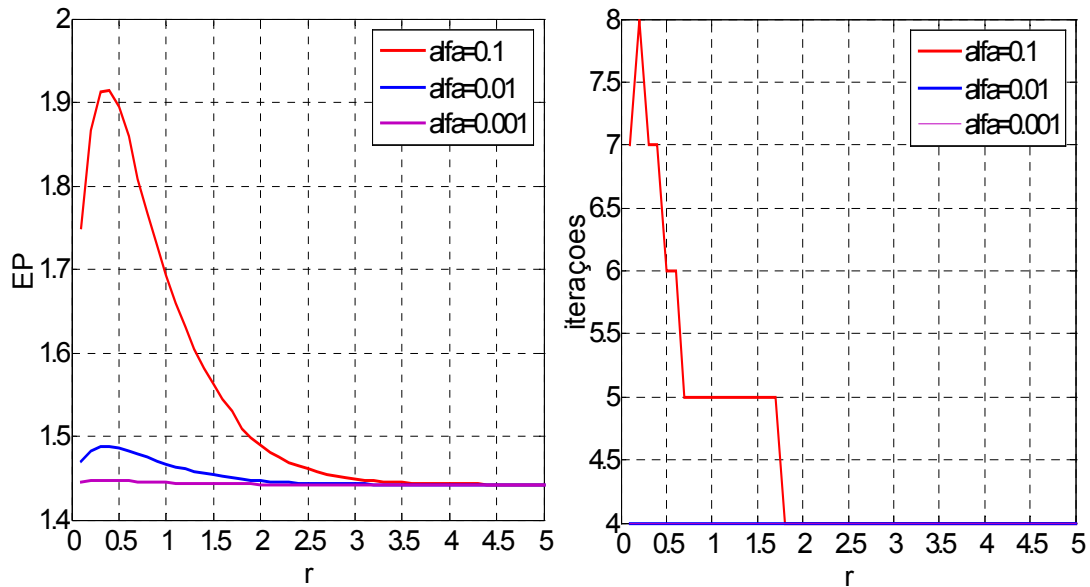


Fig. 5.74.- Erros percentuais *versus* r . Caso homogêneo, dado com 1 % de ruído

A Fig. 5.75 como na Fig. 5.74, apresenta os erros percentuais e as iterações obtidas com o funcional de Tikhonov para valores de r entre 0.1 até 5, empregando dados com ruído do 2%.

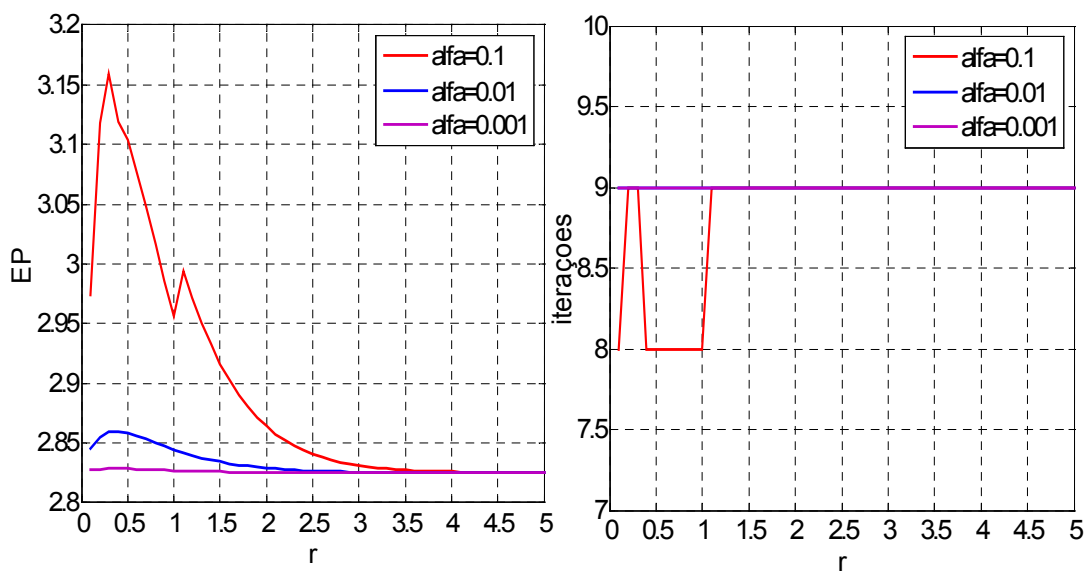


Fig. 5.75.- Erros percentuais *versus* r . Caso homogêneo, dado com 2% de ruído

A Fig. 5.76 apresenta a reconstrução de um meio homogêneo obtido com dados com ruído de 1 % e 2% e 10%, para $r=0.1$ e $\alpha=0.01$

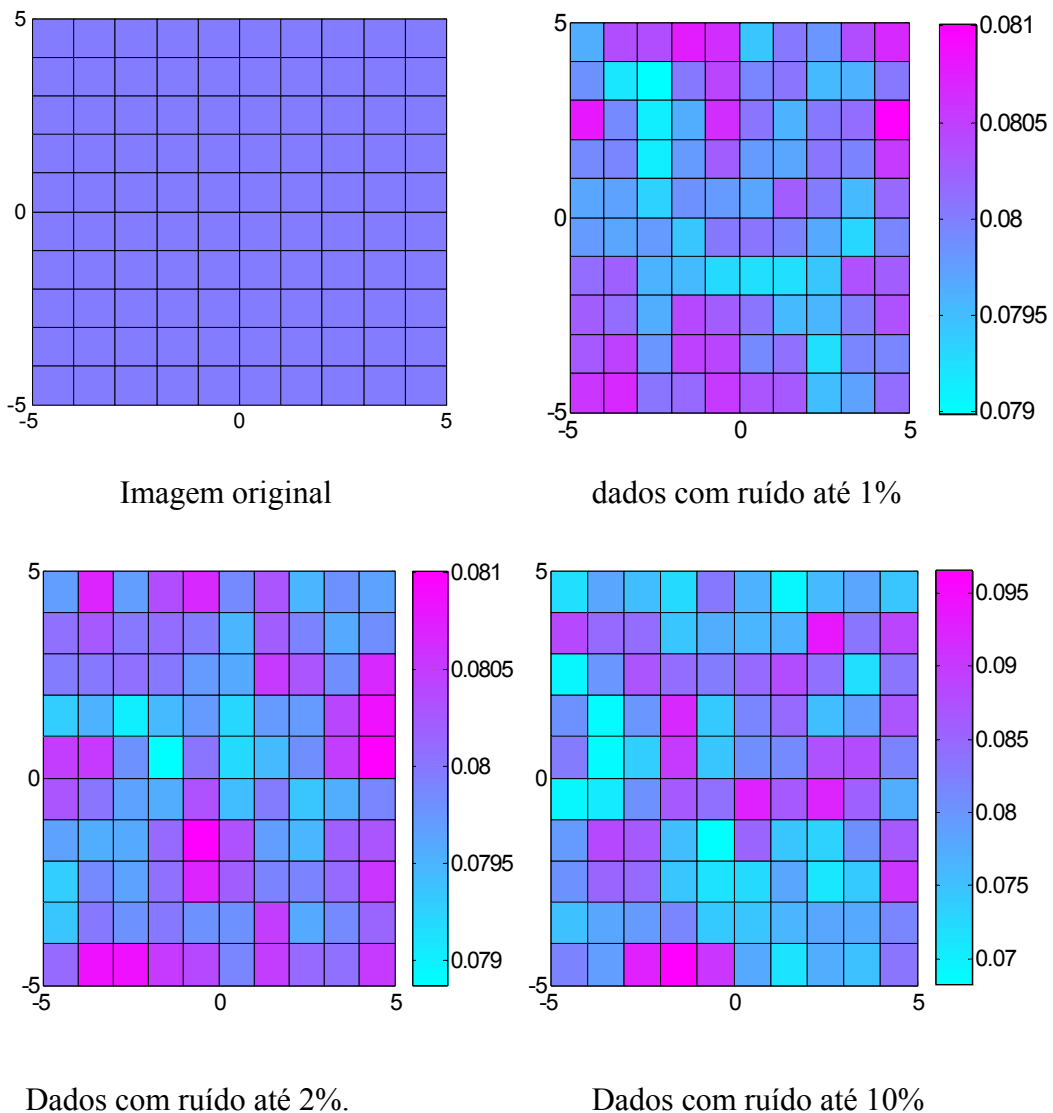


Fig. 5.76.- Reconstrução de imagem. Dados com ruído

5.3.2.2.- Exemplo 2 meio heterogêneo com 1 perturbação

Neste exemplo considera-se um meio com $\sigma_a = 0.04$, com uma perturbação dada por um material incrustado com $\sigma_a = 0.4$. A imagem a ser reconstruída nos casos testes é similar à figura apresentada no final desta item.

