Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 3, N. 1, 2015.
Trabalho apresentado no XXXV CNMAC, Natal-RN, 2014.

Desempenho do Método de Kaczmarz Aleatório com Parâmetros de Relaxação Bem Escolhidos

Leonardo Bravo Estácio * **Elias Salomão Helou Neto**
Instituto de Ciências Matemáticas e de Computação, ICMC – USP
Universidade de São Paulo – Programa de Ciências de Computação e Matemática Computacional
13566-590, São Carlos – SP
leobravo@icmc.usp.br elias@icmc.usp.br

RESUMO

O Método de Kaczmarz [6], também conhecido na literatura como Técnica de Reconstrução Algébrica (ART - *Algebraic Reconstruction Technique*), é um algoritmo iterativo para solução de sistemas lineares $Ax = b$. Consiste em uma série de projeções ortogonais alternadas realizadas nos hiperplanos definidos pelas equações do sistema.

Uma variação do método clássico, denominada Método de Kaczmarz Aleatório (RK - *Randomized Kaczmarz*), apresentada em 2009 por Strohmer e Vershynin em [7] consiste em selecionar os hiperplanos para serem realizadas as projeções ortogonais através de uma probabilidade proporcional à sua norma euclidiana. Sua abordagem pode ser descrita por

$$x_{k+1} = x_k + \lambda \frac{b_{p(i)} - \langle a_{p(i)}, x_k \rangle}{\|a_{p(i)}\|_2^2} a_{p(i)}^t, \tag{1}$$

sendo que $p(i)$ assume valores entre $\{1, ..., m\}$ com probabilidade $\frac{\|a_{p(i)}\|_2^2}{\|A\|_F^2}$, $\|\cdot\|_F$ denota a norma de Frobenius, $\|\cdot\|_2$ denota a norma Euclidiana ou norma espectral para vetores ou matrizes e λ o parâmetro de relaxação. Através desta probabilidade foi provada a taxa exponencial de convergência esperada de

$$\mathbb{E}\|x_k - x\|_2^2 \leq \left(1 - \frac{1}{\hat{\kappa}(A)}\right)^k \|x_0 - x\|_2^2,$$

sendo que $\hat{\kappa}(A) = \|A^{-1}\|_2^2 \|A\|_F^2$ é o número de condição escalado de A, x_0 é uma estimativa inicial arbitrária e $\mathbb{E}$ denota a esperança sobre a escolha das linhas. Isto assumindo que a matriz A é de posto completo de modo que $\|A^{-1}\| \stackrel{def}{=} \inf \{M : M\|Ax\|_2 \geq \|x\|_2, \forall\, x\}$ esteja bem definido.

Herman iniciou em 1978 [5] os estudos a respeito do parâmetro de relaxação provando que o algoritmo converge caso $0 < \lambda < 2$. A partir daí, vários trabalhos [2, 3, 7] comentam que a taxa de convergência do Método de Kaczmarz pode ser acelerada através da seleção de um bom parâmetro de relaxação (vide [4]).

Este parâmetro de relaxação define como será realizada a projeção ortogonal em cada iteração do algoritmo, por exemplo, caso $\lambda = 1$ a projeção é realizada de maneira exata e gera um ponto pertencente ao hiperplano selecionado. Porém, caso $\lambda > 1$ a projeção ortogonal vai além do ponto pertencente ao hiperplano selecionado (forma sobrerelaxada) e, caso $\lambda < 1$ a projeção ortogonal fica em uma posição anterior ao ponto pertencente ao hiperplano selecionado (forma sub-relaxada).

O objetivo deste trabalho foi utilizar nosso método de escolha do parâmetro de relaxação e aplicar ao Método de Kaczmarz Aleatório de Sthromer e Vershynin [7] de forma a acelerar a convergência do algoritmo e comparar o seu desempenho com o Método dos Gradientes Conjugados para Equações Lineares Não Simétricas e Problemas de Mínimos Quadrados (CGLS - *Conjugate Gradient Method for Unsymmetric Linear Equations and Least Squares Problems*) [1].

De (1) podemos chegar, através de manipulações algébricas, em: $\|x^{k+1} - x^*\|_2^2 = \|x^{k-1} - x^*\| + (\lambda^2 - 2\lambda)\left[\langle Z_{k-1}, x^{(k-1)} - x^* \rangle^2 + (-\lambda \langle Z_{k-1}, x^{k-1} - x^* \rangle \langle Z_k, Z_{k-1} \rangle + \langle Z_k, x^{k-1} - x^* \rangle)^2\right]$, dado que $Z_k = A_{i(k)} / \|A_{i(k)}\|_2$ foram as linhas selecionadas para o cálculo das iterações do algoritmo. A fim de obtermos (2) e (3), partimos deste resultado e avaliamos primeiro a esperança condicional a respeito da escolha de $Z_0, ..., Z_{k-1}$ e então a esperança condicional a respeito da escolha de $Z_0, ..., Z_{k-2}$.

