

Métodos iterativos para reconstrução de imagens

Lila Lisbeth Tenorio Paredes

UFSC - Departamento de Matemática

Campus Trindade

88040-900, Florianópolis, SC

E-mail: lilablossom@gmail.com

Fermín S. Bazán Viloche

UFSC - Departamento de Matemática

Campus Trindade

88040-900, Florianópolis, SC

E-mail: fermin@mtm.ufsc.br

Resumo: Neste trabalho apresentamos um estudo comparativo de métodos iterativos da família *SIRT* e *ART* com o método *LSQR* na solução de problemas discretos mal-postos. Incluímos critérios de parada para os métodos *SIRT* baseados no princípio da discrepância e a regra monótona [3], enquanto para *LSQR*, usamos a regra do produto mínimo introduzido recentemente na literatura [1]. Apresentamos resultados obtidos para um problema da área de tomografia chamado Shepp-Logan phantom incluindo uma comparação com os resultados do método iterativo *LSQR*.

Palavras-chave: *SIRT*, *ART*, critérios de parada, *LSQR*

1 Introdução

Consideramos soluções numéricas para problemas lineares

$$Ax = b, \quad (1.1)$$

onde o vetor de dados b é obtido experimentalmente e portanto sujeito a incertezas: $b = b^{\text{exato}} + e$, em que e é um vetor de ruídos tal que

$$\|b - b^{\text{exato}}\| \leq \delta. \quad (1.2)$$

Em muitas aplicações a matriz A é mal condicionada e o problema acima é dito problema discreto mal-posto. Uma consequência do problema original ser mal posto é que a matriz A do problema é mal condicionada, isto é, pequenos ruídos no vetor b podem produzir grandes alterações na solução x . Uma maneira intuitiva para resolver um problema mal posto é através da substituição do problema original por um problema associado bem posto. Esta substituição é feita conforme a teoria de regularização. A idéia básica nos métodos de regularização é incorporar informações adicionais ao problema, visando soluções estáveis e compatíveis com os dados de entrada. Na literatura encontramos os diferentes métodos de regularização tais como o método de Tikhonov e métodos iterativos tais como *LSQR*, *GMRES*, *RRGMRES*.

Para problemas mal posto de grande porte (tais como tomografia, electrocardiologia, etc) em que o cálculo dos valores singulares da matriz não pode ser realizado, são utilizados métodos de regularização iterativa. A idéia nesses casos consiste em calcular as iteradas e parar o processo iterativo quando a qualidade da iteração começa a incorporar ruído nos dados. Assim, uma

dificuldade comum nos métodos iterativos é determinar a iteração de parada, já que uma vez alcançado o número ótimo de passos, as iteradas tendem a incluir mais informações do ruído ao longo do processo.

Neste trabalho apresentamos um estudo comparativo dos métodos de regularização iterativos SIRT, ART e LSQR aplicado ao problema de reconstrução de imagens médicas, que é uma das muitas aplicações dos problemas inversos. A propriedade de regularização dos métodos iterativos depende fortemente de um apropriado critério de parada que determine a transição entre a convergência e divergência do método. Aqui utilizamos dois critérios de parada, o princípio de discrepância e a regra monótona aplicados ao problema de Shepp-Logan do pacote AIR-Tools [4].

2 Métodos iterativos

Nesta seção apresentaremos alguns métodos iterativos do tipo SIRT e ART.

2.1 SIRT

Os métodos SIRT são do type Landweber e descritos, em forma geral, como

$$x^{k+1} = x^k + \lambda_k T A^T M (b - Ax^k), \quad k = 0, 1, \dots \quad (2.1)$$

onde x^k denota o vetor atual da iteração, λ_k denota o parâmetro de relaxação, e as matrizes M e T são simétricas e definidas positivas. As condições de convergência do método é descrita no teorema abaixo [2].