A tabela 5.22 apresenta os erros percentuais obtidos com os métodos: mínimos quadrados, Levenberg Marquardt ($\alpha=0.1$), o funcional de Tikhonov (quando $r \rightarrow 1$ e $s \rightarrow 1, \alpha=0.1$), LSQLIN

Tabela 5.22.- Os erros percentuais obtidos com diferentes métodos

	Min. Quadrado	Levenberg Marquardt	Método de Tikhonov		LSQLIN
			r=1	r=0.1	
E.P. sem ruído	14.024	15.437	3.689	2.22	2.7668
E.P. ruído 1%	14.376	15.493	4.251	2.958	12.722
E.P. ruído 2%	15.568	16.602	6.292	4.36	20.716
E.P. ruído 3%	24.743	20.040	7.814	5.52	31.714
E.P. ruído 4%	31.993	20.969	11.062	8.56	39.184
E.P. ruído 5%	48.751	23.800	12.713	9.62	42.954
E.P. ruído 6%	55.160	26.587	14.421	10.89	43.220
E.P. ruído 7%	63.358	30.251	15.318	12.04	42.523
E.P. ruído 8%	71.311	30.158	18.163	14.22	40.037
E.P. ruído 9%	73.326	30.270	21.604	15.64	46.638
E.P. ruído 10%	85.011	32.823	25.366	20.42	48.854

As Figs. 5.77 e 5.78 apresentam os erros percentuais e as iterações, respectivamente, obtidas com o funcional de Tikhonov para valores de r entre 0.1 até 2.1, empregando dados com ruído do 1%. A curva em vermelho foi obtida com $\alpha = 0.1$, a curva azul com $\alpha = 0.01$, e a de cor rosa com $\alpha = 0.001$.

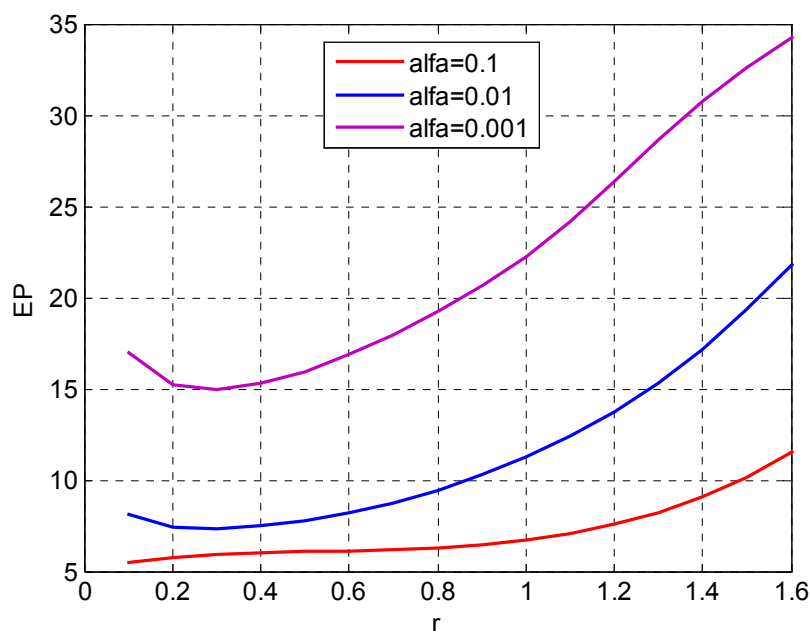


Fig. 5.76. – Os erros percentuais *versus* r. Meio heterogêneo com 1% de ruído

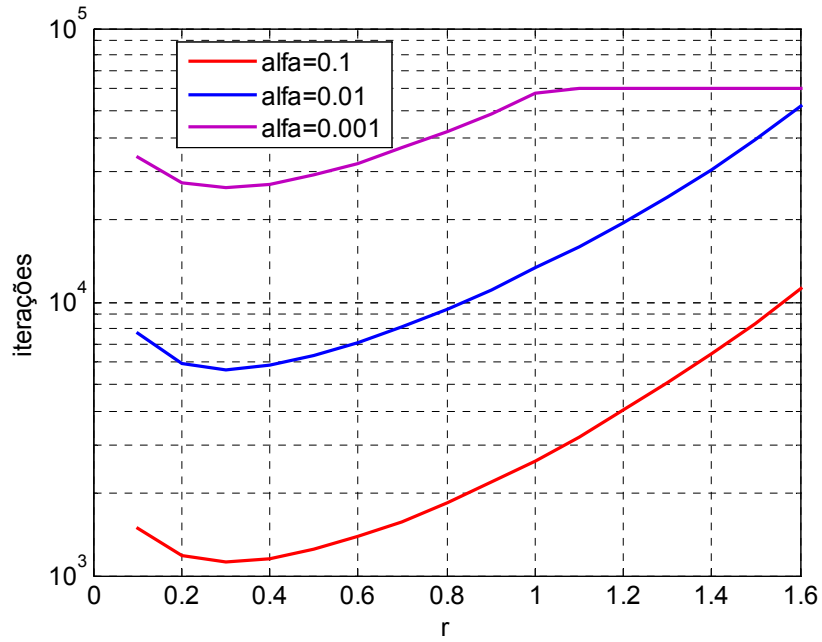


Fig. 5.77. – O número de iterações *versus* r . Meio heterogêneo com 1% de ruído

As Figs. 5.79 e 5.80 apresentam os erros percentuais e as iterações, respectivamente, obtidas com o funcional de Tikhonov para valores r de 0.1 até 2.1, empregando dados com ruído de 2%.

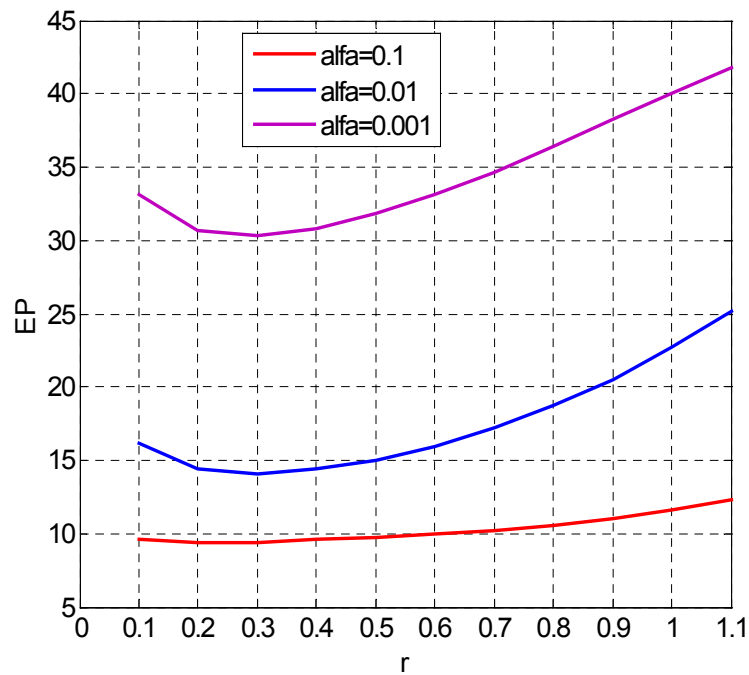


Fig. 5.79. – Os erros percentuais *versus* r . Meio heterogêneo com 2% de ruído.

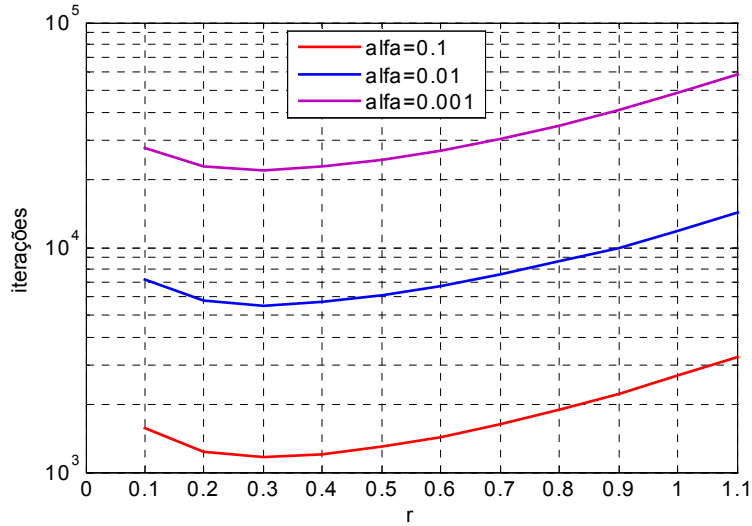


Fig. 5.80. – O número de iterações *versus* r. Meio heterogêneo com 2% de ruído.

Na Fig. 5.81 são apresentadas a imagem original e as imagens reconstruídas com 1%, 2% e 10 % de ruído, empregando $\alpha = 0.1$ e $r=0.2$.