*bolsista de mestrado FAPESP

Para uma matriz $A \in \mathbb{C}^{m\times n}$ de posto completo, nossas estimativas do parâmetro λ são obtidas por:

$$\min_{\lambda\in(0,2)} \left(\lambda^2 - 2\lambda\right)\left(2\|A\|_F^2 + \lambda^2\delta - 2\lambda\|A\|_2^2\right) \quad (2) \qquad \min_{\lambda\in(0,2)} \left(\lambda^2 - 2\lambda\right)\left(2\|A\|_F^2 + \lambda^2\rho - 2\lambda\underline{\sigma}^2\right). \quad (3)$$

sendo que $\rho = \max_{i=1\ldots m} \frac{\|Aa_i\|_2^2}{\|a_i\|_2^2}$, $\delta = \min_{i=1\ldots m} \frac{\|Aa_i\|_2^2}{\|a_i\|_2^2}$ e $\underline{\sigma}$ é o menor valor singular da matriz A. Nossos resultados numéricos apontam que a escolha entre as estimativas (2) e (3) se dá a partir da distância $\|x_0 - x^*\|_2$ ou do resíduo $\|b - Ax_0\|_2$. Caso seja mínima, escolhemos (2), se não, (3).

A minimização (3) é otimista em relação a taxa de convergência do algoritmo e, desta forma, calcula o parâmetro de relaxação λ acreditando que o algoritmo terá uma boa convergência. Por outro lado, a minimização (2) é pessimista neste mesmo aspecto. Em geral minimizando (3) obtemos parâmetros $\lambda > 1$ e minimizando (2) obtemos $\lambda < 1$.

Utilizando o Método RK e o CGLS em um sistema $Ax = b$, sendo que $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$, $m = 5000$, $n = 500$, $\|A\|_F^2 = 1.29 \times 10^3$, $\underline{\sigma} = 19.82$, $\|A\|_2 = 1.12 \times 10^3$, a parte real dos elementos aleatórios da matriz A e do vetor b estão no intervalo $[0, 1]$ e a parte imaginária estão no intervalo $[0, i]$. Executamos ambos os algoritmos e analisamos o tempo de convergência do método através do erro $\|x_k - x^*\|_2$ utilizando, no método de Strohmer, projeções com o valor do parâmetro igual a $\lambda = 1$ (projeções exatas do artigo original [7]) e também o valor calculado pela estimativa (3) $\lambda = 1.2358$, obtendo a Figura 1.

Analisando a Figura 1 vemos que apesar de reduzir bastante o erro em cada iteração, o algoritmo CGLS se mostrou bem mais custoso do que o Método de Kaczmarz Aleatório de Strohmer. Além disso, vemos o ganho na velocidade de convergência utilizando a estimativa calculada minimizando (3), que é um resultado bastante satisfatório.

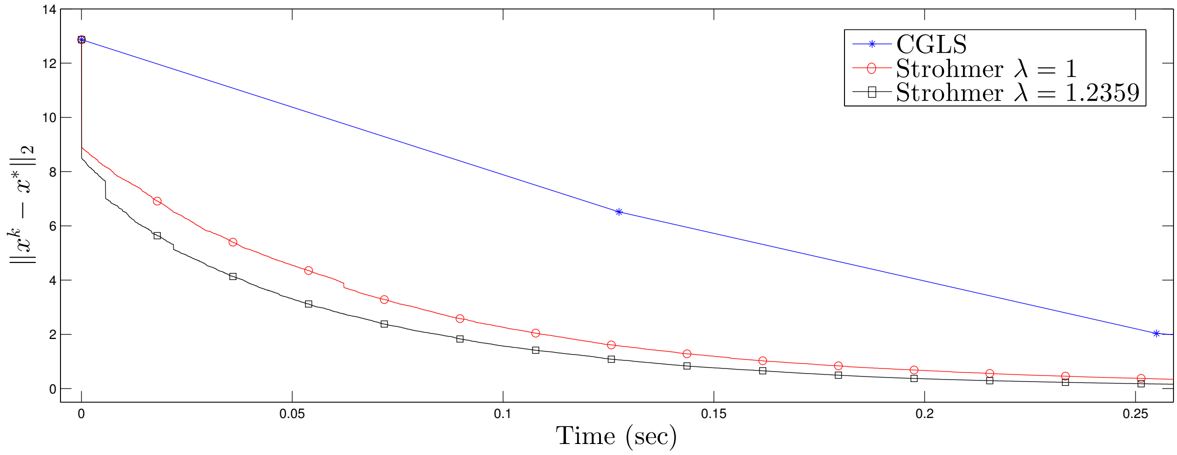

Figura 1: Comparação da velocidade de convergência do Método RK e do Método CGLS

Palavras-chave: *Método de Kaczmarz, Parâmetro de Relaxação, CGLS*

Referências

[1] O. Axelsson. Conjugate Gradient type methods for unsymmetric and inconsistent systems of linear equations. *Lin. Alg. and its Appl.*, 29(1):1–16, 1980.

[2] Y. Censor, P. Eggermont, and D. Gordon. Strong underrelaxation in Kaczmarz's method for inconsistent systems. *Numer. Math.*, 41(1):83–92, 1983.

[3] Y. C. Eldar and D. Needell. Acceleration of randomized Kaczmarz method via the Johnson-Lindenstrauss lemma. *Numer. Alg.*, 58(2):163–177, 2011.

[4] L. B. Estácio and E. S. Helou Neto. Escolha do parâmetro de relaxação e implementações paralelas do Método de Kaczmarz. In *Anais do CMAC (Sudeste)*, volume 1, pages 701–706, 2013.

[5] G. T. Herman, A. Lent, and P. H. Lutz. Relaxation methods for image reconstruction. *Commun. ACM*, 21(2):152–158, 1978.

[6] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bull. Acad. Pol. Sci.*, 35(1):355–357, 1937.

[7] T. Strohmer and R. Vershynin. A randomized Kaczmarz algorithm with exponential convergence. *J. Fourier Anal. Appl.*, 15(2):262–278, 2009.