Teorema 2.1. *O algoritmo 2.1 com $T = I$ converge para \hat{x} , onde \hat{x} é solução do problema $\min \|Ax - b\|_M$ se e somente se*

$$0 < \epsilon \leq \lambda_k \leq \frac{2}{\sigma_1^2} - \epsilon, \quad (2.2)$$

onde $\epsilon > 0$ é um número fixo, arbitrário e muito pequeno, e σ_1 é o maior valor singular de $M^{\frac{1}{2}}A$. Adicionalmente, se $x^0 \in \mathcal{R}(A^T)$, então \hat{x} é a solução de norma mínima na norma 2.

2.1.1 Método clássico de Landweber

O método clássico Landweber tem a seguinte forma,

$$x^{k+1} = x^k + \lambda_k A^T (b - Ax^k), \quad k = 0, 1, \dots \quad (2.3)$$

o que corresponde de definir $M = T = I$ em (2.1).

2.1.2 Método de Cimmino

O método de Cimmino foi originalmente baseado em reflexões sobre hiperplanos mas também tem outra versão com projeções. Seja \mathcal{H}_i o hiperplano definido pela equação linear $\langle a^i, x \rangle = b_i$

$$\mathcal{H}_i = \{x \in \mathbb{R}^n \mid \langle a^i, x \rangle = b_i\}, \quad i = 1, \dots, m.$$

A versão do método baseado em projeção calcula as iteradas através de:

$$x^{k+1} = x^k + \lambda_k \frac{1}{m} \sum_{i=1}^m w_i (\mathcal{P}_i(x^k) - x^k)$$

onde \mathcal{P}_i é o operador projeção no hiperplano \mathcal{H}_i .

2.1.3 Método CAV(Component Averaging)

Este método é uma extensão do método Cimmino que inclui um termo proporcional ao número de elementos não zeros da matriz A . Seja s_j o número de elementos não zeros da coluna j para cada $j = 1, \dots, n$, e $\|a^i\|_S^2 = \sum_{j=1}^n a_{ij}^2 s_j$. O algoritmo do método CAV calcula as iteradas na forma:

$$x_j^{k+1} = x_j^k + \lambda_k \sum_{i=1}^m w_i \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_S^2} a_j^i \quad k = 0, 1, \dots \quad (2.4)$$

onde $w_i > 0$ são os pesos definidos pelo usuário.

2.1.4 Método DROP(Diagonally Relaxed Orthogonal Projections)

O método DROP é outra extensão do método Cimmino inspirado no método CAV. No método DROP também temos os pesos correspondentes a cada equação. Seja $w_i > 0$ o peso definido pelo usuário. O algoritmo DROP é da forma

$$x_j^{k+1} = x_j^k + \lambda_k \frac{1}{s_j} \sum_{i=1}^m \frac{b_i - \langle a^i, x^k \rangle}{\|a^i\|_2^2} a_j^i \quad k = 0, 1, \dots \quad (2.5)$$

para todo $j = 1, 2, \dots, n$ e os $s_j, w_i > 0$ como no método anterior.

2.1.5 Método SART (Simultaneous Algebraic Reconstruction Technique)

O método SART tem a seguinte forma

$$x^{k+1} = x^k + \lambda_k V^{-1} A^T W (b - Ax^k), \quad (2.6)$$

onde $V = \text{diag}(\zeta_j)$ e $W = \text{diag}(\frac{1}{\zeta_i})$, onde ζ^i e ζ_j denotam as somas dos elementos das linhas e colunas da matriz de A respectivamente. Isto é

$$\zeta^i = \sum_{j=1}^n a_j^i \quad \text{para } i = 1, \dots, m \quad (2.7)$$

$$\zeta_j = \sum_{i=1}^m a_j^i \quad \text{para } j = 1, \dots, n. \quad (2.8)$$

2.2 Algebraic Reconstruction Techniques(ART)

Os métodos da família ART tratam o problema projetando sucessivamente em cada hiperplano definido pelas linhas da matriz.