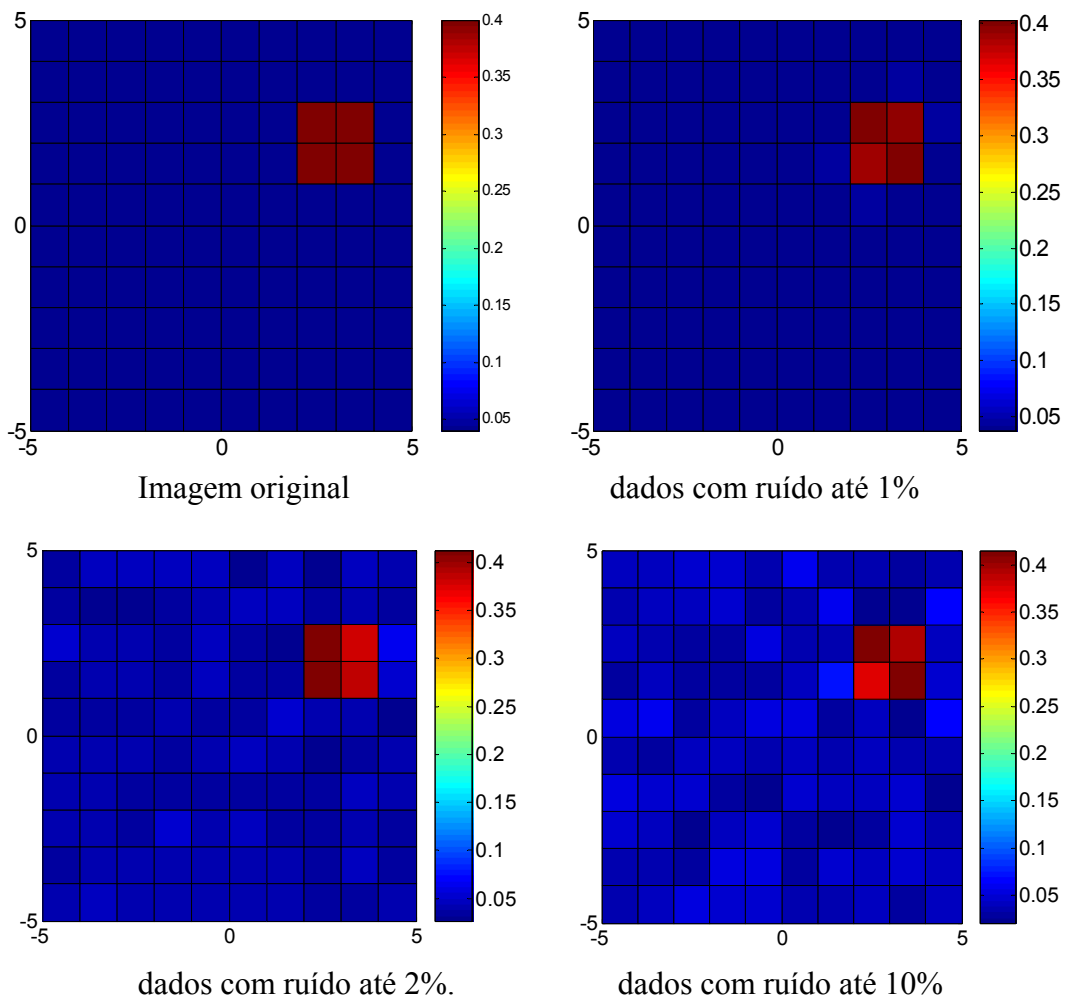


Fig. 5.81.- Reconstrução de imagem. Dados com ruído.

5.3.2.3.- Exemplo 3. Meio Heterogêneo com 2 perturbações

Neste exemplo considera-se um meio com $\sigma_a = 0.04$, e duas perturbações dada por um material incrustado com $\sigma_a = 0.2$ na parte superior e $\sigma_a = 0.4$ na parte inferior do material.

Na tabela 5.23 apresenta-se os erros percentuais obtidos com os métodos: mínimos quadrados, Levenberg - Marquardt ($\alpha = 0.1$), funcional de Tikhonov (quando $r \rightarrow 1$ e $s \rightarrow 1, \alpha = 0.1$), e LQSLIN.

Tabela 5.23.- Os erros percentuais obtidos com diferentes métodos

	Min. Quadrado	Levenberg Marquardt	Método de Tikhonov		LSQLIN
			r=1	r=0.1	
E.P. sem ruído	18.5514	18.6936	5.0473	3.33	1.6197
E.P. ruído 1%	18.5557	18.6272	5.49	3.74	14.3833
E.P. ruído 2%	22.2621	21.3361	7.99	5.36	30.4874
E.P. ruído 3%	27.2871	23.0812	9.21	6.53	31.9666
E.P. ruído 4%	27.3766	22.7333	11.61	9.09	36.5828
E.P. ruído 5%	31.9783	24.8754	13.56	10.15	35.4082
E.P. ruído 6%	53.1881	29.8295	15.28	11.44	43.3586
E.P. ruído 7%	68.9399	30.2696	16.89	13.41	43.3718
E.P. ruído 8%	71.6844	38.1292	19.47	15.12	41.3229
E.P. ruído 9%	65.4074	39.9931	23.65	17.66	51.8408
E.P. ruído 10%	141.4297	43.9278	27.55	21.82	53.4935

As Figs. 5.82 e 5.83 apresentam os erros percentuais e as iterações obtidas com o funcional de Tikhonov para valores de r entre 0.1 até 2.1 e empregando dados com ruído de 1%. A linha vermelha é obtida com $\alpha = 0.1$, a azul com $\alpha = 0.01$, e a de cor rosa com $\alpha = 0.001$.

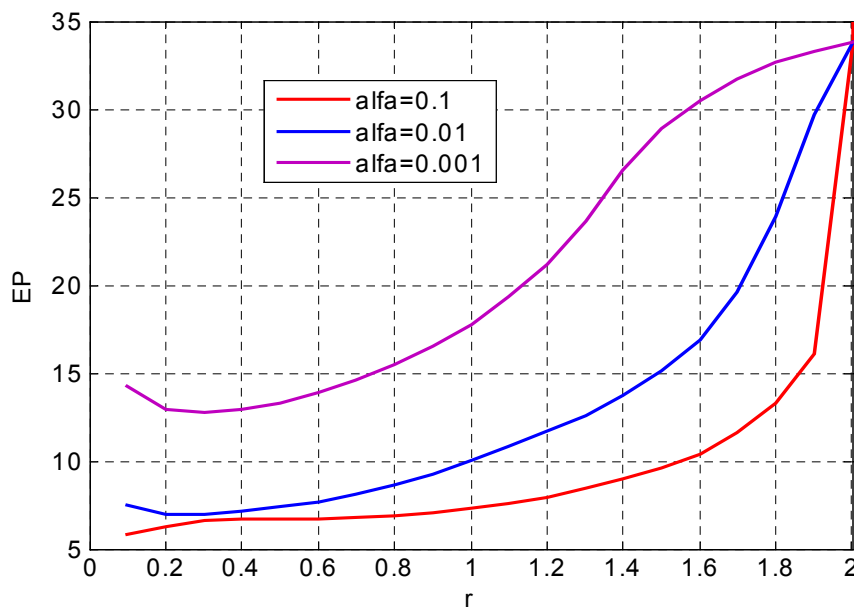


Fig. 5.82. – Os erros percentuais *versus* r. Meio heterogêneo 2 com 1% de ruído.

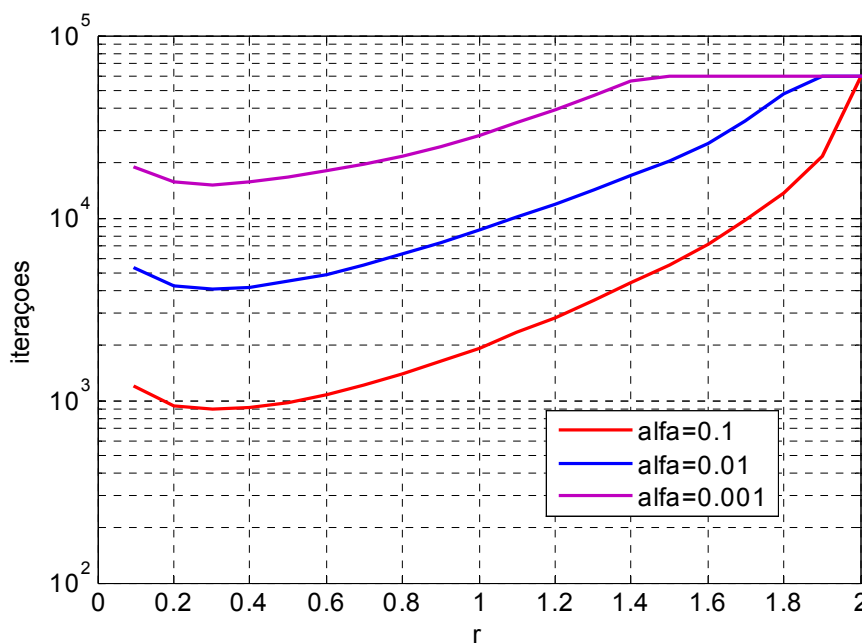


Fig. 5.83. – O número de iterações *versus* r. Meio heterogêneo 2 com 1% de ruído.

As Figs. 5.84 e 5.85 apresentam os erros percentuais e as iterações, respectivamente, obtidas com o funcional de Tikhonov para valores de r entre 0.1 até 2.0, empregando dados com ruído de 2%.

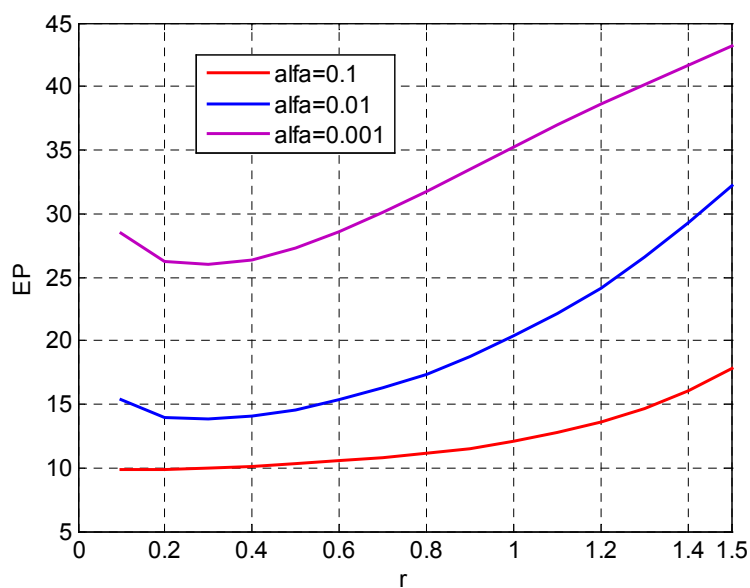


Fig. 5.84. – Os erros percentuais *versus* r. Meio heterogêneo 2 com 2% de ruído.

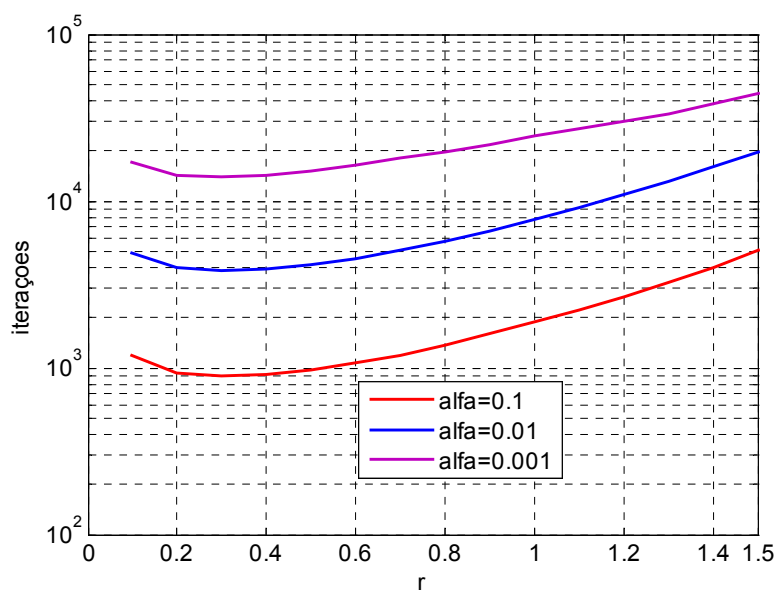


Fig. 5.85. – O número de iterações *versus* r. Meio heterogêneo 2 com 2% de ruído.

Na Fig. 5.86 são apresentadas a imagem original e as imagens reconstruídas com ruído de 1%, 2% e 10% de ruído, empregando $\alpha = 0.1$ e $r=0.2$

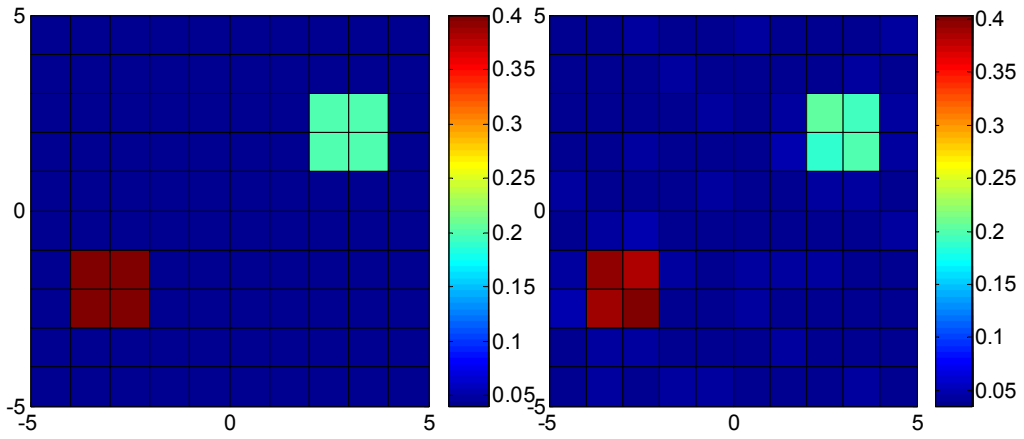
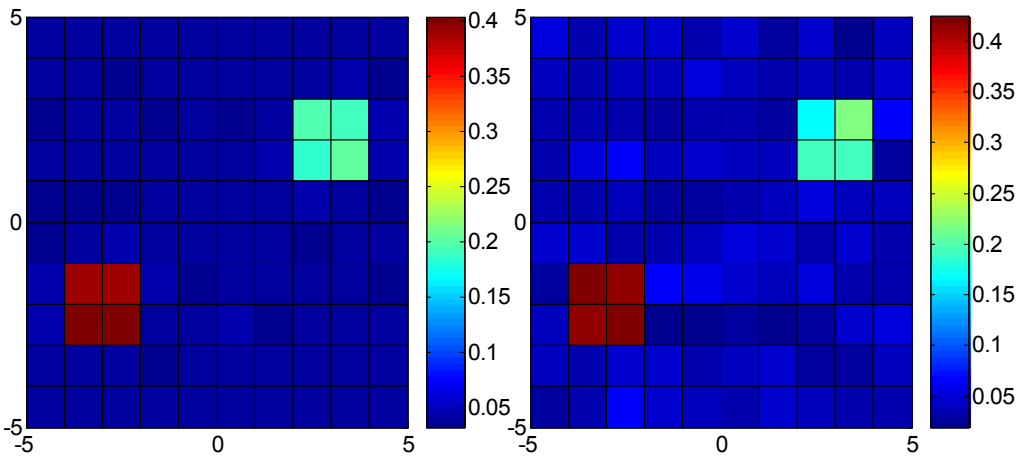


Imagem original

dados com ruído até 1 %



Dados com ruído até 2%.

Dados com ruído até 10%

Fig. 5.86.- Reconstrução de imagem. Dados com ruído.

CAPÍTULO 6

CONCLUSÕES E FUTUROS TRABALHOS

Neste trabalho foram apresentados diversos casos - testes relacionados a problemas de transferência radiativa. Para a solução destes casos - testes foram empregados os métodos de máxima entropia generalizada, algoritmos *ART*, e a regularização de Tikhonov.

Para a estimativa do coeficiente de absorção e espalhamento em meios unidimensionais o método de máxima entropia generalizada demonstrou ser mais eficiente quando a distância de Bregman foi construída com o funcional $\eta_{s,r}$, relacionado à entropia de Sharma e Taneja, com $s=1$, $r=0.2$, $\sigma_{ao} = 0.0001$, $\sigma_{so} = 0.0001$, em comparação com os funcionais relacionadas às entropias de Shannon, Kenyi, Varma e Burn.

Os casos – testes, apresentados para a reconstrução de imagens tomográficas. mostram que o funcional $\eta_{r,s}$, ofereceu melhores resultados ao ser empregado:

- (i) método de Tikhonov em comparação com os métodos de mínimos quadrados e Levenberg Marquardt; e
- (ii) algoritmo *ART* em comparação com os algoritmos *q-ART*, desenvolvido por Carita et al., 2004.

É importante comentar que,

- o melhor valor de tentativa inicial para as incógnitas em todos os casos testes é aquele que é pequeno em comparação ao valor a ser estimado.
- no algoritmo *ART* tem-se que com a maior discretização da malha o erro percentual diminui.
- Na reconstrução de imagem empregando o método de Tikhonov os valores ótimos de s e r foram:
 $s = 1$, $r = 0.1$, $\sigma_{ao} = 0.0005$, $\alpha = 0.1$
- Na reconstrução de imagem empregando o algoritmo *ART*
 $s = 1$, $r = 0.1$ ou 0.2 , $\sigma_{ao} = 0.0005$, $\alpha = 0.1$

Como trabalho futuro propomos refinar a malha empregada nos algoritmos *ART* e a malha empregada na regularização de Tikhonov para a reconstrução de imagem.

Aplicar os algoritmos desenvolvidos nesta tese para a reconstrução de imagem com feixes divergentes.

CAPÍTULO 7

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDALA, F. J. P., 1995, “*Reconstrução de Imagens Tomográficas usando o método de máxima entropia*”. Tese M.Sc. COPPE/UFRJ, Rio de Janeiro, RJ, Brasil
- AHMAD, F., OPPENHEIMER, S. e RAZZAGHI, M., 2003, "A Discrete Bidirectional Reflectance Model in Remote Sensing", *J. Quant. Spectrosc. Radiat. Transfer*, V. 77, pp. 335-343.
- ALIFANOV, O.M., 1974, "Solution of an Inverse Problem of Heat-Conduction by Iterative Methods", *J. Eng. Phys.*, Vol. 26, No. 4, pp. 471-476
- ARRIDGE, S. R., 1999, "Optical tomography in medical imaging", *Inverse Problems*, V. 15, pp. R41-R93.
- BECK, J. V. 1988, ‘Combined parameter and function estimation in heat transfer with application to contact conductance’, *Journal of Heat Transfer* 110, 1046–1058.
- BERROCAL TITO, M. J., ROBERTY, N. C., SILVA NETO, A. J., 2000 ,“Estimativa do coeficiente de absorção e do coeficiente de espalhamento anisotrópico em um meio homogêneo”. *Anais do III Encontro de Modelagem Computacional*. Friburgo, 2000. v.unico. p. 192-195,
- BERROCAL, TITO, M. J., 2001a, “Problemas inversos em transferência radiativa em meios heterogêneos bidimensionais”, Tese M.Sc, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- BERROCAL TITO, M. J., ROBERTY, N. C., e SILVA NETO, A. J., 2002a, ”Absorption and Scattering Coefficient Estimation in two Dimensional Participating Media Using the Generalized Máxima Entropy and Levenberg Marquardt”, *Proc. XIII International Nuclear Atlantic Conference, INAC- ENFIR in CD-ROM*, Rio de Janeiro, Brasil.
- BERROCAL TITO, M. J., ROBERTY, N. C., 2002b, SILVA NETO, A. J., e CABREJOS, J. B., 2002b, ” Estimación de los Coeficientes de Absorción en un Medio Participante Mediante una Aproximación desde el Unidimensional”, *Proc. 5^{to} Encontro de Modelagem Computacional*, Nova Friburgo, Brasil. V. 15, pp. R41-R93.
- BERROCAL, TITO, M. J., 2002c, “*Estimación del coeficiente de absorción y*