2.2.1 O Método de Kaczmarz

O clássico e mais famoso método do tipo ART é denominado método de Kaczmarz. O algoritmo é dado por

$$\begin{aligned} x^{k,0} &= x^k \\ x^{k,i} &= x^{k,i-1} + \lambda_k \frac{b_i - \langle a^i, x^{k,i-1} \rangle}{\|a^i\|_2^2}, \quad i = 1, 2, \dots, m, \\ x^{k+1} &= x^{k,m}. \end{aligned}$$

Uma variante do método acima é o método simétrico de Kaczmarz.

2.2.2 O Método Simétrico de Kaczmarz

Este método considera um passo do método de Kaczmarz seguido de outro passo do mesmo método onde as equações são usadas em ordem reversa.

$$\begin{aligned} x^{k,0} &= x^k \\ x^{k,i} &= x^{k,i-1} + \lambda_k \frac{b_i - \langle a^i, x^{k,i-1} \rangle}{\|a^i\|_2^2} a^i, \quad i = 1, 2, \dots, m, \\ x^{k+1} &= x^{k,1}. \end{aligned}$$

2.3 LSQR

O método LSQR é um método iterativo baseado na fatoração QR e no processo de bidiagonalização de Lanczos para o problema

$$\min_{x \in \mathbb{R}^n} \|b - Ax\|_2^2 \tag{2.9}$$

O método produz uma sequência de soluções aproximadas x^k para o problema (2.9) minimizando no subespaço de Krylov $\mathcal{K}_k(A^T A, A^T b)$. Isto é, a k -ésima iterada é definida por

$$x^k = \operatorname{argmin}_{x \in \mathcal{K}_k(A^T A, A^T b)} \|b - Ax\|_2^2. \tag{2.10}$$

3 Critério de parada

Nesta seção introduzimos uma regra geral para determinar um índice apropriado de parada k^* para os métodos iterativos SIRT, ART e LSQR.

Para o caso dos métodos SIRT e ART apresentamos o princípio de discrepância e regra monótona descritos em [2].

Proposição 3.1. *Seja $\{x_k\}$ dado por (2.1), onde $T = I$ e $r^k = M^{\frac{1}{2}}(b - Ax^k)$. Sejam $Q = M^{\frac{1}{2}}AA^T M^{\frac{1}{2}}$, $W = I - \frac{\lambda\beta}{2(1-\alpha)}Q$, onde α e β são números reais. Seja $b^{\text{exato}} \in \mathcal{R}(A)$ e \bar{x} qualquer solução do problema $Ax = b^{\text{exato}}$, e $-1 \leq \tau_k \leq 1$. Também, seja $e_k = \|\bar{x} - x^k\|$ e $t_1 = 2\lambda(1-\alpha)\langle r^k, Wr^k \rangle$. Então*

$$e_{k+1}^2 = e_k^2 - \lambda(d_{\alpha,\beta} - 2\tau_k\delta\|M^{\frac{1}{2}}\|r^k) - t_1 \tag{3.1}$$

onde

$$d_{\alpha,\beta} = \langle r^k, (2\alpha + \beta - 1)r^k + (1 - \beta)r^{k+1} \rangle. \tag{3.2}$$

Baseado na proposição acima temos a seguinte regra de parada para os métodos SIRT e ART, chamada de *regra α, β* , definida pela desigualdade

$$\text{regra } \alpha, \beta : \frac{d_{\alpha,\beta}}{\|r^k\|} \leq \tau\delta\|M^{\frac{1}{2}}\|. \tag{3.3}$$