- dispersión en medios bidimensionales utilizando el método de máxima entropía generalizada*”, Tese M.Sc. dissertação, Universidad Nacional de Ingeniería, Lima-Perú..
- BERROCAL TITO, M. J, ROBERTY, N. C., SILVA NETO, A. J., 2002d, “Estimación de los coeficientes de absorción en un medio bidimensional participante mediante una aproximación desde el unidimensional”. *Anais do V Encontro de Modelagem Computacional. Friburgo*,. v.unico. p. 284 – 294
- BERROCAL TITO, M. J., MONTERO, R. F. C., ROBERTY, N. C., 2004, ”Uma comparação do algoritmo máxima entropia e uma família de algoritmo q-art para a estimativa do coeficiente de absorção” *Proc. Iberian Latin American Congress on Computational Methods. in CD-ROM*.
- BERTERO, M. AND BOCACCI P., “Introduction to Inverse Problems in Imaging”, IOP Pub. Ltd., London
- BRANCO C. L.M. 2000, “*Reconstrução de imagens tomográficas baseada em arquiteturas reconfiguráveis*”. Tese M.Sc. COPPE/UFRJ, Rio de Janeiro, RJ, Brasil
- BREGMAN, L.M., "The Relaxation Method of Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming", *USSR Computational and Mathematical Physics Journal*, V. 7, pp. 200-217, 1967.
- BURACHIK, R.S., IUSEM, A.N., 1998, “A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Spaces”, *SIAM Journal on Optimization* , v. 8, pp. 197-216.
- BURACHIK, R.S., 1995, *Generalized Proximal Point Methods for the Variational Inequality Problem*. Ph.D. Dissertation, Instituto de Matemática Pura e Aplicada, Riode Janeiro, Brazil, 1995
- BURACHIK, R.S AND SCHEIMBERG, S., 2003, A Proximal Point Method for the Variational Inequality Problem in Banach Spaces, *SIAM Journal on Optimization* , v. 39, pp. 1633-1649,
- BUTNARIU, D. IUSEM, A. N., 1997, “Local Moduli of Convexity and their Application to Finding almost Common Fixed Points of Measureble Families of Operators”. In *Recent Development in Optimization Theory and Nonlinear Analysis*, v. 204, Amer.Math. Soc., Contemporary Mathematics Series, Y Censor and S. Reich.
- CARITA MONTERO R.F. (2000) ,”*Estimativa do coeficiente de absorção em meios participantes heterogêneos bidimensionais com uma partição de domínio*

- construída a partir de feixes divergentes de radiação*”, Tese M.Sc. COPPE/UFRJ, Rio de Janeiro, RJ, Brasil
- CARITA MONTERO, R. F., ROBERTY, N. C. e SILVA NETO, A. J., 2004, "Reconstruction of a combination of the absorption and scattering coefficients with a discrete ordinates method consistent with the source-detector system", *Inverse problem in Science and Engineering*, V. 12, Nro. 1, pp. 81-101.
- CARITA MONTERO, R.F., ROBERTY, N.C. e SILVA NETO, A.J., 2003, "Absorption and Scattering Coefficients Estimation, Solved with a q -ART algorithm", *XXVI Congresso Nacional de Matemática Aplicada e Computacional, Santos, XXVI CNMAC*, São José do Rio Preto, São Paulo.
- CARITA MONTERO, R.F., ROBERTY, N.C. e SILVA NETO, A.J., 2002, "Reconstruction of Absorption and Scattering Coefficients with a Discrete Ordinates Method Consistent with The Source-Detector System", *Proc. 4th International Conference on Inverse Problem in Engineering: Theory and Practice*, Angra dos Reis, Brazil, V. 2, pp.267-274.
- CARITA MONTERO, R. F., ROBERTY, N. C., e SILVA NETO, A. J., 2001, "Absorption coefficient estimation in heterogeneous media using a domain partition consistent with divergent beams", *Inverse Problems in Engineering*, V. 9, pp. 587-617.
- CARITA MONTERO, R. F., ROBERTY, N. C., e SILVA NETO, A. J., 2000a, "A method of discrete ordinates consistent with the source-detector system for radiation transport with anisotropic scattering", *Proc. 3^{er} Encontro de Modelagem Computacional*, Nova Friburgo, Brasil, pp. 190-191.
- CENSOR Y. AND A. LENT, 1981, "An Iterative Row-Action Method for Interval Convex Programming". *Journal of Optimization Theory and Applications*, v. 34 pp.:321—353.
- CHANDRASEKHAR, "Radiative Transfer", Dover Publications. 1960,
- CIDADE, G. A. G., ANTENEODO, C., ROBERTY, N. C. E SILVA NETO, A. J., 2000. "A Generalized Approach for Atomic Force Microscopy Image Restoration with Bregman Distances as Tikhonov Regularization Terms", *Inverse Problem in Engineering*. v. 8, pp 457-472.
- CORMACK, A. M., 1963 "Representation of a Function by its Line Integrals, with some Radiological Application", *J. Appl. Phys.*, v. 34, pp. 2722-2727.

- CSISZER, I., 1991, "Why Least Squares and Maximum Entropy? an Axiomatic Approach to Inference for Linear Inverse Problems ", *The Annals of Statistics*, v. 19, n. 4, pp. 2032-2066.
- DE PIERRO, A.R., IUSEM, A.N. ,1988, "A Finitely Convergent "Row-Action" Method for the Convex Feasibility Problem". *Applied Mathematics and Optimization* v 17 ,pp. 225-235.
- ESTEBAN, D, MORALES, D, 1995, A Summary on Entropy Statistics. *Kybernetika*. v. 31, N.4, pp. 337-346.
- GORDON, R., 1974, "A Tutorial on ART", *IEEE Transactions on Nuclear Science*, v. N. 21, pp.78-93.
- ENGL H. W., HANKE, M. AND A. NEUBAUER, 1996, "Regularization of Inverse Problems", *Kluwer Academic Publishers*, Dordrecht, The Netherlands.
- FIVELAND, W. A, 1984, "Discrete – Ordinates Solutions of the Radiative Transport Equation for Rectangular Enclosures", *Journal of Heat transfer*, v. 106, pp 600-706.
- FIVELAND, W. A., 1987, "Discrete Ordinate Methods for Radiative Heat Transfer in Isotropically and Anisotropically Scattering Media, J". *Heat Transfer*, v. 109, pp. 809-812.
- FIVELAND, W. A., 1991, "The selection of discrete ordinate quadrature sets for anisotropic scattering, HTD-Vol. 160", *Fundamentals of Radiation Heat Transfer*, *ASME*, pp. 89-96.
- IUSEM, A.N., 1995, "On Some Properties of Generalized Proximal Point Methods for Quadratic and Linear Programming". *Journal of Optimization Theory and Applications* v. 85 pp. 593-612.
- HAKIM, A.H., e MCCORMICK, N.J., 2003, "Ocean Optics Estimation for Absorption, Backscattering, and Phase Function Parameters", *Applied Optics*, V. 42, No. 6, pp. 931-938.
- HAVDRA, J.; CHARVAT, F, 1967, "Concept of Structural α -Entropy". *Kybernetika*, v. 3 pp. 30-35.
- HENSEL, E., 1991, "Inverse Theory and Applications for Engineers", Chaps. 1 and 2, *Prentice-Hall*, New Jersey, USA
- HERMAN, G. T., 1980, "Image Reconstruction from Projections", *The Fundamentals of Computerized Tomography*, 1st ed. New York, Academic Press.

- HOUNSFIELD, G. N., 1972, "A Method and Apparatus for Examination of a Body by Radiation such as X or gamma Radiation", In: Patent Specification 1283915, The Patent office, London, England
- IYER H.M AND HIRAKARA K., 1993, "Mathematical Introduction to Seismic Tomography", Chapman & Hall, USA.
- KAPUR, J. N., e KESAVAN, H. K., 1992, "Entropy Optimization Principles with Applications", Academic Press. 1st ed., New York, USA
- KIRSCH, A.,1996, "Ill -Posed Problems in an introduction to the mathematical theory of inverse problems", *Springer-Verlag.*, v. 120, New York, USA.
- KIWIEL, K., 1997, "Additional contact information", *Mathematics of Operations Research*, v.22 pp. 326-349.
- KRESS, R., 1989, "Linear Integral Equations", 1st ed., *Springer Verlag.* New York, USA
- KREYSZIG, E., 1972, "Advanced Engineering Mathematics", *Wiley International.* 3rd ed USA.
- LUENBERGER, D. G. 1970. "Introduction to linear and no linear programming". *Addison – Wesley Iberoamericana.*
- LOPES, R.T., RODEIGUES, J.L, ASSIS,J.T, JESUS, E.F.O. e OLIVEIRA, L.F., 1997, "Evaluation of Microtomography system with an X-Ray Microfocus Tube", *Appl. Radiat. Isot.*, v48, pp1437-1442.
- MCCORMICK, N. J. 2000, "Algorithms for marine optics inverse problems, mini-symposium on inverse problems in medicine, engineering and geophysics", in 'XXIII CNMAC', Santos, SP, pp. 14–23.
- MATLAB – TOOLBOX. Disponível em: <<http://www.mathworks.com>>. Acesso em: 24 nov 2005.
- MENGÜÇ, M.P. e DUTTA, P., 1994, "Scattering Tomography and its Application to Sooting Diffusion Flames", *Journal of Heat Transfer*, v. 116, pp. 144-151.
- MILANDRI, A., ASLLANAJ, F. e JEANDEL, G., 2002, "Determination of Radiative Properties of Fibrous Media by an Inverse Method - Comparison with the Mie Theory", *J. Quant. Spectrosc. Radiat. Transfer*, v. 74, pp. 637-653.
- MODEST, F. M, 1993, "Radiative Heat Transfer". *McGraw-Hill*, Inc. Princeton Road,
- OZISIK M. N., 1973, "Radiative Transfer and Interactions with Conduction and Convection", *John Wiley and Sons*,USA.

- OZISIK, MN AND ORLANDE, HRB, 2000, "Inverse Heat Transfer: Fundamentals and Applications", *Taylor & Francis*, New York.
- PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T., FLANNERY, B.P. (1996), "Numerical Recipes in Fortran 77 - The Art of Scientific Computing", 2ª edição, vol.1, Cambridge Univ. Press. USA
- RAMOS, D. S., "O método do Problema Auxiliar com Regularização de Bregman para Inequações Variacionais". Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brazil.
- REIS, M. L., 1990, "Base Natural e entropia em problemas de Reconstrução de Imagens ". Tese M.Sc, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- REIS, M. L., e ROBERTY, N. C., 1992, "Maximum Entropy Algorithms For Image Reconstruction from Projections", *Inverse Problems*, v. 8, pp. 623-644.
- RÉNYI, A., 1961, "On the Measures of Entropy and Information". Proc. 4th Berkeley Symp. Math. Statist. and Prob., v. 1, pp. 547-561
- RIBEIRO, J. S., 1994, "*Pre-processamento e segmentação de imagens tomográficas*", Tese M.Sc, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- ROBERTY, N.C. e OLIVEIRA, C.R.E., 1995, "Transport Models for Tomography: Theoretical and Practical Considerations", *Anais do X Encontro Nacional de Física de Reatores e Termohidráulica*, pp. 437-442.
- TANEJA, I. J. "Generalized Information Measure and their Application". livro on-line, Brazil. <http://www.mtm.ufsc.br/~taneja/dados/public2.html>
- TANEJA, I. J., 1975, *A Study of Generalized Measures in Information Theory*. Ph.D. Dissertation. University of Delhi, USA.
- TIKHONOV, A. N.; ARSENIN, V. Y. Solutions of Ill-Posed Problems. Wiley, New York, 1977
- TRUJILLO, D.M. & BUSBY, H.R., 1997, Practical Inverse Analysis in Engineering" Ohio State University
- TWOMEY, S., 1996, "Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements". *Dover Publications*, Inc., Mineola, New York.
- SHANNON C. E., 1948, "A Mathematical Theory of Communication. Bell". *System Tech. J.*, v. 27, pp. 379-423.
- SHARMA B.D.; MITTAL D.P., 1975, "New non-Additive Measures of Relative Information". *J. Comb. Inform. & Syst. Sci.*, v. 2, pp. 122-133.
- SHARMA B.D.; TANEJA, I.J., 1975, "Entropy of Type (α, β) and other Generalized Measures in Information Theory", *Metrika*, v. 22, pp. 205-215.

- SHARMA B.D., TANEJA, I. J., 1977, "Three Generalized Additive Measures of Entropy". *Elect. Infor. Kybern*, v. 13, pp. 419-433.
- STRANG, G. "Introduction to applied mathematics". Wellesley – Cambridge press, 1986
- SÓLOV, I. M., 1976, "Método de Monte Carlo", *ed. Mir Moscú*.
- VARMA, R. S., 1966, "Generalizations of Rényi's Entropy of Order α ". *J. Math. Sci.*, v. 1, pp. 34-48.
- VAN TIEL, J., 1984, "Convex Analysis", *John Wiley and Sons Ltd . USA*.

APÊNDICE I

CONCEITOS BÁSICOS DOS MULTIPLICADORES DE LAGRANGE

A1- Conceitos Básicos

Seja um sistema formado por: K equações e p incógnitas, dependendo da relação entre p e K em nosso sistema obtemos:

- i) se $K > p$ pode-se estender o valor de p até que coincida com o de K . isso permitirá que nosso novo sistema tenha solução.
- ii) Se $K = p$, nosso sistema é consistente e a solução é única
- iii) Se $K < p$, esta situação leva a um sistema indeterminado que fornecerá um conjunto infinito de soluções. Em outras palavras, pode-se dizer que não há dados suficientes para resolver o problema na solução desejada.

No item iii) é necessário outro tipo de análise, na qual não mais se procura a unicidade da solução para a formulação inicial, mas a melhor solução. A melhor solução tem um significado abstrato, pois nós não temos muitas referências de quais são os valores ótimos de nossas incógnitas.

Assim inúmeros métodos foram criados, cada um usando um critério que para o seu autor leva à um melhor solução entre aquelas compatíveis com os dados experimentais. Todos esses métodos usam o critério de minimização (ou maximização) de um funcional.

Deste modo, é formulado o seguinte problema de programação convexa, considerando as seguintes definições:

Definição 1 conjunto convexo:: Um subconjunto A de um espaço vetorial V é dito convexo se, para todo $x, y \in A$ temos. Ver Figura A1.1.

$$\lambda x + (1 - \lambda)y \in A, \quad \forall \lambda \in]0, 1[. \quad (A1.1)$$

Definição 2 função convexa: Uma função f , $f : A \subset V \rightarrow R$ é dita convexa se A é convexo e para todo $x, y \in A$, vale a desigualdade

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall \lambda \in]0, 1[. \quad (A1.2)$$

Uma função convexa tem um mínimo e a derivada da função convexa no ponto mínimo é zero (Figura A1.2).

Problemas deste tipo: Seja $F(x, y)$ a função convexa, e $N(x, y)$ uma reta, procura-se encontrar o ponto (x, y) que minimiza a função $F(x, y)$, tal que $x, y \in N(x, y)$. Ver Fig. A1.3. Estes são problemas que podem ser resolvidos utilizando o método de os multiplicadores de Lagrange. Segundo este método, $F(x, y)$ e $N(x, y)$ ajudados de uma nova incógnita, λ , vão a formar uma nova função, $L(x, y, \lambda)$. Para obter os valores de x e y , precisa-se minimizar a nova função $L(x, y, \lambda)$. Então derivando $L(x, y, \lambda)$ com respeito a seus 3 incógnitas e igualando as derivadas a zero, pode-se encontrar os valores x, y, λ . Os λ são denominados multiplicadores de Lagrange, eles fazem possível que sistemas com maior número de incógnitas que de equações sejam resolvidas. Para um melhor entendimento apresentamos um exemplo que nos ajudará a visualizar (Strang, 1986):

Exemplo:

Seja a função convexa $F(x, y) = \frac{1}{2}(x^2 + y^2)$ e a reta $N: 2x - y = 5$

Se pede encontrar o mínimo valor da $F(x, y)$ considerando (x, y) pertence à reta N

Solução 1

Resolvendo: $2x - y = 5$ para y , logo substituindo em $F(x, y)$ para minimizar-lo

$$y = 2x - 5; \Rightarrow F(x, y) = \frac{1}{2}(x^2 + (2x - 5)^2); \Rightarrow F'(x, y) = x + 2(2x - 5) = 0$$

ficando $x = 2 \Rightarrow y = 2x - 5 = -1$

Solução 2

Construímos uma nova função L

$$L = F(x, y) + \lambda (2x - y - 5) = \frac{1}{2}(x^2 + y^2) + \lambda (2x - y - 5) \quad (\text{A1.3})$$

Agora a função $F(x, y)$ será minimizado dentro de L , para encontrar seu mínimo valor deriva-se e iguala-se a zero:

$$\frac{\partial L}{\partial x} = 0 \quad x + 0y + 2\lambda = 0$$

$$\begin{aligned} \frac{\partial L}{\partial y} = 0 & \quad 0x + y + \lambda = 0 \\ \frac{\partial L}{\partial \lambda} = 0 & \quad 2x - y + 0\lambda = 5 \end{aligned} \tag{A1.4}$$

logo $x = -2\lambda, y = \lambda, -5\lambda = 5 \Rightarrow \lambda = -1, x = 2$ e $y = -1$

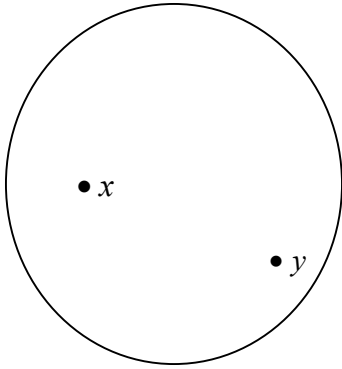


Figura A1.1 - Conjunto convexo

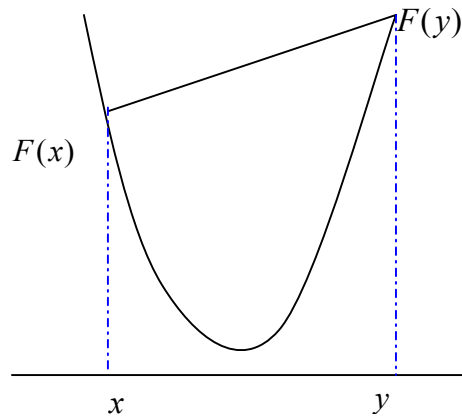


Figura A1.2 - Função convexa

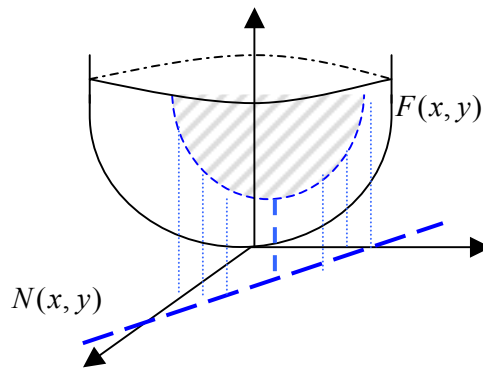


Figura A1.3 Função convexa $F(x, y)$ e a reta $N(x, y)$

Observamos que a solução 1 remove a incógnita e a substitui em $F(x, y)$ e logo minimiza a função, considerando um sistema maior de incógnitas este método emitiria uma solução dependendo da escolha das primeiras variáveis a remover. A solução 2 evita esse problema ao tratar a todas as incógnitas de igual forma o que garanta que sejam obtidos os melhores resultados para nosso problema. Note que a derivada com respeito a λ é a função $F(x, y)$ original.