Utilizando a regra (3.3), se busca o menor $k = k_{\alpha,\beta}$ tal que a desigualdade (3.3) é satisfeita para α, β e $d_{\alpha,\beta}$ escolhidos pelo usuário. Se $\alpha = 0.5, \beta = 1$, então $d_{0.5,1} = \|r^k\|^2 = d_{DP}$, o índice de parada $k = k_{0.5,1} = k_{DP}$. Ou seja, a regra α, β coincide com o critério de discrepância de Morozov:

$$\text{regra } DP : \|r^k\| \leq \tau\delta\|M^{\frac{1}{2}}\| \tag{3.4}$$

Se $\alpha = 1, \beta = 0$, então $d_{1,0} = d_{ME} = \langle r^k, r^k + r^{k+1} \rangle$. Neste caso obtemos o critério de parada conhecido como *regra monótona*, a qual sugere tomar como índice de parada o primeiro índice tal que

$$\text{regra } ME : \frac{d_{ME}}{\|r^k\|} \leq \tau\delta\|M^{\frac{1}{2}}\|. \tag{3.5}$$

No método LSQR o critério de parada está baseado no comportamento da norma do resíduo $\|r^k\|_2$ e da solução aproximada $\|x^k\|_2$ em cada iteração. Para este método $\|r^k\|_2 = \|b - Ax^k\|_2$ decresce com k enquanto $\|x^k\|_2$ cresce. Conseqüentemente, se busca escolher a iteração k que minimiza a função $\psi_k = \|r^k\|_2 \|x^k\|_2$ [1], ou seja o critério de parada para LSQR define

$$k^* = \operatorname{argmin}\{\psi_k\}, \quad k = 1, 2, \dots \tag{3.6}$$

4 Resultados Numéricos

Os problemas aqui apresentados são de reconstrução de imagens. O objetivo é obter uma imagem de melhor qualidade, ou seja, com mais detalhes, a partir de uma imagem de qualidade inferior.

4.1 Reconstrução de Imagens

Nos testes consideramos $x^0 = 0$ e o lado direito b contém incertezas; isso ocorre frequentemente pois, em geral, b provém da discretização de uma função contínua, ou é obtida experimentalmente e, portanto, sujeita a erros de medição.

Em todos os problemas, o vetor de perturbações e , gerado pela rotina *randn* do MATLAB,

$$b = b^{\text{exato}} + \|b^{\text{exato}}\|_2 0.01 NL(e/\|e\|_2) \tag{4.1}$$

onde NL é o nível de ruído relativo desejado.

4.2 Imagem Shepp-Logan phantom

A imagem exata é uma modificação do problema Shepp & Logan. A imagem tem tamanho 50×50 pixels e o objetivo é recuperar a imagem original. A matriz A utilizada é gerada pela rotina *paralleltomo* contida no pacote AIR-tools [4]. A imagem embaçada e com ruído é obtida também pela rotina *paralleltomo*.

A figura 1 mostra a imagem exata e as imagens obtidas pelos metodos iterativos Landweber, Cimmino, CAV, DROP, SART, Kaczmarz, Kaczmarz simétrico e LSQR.

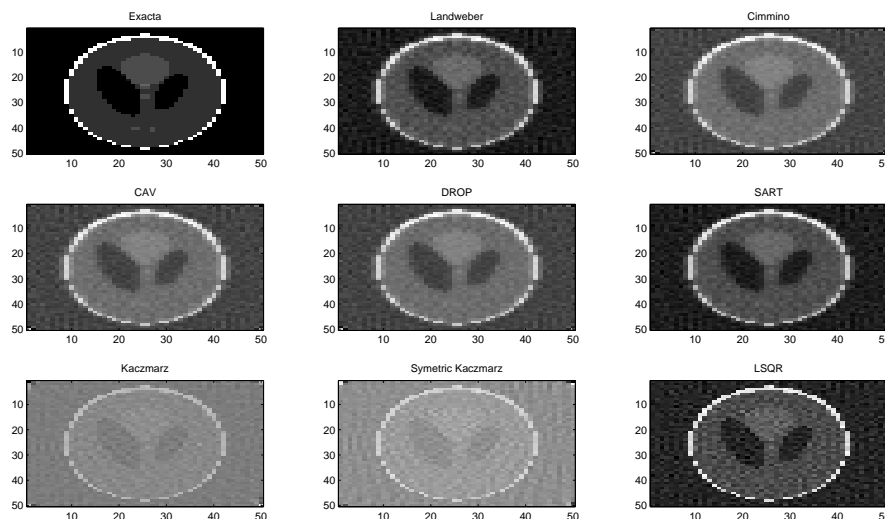


Figura 1: Imagen exata e imagens estimadas com 5% de ruído nos dados.