APÊNDICE II

ALGORITMOS PARA FUNÇÕES NÃO LINEARES

Considerando que: a distância de Bregman $D(\sigma)$ é a função objetivo, σ são as incógnitas do sistema, $F(\sigma)$ é a função erro definida pela Eq.4.2..

1. Condição de primeira ordem

$$\nabla D(\sigma) = 0 \tag{A2.1}$$

2. Condição de segunda ordem

$$\nabla^2 D(\sigma) \geq 0 \tag{A2.2}$$

Em geral, os métodos de otimização começam com uma suposição inicial de um σ_0 que geram uma sucessão de σ_k estimativas até que uma solução σ' seja alcançada. Estes algoritmos iterativos caracterizam-se pela sua robustez, eficiência e precisão.

A2.1 Algoritmos baseados no gradiente

Uma das técnicas aplicadas comumente são os métodos baseados no gradiente. LUEBERGERGER (1984), NASH (1996) e NOCEDAL (1999); provaram que técnicas de otimização baseadas no gradiente, são eficientes computacionalmente, para problemas de grande porte como em tomografia ótica onde ter 10^3 e 10^5 incógnitas é comum. Estas técnicas de otimização usa o gradiente da função objetivo ($\nabla D(\sigma)$), para calcular a seguinte direção de busca v_k (∇D) e o comprimento do passo α_k . A expressão seguinte determina uma seqüência de novas estimativas de $\{\sigma_k\}$:

$$\sigma_{k+1} = \sigma_k + \alpha_k v_k \tag{A2.3}$$

Calculando as novas estimativas de σ_{k+1} apresentam-se duas tarefas:

1. Achar um próprio comprimento de passo (α_k)
2. Calcular a direção de busca, v_k .

Os algoritmos de otimização baseados no gradiente têm a seguinte forma:

1. Defina o valor inicial σ_0 da solução de σ_k
2. Se a função objetivo D é mínimo em σ_k , então fim.
3. caso contrario determina uma nova estimativa $\sigma_{k+1} = \sigma_k + \alpha_k v_k$
4. ir ao passo 2.

As técnicas de gradiente diferem-se principalmente na forma que o comprimento de passo α_k e a direção de busca v_k é calculado.

Em tomografia ótica o método mais comum para determinar a direção de busca, v_k , é a técnica do gradiente conjugada não linear, ARRIDGE (1998) e HIELSCHER (1999). Um método mais complexo é o método Quase Newton, que também será estudado. Lembrar: “todas as técnicas de otimização utiliza o algoritmo *busca de linha* para determinar o comprimento de passo α_k ”.

A2.2 -Algoritmo de busca de linha

O algoritmo de busca de linha é empregada pelos métodos de gradiente conjugado e o método Quase Newton, para achar o seguinte $\sigma_{k+1} = \sigma_k + \alpha_k v_k$ trás a direção de busca v_k .

O algoritmo de busca de linha escolhe uma sucessão de comprimentos de passo e aceita o comprimento de passo, α_k , que cumpre certas condições. Uma condição simples é que a busca de linha proporcione um valor novo da função objetivo com $D(\sigma_k + \alpha v_k) < D(\sigma_k)$. Porém, esta condição nem é sempre conduz a uma redução suficiente em D . Então são formuladas duas condições:

1.- Para um decrescimento suficiente em a função objetivo D deve cumprir com a inequação seguinte

$$D(\sigma_k + \alpha v_k) < D(\sigma_k) + C1 \cdot \alpha \cdot \nabla D(\sigma_k)^T v_k \quad (A2.4)$$

A constante $C1 \in (0,1)$ ela é também chamada de condição de Armijo ou condição de decrescimento suficiente. Ainda esta condição só, não assegura que o algoritmo avance razoavelmente.

2- A uma segunda condição, chamada a condição de curvatura, que exige satisfazer

$$\nabla D(\sigma_k + \alpha v_k)^T v_k \geq C2 \cdot \nabla D(\sigma_k)^T v_k \quad (A2.5)$$

onde a constante $C2 \in (C1,1)$.

Ambas condições 1 e 2, são conhecidas como as condições de Wolfe. Valores típicos de $C2$ é 0.9 quando a direção de busca é escolhida para o método de Quase Newton, e é 0.1 quando a direção de busca é escolhido pelo método do gradiente conjugado não linear.

Além disso, a buscas de linha são distintas entre as buscas exatas e inexatas dependendo do método empregados por calcular a direção de busca v_k .

Uma busca de linha exata executa uma minimização da linha unidimensional $D(\alpha) = D(\sigma_k + \alpha v_k)$ ao longo da direção v_k , achando um apropriado comprimento de passo α . Este método é principalmente usado pelo método do gradiente conjugado. Aqui, a busca de linha é acompanhada em dois passos.

a) O primeiro passo é chamado de seleção, (PRESS, 1992). Esta seleção do mínimo achando três pontos $\sigma_a < \sigma_b < \sigma_c$ ao longo da direção de busca v_k tal que

$$D(\sigma_a) > D(\sigma_b) \text{ e } D(\sigma_c) > D(\sigma_b) \quad (\text{A2.6})$$

Assim, o mínimo está em algum lugar dentro do intervalo $[\sigma_a, \sigma_c]$. Estas seleções podem ser realizadas por um algoritmo eficiente como a regra de seção de ouro, (Press,1992).

b) O segundo passo localiza o próprio mínimo no intervalo $[\sigma_a, \sigma_c]$, que pode ser feito executando uma interpolação parabólica que corrija ajustando um polinômio quadrático à função objetivo D nos três pontos σ_a, σ_b e σ_c . O mínimo desta parábola é achado. Este processo é repetido varias vezes corrigindo uma nova parábola no novo ponto e os prévios dois pontos com a função objetivo menor. Um método geralmente usado por executar esta tarefa é o algoritmo de Brent (PRESS,1992). No mínimo obtemos o comprimento de passo α_k .

Em uma busca de linha inexata ao longo de v_k não vai requerer que $D(\alpha) = D(\sigma_k + \alpha v_k)$ seja mínimo. Neste caso a segunda condição Eq. (A2.5), não é aplicável, é possível utilizar um algoritmo de aproximação regressiva (PRESS, 1992). Esta técnica é usada no método Quase Newton, que no requiere busca de linha exata.

O algoritmo de aproximação regressiva busca os comprimentos de passo candidato α_k empregando a condição suficiente de decrescimentos dada pela Eq. (A2.4) que deve ser cumprida.

A2.3 - Método do gradiente conjugado não linear

Porém, nos não podemos usar o método de gradiente conjugado linear em tomografia óptica porque este método só minimiza bem função quadrática $h(x)$. Geralmente para funções não lineares, como a função objetiva D , nós temos que usar o método gradiente conjugado não linear

O método de gradiente conjugado não linear gera uma seqüência de $\{\sigma_k\}$ convergindo à solução σ na equação (A2.1)

$$\nabla D(\sigma) = 0$$

Este método é derivado do método de gradiente conjugado linear por minimização da função $h(x)$ quadrática convexa, com uma matriz A simétrica e positiva definida:

$$h(x) = \frac{1}{2} x^T A x - b^T x \quad (\text{A2.7})$$

Usando a condição necessária de primeira ordem ($\nabla h(x) = 0$), o método gradiente conjugado linear também pode ser considerado como uma técnica que resolve um sistema de equações linear.

$$\nabla h(\sigma) = A x - b = 0 \quad (\text{A2.8})$$

Os dois problemas, minimizar uma função quadrática convexa e resolver um sistema linear de equação, são equivalentes.

O método de gradiente conjugado linear resolve a Eq. (A2.8) usando uma seqüência de direções de busca $\{w_k\}$ de $\{h_k\}$ ao longo de w_k para determinados x_{k+1} desde um prévio x_k . O esquema de atualização se torna:

$$x_{k+1} = x_k + \alpha_k w_k \quad (\text{A2.9})$$

As direções w_k são conjugadas com respeito à matriz A . Logo temos a definição

$$w_k^T A w_{k'} = 0 \quad \text{para } k \neq k' \quad (\text{A2.10})$$

Por conseguinte nós achamos uma direção conjugada nova

$$w_k = -r_k + \beta_k w_{k-1} \quad (\text{A2.11})$$

onde r_k é o resíduo definido por

$$r_k = Ax_k - b \quad (\text{A2.12})$$

e o escalar β_k

$$\beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}} \quad (\text{A2.13})$$

O comprimento de passo unidimensional, α_k , da função quadrática $h(x)$ ao longo de w_k é determinado por

$$\alpha_k = \frac{r_k^T r_k}{w_{k-1}^T A w_{k-1}} \quad (\text{A2.14})$$

O método de gradiente conjugado não linear difere do modo como o comprimento do passo α_k e o resíduo r_k são determinados.

Em vez de calcular α_k diretamente da Eq. (A2.14), se vai utilizar o algoritmo de busca de linha para minimizar a função objetivo $D(\sigma_k + \alpha v_k)$. Logo o valor do resíduo r_k , para a determinação do escalar β_k , na Eq. (A2.13) é substituído pelo gradiente da função objetivo ($\nabla D(\sigma_k)$).

$$\beta_k = \frac{\nabla D(\sigma_k)^T \nabla D(\sigma_k)}{\nabla D(\sigma_{k-1})^T \nabla D(\sigma_{k-1})} \quad (\text{A2.15})$$

Finalmente a nova direção conjugada é calculada modificando

$$V_k = -\nabla D(\sigma_k) + \beta_k v_{k-1} \quad (\text{A2.16})$$

Os métodos de gradiente conjugado não linear se diferenciam no modo de calcular β_k

O algoritmo de conjugado gradiente não linear pode ser resumido como segue

Convergência fixa, tolerância $\epsilon > 0$

Inicialize x_0 , ponto de partida σ_0

Avalie $D(\sigma_o)$ e $\nabla D(\sigma_o)$

Logo $v_o = -\nabla D(\sigma_o)$

$K=0$

Enquanto $\|\nabla D(\sigma_k)\| > \epsilon$

Determine executando uma minimização de linha de $D(\sigma_k + \alpha v_k)$

Calcule uma estimativa nova $\sigma_{k+1} = (\sigma_k + \alpha v_k)$

Calcule um novo gradiente $\nabla D(\sigma_{k+1})$

Determine β_{k+1} usando Eq. (A2.15)

Calcule a nova busca de direção v_{k+1} usando a equação $v_k = -\nabla D(\sigma_k) + \beta_k \sigma_{k-1}$

Faça $k = k + 1$

Fim enquanto

A2.4 O método de Newton para equações não linear

Os métodos de gradiente conjugado não linear sofrem freqüentemente de taxas de convergência lentas e requerem um número relativamente grande de avaliações de função durante a busca linear exata.

Os métodos de otimização tipo Newton em vez de resolver uma equação não linear diretamente eles são resolvido como uma seqüência de equações lineares. Pelo que, a função objetiva $D(\sigma)$ é substituída por um modelo quadrático $Q(v_k)$ em σ_k

$$D(\sigma_k + v_k) \sim Q(v_k) = D(\sigma_k) + v_k^T \nabla D(\sigma_k) + \frac{1}{2} v_k^T \nabla^2 D(\sigma_k) v_k \quad (\text{A2.17})$$

onde $\nabla^2 D$ é chamado a matriz Hessiana. Logo igualando a zero a derivada de $Q(v_k)$ com respeito a v_k , nós adquirimos um sistema linear de equações, chamadas de equações de Newton, dada por

$$\nabla^2 D(\sigma_k) \cdot v_k = -\nabla D(\sigma_k) \quad (\text{A2.18})$$

então podemos calcular os valores v_k e fazer logo uma estimativa nova dos nossos parâmetros incógnitos $\sigma_{k+1} = \sigma_k + v_k$ pode ser encontrado

Os métodos de Newton raramente são usado em problemas não lineares. Devido que o modelo quadrático $Q(v_k)$ não é uma boa aproximação da função objetivo não linear $D(\sigma_k + v_k)$ em v_k . Um segundo possível problema é que pode não existir a definição positiva da Hessiana ($\nabla^2 D$). Quando ($\nabla^2 D$) não é positivo definido o método de Newton, os valores de v_k atualizado não pode ser definido, pelo que o inverso $(\nabla^2 \phi)^{-1}$ pode não existir. Além disso, até mesmo quando está definido pode não satisfazer a propriedade de descida (NOCEDADL, 1999)

$$v_k^T \nabla D(\sigma_k) < 0 \quad (\text{A2.19})$$

Por conseguinte foram propostos refinamentos para o método de Newton, são chamados de métodos de Quasi Newton. Eles diferem na velocidade de computacional e estabilidade (NOCEDAL,1999). Geralmente, todo o método Quasi-Newton substituem Eqs. (A2.18) com várias alternativas por calcular a direção de busca v_k que tem a forma geral

$$B_k \cdot v_k = -\nabla D(\sigma_k) \quad (\text{A2.20})$$

onde B_k é uma matriz definida positiva, Intuitivamente B_k deveria ser alguma aproximação para $\nabla^2 D(\sigma_k)$. A direção de busca v_k é exigida ser uma direção de descida (Eq. (A2.19)). Além disso, a Eq. (A2.18) é substituída pela atualização da fórmula de Eq.(A2.3). Em resumo, os métodos Quasi Newton requerem da primeira derivada de nossa função objetivo $\nabla D(\sigma_k)$, uma aproximação de B_k à Hessiana $\nabla^2 D(\sigma_k)$, a solução v_k do sistema linear, Eq. (A2.21), e o armazenamento do aproximado da Hessiana $B_k \approx \nabla^2 D(\sigma_k)$

A2.5. Método Quase Newton

Todos os métodos Quase-Newton empregam generalizações do método da secante para calcular a matriz B_k (NASH, 1996) com:

$$\nabla^2 D(\sigma_{k+1}) \cdot (\sigma_{k+1} - \sigma_k) \approx \nabla D(\sigma_{k+1}) - \nabla D(\sigma_k) \quad (\text{A2.21})$$

Substituindo o Hessiana com uma Hessiana aproximada B_{k+1} temos:

$$B_{k+1} \cdot (\sigma_{k+1} - \sigma_k) \approx \nabla D(\sigma_{k+1}) - \nabla D(\sigma_k) \quad (\text{A2.22})$$

então são definidos dois vetores que serão usado repetidamente

$$s_k = \sigma_{k+1} - \sigma_k \quad (\text{A2.23})$$

$$\varphi_k = \nabla D(\sigma_{k+1}) - \nabla D(\sigma_k) \quad (\text{A2.24})$$

substituindo as Eqs (A2.23) e (A2.24) dentro a Eq. (A2.22) obtemos:

$$B_{k+1} \cdot s_k = \varphi_k \quad (\text{A2.25})$$

que é conhecida como a condição de secante ou a condição do Qwasi Newton. Esta condição tem que ser satisfeito pela aproximação de Hessiana B_{k+1} . Além disso, a matriz B_{k+1} só é definida positiva se somente satisfaz a condição de curvatura dada a seguir:

$$s_k^T \cdot \varphi_k > 0 \quad (\text{A2.26})$$

Uma técnica para calcular a matriz B_{k+1} que satisfaz a condição de secante é determinada pela formulada de Broyden-Fletcher - Goldfarb-Shanno (BFGS) fórmula

$$B_{k+1} = B_k - \frac{(B_k s_k) \cdot (B_k s_k)^T}{s_k^T B_k s_k} + \frac{\varphi_k \varphi_k^T}{\varphi_k^T s_k} \quad (\text{A2.27})$$

Depois de cada repetição a nova da matriz B_{k+1} fica atualizada, B_{k+1} está em função de B_k então pode ser calculada começando com $B_o = I$.

O método Quase-Newton pode calcular diretamente o inverso H_{k+1} de B_{k+1} . Agora, dado um H_{k+1} , pode-se calcular mais facilmente a nova direção de busca v_{k+1} usando a Eq. (A2. 51):

$$v_{k+1} = -H_{k+1} \nabla D(\sigma_{k+1}) \quad (\text{A2.28})$$

O inverso da Hessiana é derivado de uma maneira iterativa começando com $H_o = I$ e aplicando a fórmula seguinte

$$H_{k+1} = H_k + \frac{(\varphi_k - H_k s_k) \cdot \varphi_k^T}{\varphi_k^T s_k} - \frac{(\varphi_k - H_k s_k)^T s_k}{(\varphi_k^T s_k)^2} (\varphi_k \varphi_k^T) \quad (\text{A2.29})$$

O método de BFGS requer o armazenamento do inverso da Hessiana H_k para ser usada na próxima iteração, pelo que se precisa de um computador com muita memória. O método LM-BFGS não precisa guardar a inversa da Hessiana, pelo que pode ser usado em sistemas grandes. O LM-BFGS está baseada na formulação inversa do método de BFGS, agora a matriz H_k é substituída pela matriz de identidade I a cada passo de repetição. Nós fazemos esta substituição na Eq. (A2.29) e obtemos a expressão seguinte para a direção de busca v_k usando a Eq. (A2.28) e os vetores s_k e φ_k

$$v_k = -\nabla D(\sigma_k) + \gamma s_k + \lambda \varphi_k \quad (\text{A2.30})$$

com

$$\gamma = -\left(1 + \frac{\varphi_k^T \varphi_k}{s_k^T \varphi_k}\right) \frac{s_k^T \nabla D(\sigma_{k+1})}{s_k^T \varphi_k} + \frac{\varphi_k^T \nabla D(\sigma_{k+1})}{s_k^T \varphi_k} \quad (\text{A2.31})$$

$$\lambda = \frac{s_k^T \nabla D(\sigma_{k+1})}{s_k^T \varphi_k} \quad (\text{A2.32})$$

Um algoritmo baseado na aproximação apresentada tem forma geral:

Tolerância de convergência fixa $\varepsilon > 0$

Inicialize ponto de partida σ_o

Calcule gradiente $\nabla D(\sigma_o)$

Inicializar matriz inversa da Hessiana H_k com a matriz de identidade I

Enquanto $\|\nabla D(\sigma_o)\| > \varepsilon$

BFGS: calcule direção de busca usando a Eq. (A2.28)

LM BFGS: calcule direção de busca usando Eq.(A2.30)

determine α_k executando uma busca de linha sobre v_k

calcule uma estimativa nova $\sigma_{k+1} = \sigma_k + \alpha v_k$

calcule um gradiente novo $\nabla D(\sigma_{k+1})$

defina s_k e φ_k através das Eqs. (A2.23) e (A2.24)

BFGS: calcule usando Eq. (A2.29)

LM-BFGS: calcule λ e γ usando Eqs. (A2.31) e (A2.32)

$k = k + 1$

Fim

Pode ser que a aproximação da matriz Hessiana B_k não seja sempre positiva. Neste caso o modelo quadrático é uma aproximação pobre da função objetivo D e a condição de curvatura, Eq. (A2.26), não está satisfeita.