Métodos	λ	E_{media}	E_{min}	E_{max}	k_{min}	k_{max}
Landweber	2.7295e-04	0.2305	0.2305	0.2305	376	376
Cimmino(projeção)	67.1514	0.2763	0.2754	0.2785	183	400
CAV	1.2003	0.2769	0.2763	0.2794	177	400
DROP	1.2000	0.2749	0.2738	0.2784	151	400
SART	1	0.2359	0.2359	0.2359	290	298
Kaczmarz	0.2500	0.4486	0.4486	0.4486	5	5
Kaczmarz simétrico	0.2500	0.3926	0.3926	0.3926	2	2
LSQR	-	0.2756	0.2756	0.2756	26	26

Tabela 1: Resultados obtidos para o problema Shepp-Logan phantom após 10 execuções com 5 % de ruído nos dados e utilizando o princípio de discrepância.

Métodos	λ	E_{media}	E_{min}	E_{max}	k_{min}	k_{max}
Landweber	2.7295e-04	0.2331	0.2331	0.2331	400	400
Cimmino(projeção)	67.1514	0.2629	0.2619	0.2645	179	400
CAV	1.2003	0.2629	0.2624	0.2639	205	381
DROP	1.2000	0.2618	0.2608	0.2627	224	400
SART	1	0.2412	0.2412	0.2412	302	313
LSQR	-	0.2756	0.2756	0.2756	26	26

Tabela 2: Resultados obtidos para o problema Shepp-Logan phantom após 10 execuções com 5 % de ruído nos dados e utilizando a regra monótona.

Os problemas foram resolvidos 10 vezes com diferentes vetores de perturbação. Os resultados obtidos usando critério de parada de Morozov são apresentados na tabela 1. LSQR usa a regra de parada do produto mínimo (3.6). Podemos ver que os erros nos métodos SIRT são menores em comparação com os métodos ART mas o número de iterações são maiores. Vemos também que os resultados do método LSQR produzem erros compatíveis comos métodos SIRT mas usam menos iterações. Na tabela 2 apresentamos resultados obtidos com a regra ME para os métodos SIRT. Não apresentamos resultados dos métodos ART, pois a regra monótona não é válida para estes métodos. Vemos então que todos os métodos produzem resultados muito parecidos, exceto pelo SART cuja qualidade é levemente melhor.

Na prática os problemas são de grande porte podendo chegar a milhões de variáveis. A implementação dos métodos SIRT e ART pode apresentar dificuldades pelas limitações de memória no computador. Essas dificuldades podem ser contornadas usando estratégias de implementação por blocos. Esse é um tópico de estudo em desenvolvimento.

Referências

- [1] F.S.V. Bazán, Maria C.C. Cunha, L.S. Borges, Extension of GKB-FP algorithm to large-scale general form Tikhonov regularization, Numerical Linear Algebra with Applications, Vol.20, 2013.
- [2] T. Elfving, T. Nikazad, P. C. Hansen , Semi-Convergence and relaxation parameters for a class of SIRT algorithms, Electronic Transactions on Numerical Analysis, Vol.37, pp.321-336, 2010.
- [3] T. Elfving, T. Nikazad, Stopping rules for Landweber-type iteration, Inverse Problems, Vol.23, No 4, pp. 1417-1432, 2007.
- [4] M. Saxild-Hansen, "AIR Tools - A MATLAB Package for Algebraic Iterative Reconstruction Techniques", Tese de mestrado, Technical University of Denmark , 2010.