

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Evolução Diferencial Melhorada Utilizando Processamento Paralelo Aplicado à Solução de Grandes Sistemas Lineares

Milena Almeida Leite Brandão¹

José Laercio Doricio²

Curso de Graduação em Matemática, Universidade Federal de Uberlândia,

Faculdade de Ciências Integradas do Pontal, Ituiutaba, MG

Sezimária de Fátima Pereira Saramago³

Curso de Graduação em Matemática, Universidade Federal de Uberlândia, Uberlândia, MG

Resumo. Neste trabalho é apresentado o método de otimização estocástico Evolução Diferencial Melhorada utilizando processamento paralelo (EDMP) aplicado à solução de grandes sistemas lineares de um problema de identificação de forças dinâmicas. Os resultados obtidos com EDMP são comparados com os encontrados com os algoritmos Evolução Diferencial, GMRES e com as soluções analíticas. O objetivo é aplicar EDMP sem se preocupar com pré-condicionadores e com as características da matriz de coeficientes, mesmo que a mesma seja não quadrada, não simétrica e nem positiva definida. Embora a matriz do sistema seja retangular formada por três blocos triangulares, o problema de identificação de forças dinâmicas foi resolvido sem usar nenhuma técnica de regularização da matriz de coeficientes. Os resultados obtidos com EDMP mostraram-se muito satisfatórios.

Palavras-chave. Evolução Diferencial Melhorada, Processamento Paralelo, Grandes Sistemas Lineares.

1 Introdução

Segundo [4], mais de 75% dos problemas matemáticos encontrados em aplicações científicas e industriais envolvem a resolução de sistemas lineares em algum estágio. No entanto, estes sistemas geralmente possuem dimensão elevada o que dificulta a obtenção da sua solução de forma algébrica. De forma geral, os sistemas lineares podem ser resolvidos pelos métodos diretos e indiretos.

Os métodos diretos são aqueles que após um número finito e bem determinado de operações obtém-se a solução exata, a menos de erros de arredondamento, de um sistema linear.

¹milena@pontal.ufu.br

²jldoricio@pontal.ufu.br

³saramago@ufu.br

Os métodos indiretos ou iterativos são aqueles que, a partir de uma aproximação inicial, obtém-se a solução aproximada do sistema linear por meio de um processo iterativo. Os métodos iterativos são classificados em estacionários e não-estacionários. Nos métodos estacionários cada solução aproximante é obtida a partir da solução anterior aplicando-se sempre o mesmo processo. Nos métodos não-estacionários a matriz de iteração não é constante e por isso a cada iteração procura-se obter a melhor aproximação da solução de acordo com certas restrições e utilizando informações das iterações anteriores. Estes métodos se baseiam na teoria de otimização, procurando o mínimo de um funcional em uma direção de busca bem definida. A maioria destes métodos trata apenas de sistemas lineares onde a matriz dos coeficientes é simétrica e positiva definida. Existem algumas técnicas que podem ser utilizadas nos métodos iterativos para acelerar a convergência, por exemplo, o pré-condicionamento da matriz de coeficientes, que consiste em determinar uma matriz não singular, de forma que o novo sistema possua uma taxa de convergência maior.

O objetivo desta pesquisa é aplicar técnicas heurísticas para encontrar a solução de sistemas lineares sem se preocupar com as características da matriz de coeficientes, mesmo que a mesma seja não simétrica e nem positiva definida. Muitos métodos de otimização podem ser usados por exemplo, em [2] e [3], é utilizado Evolução Diferencial para resolver sistemas lineares. O algoritmo desenvolvido para a Evolução Diferencial Melhorada, implementada em paralelo, será aplicado na solução de grandes sistemas lineares sem a necessidade de utilização de pré-condicionadores. A metodologia utilizada consiste em minimizar o vetor resíduo $r = Ax - b$ do sistema linear.

2 Evolução Diferencial Melhorada implementada em processamento paralelo

O algoritmo da Evolução Diferencial Melhorada implementada em processamento paralelo (EDMP) é uma combinação dos métodos de otimização Evolução com Conjuntos Embaralhados [5] e Evolução Diferencial [9].

O algoritmo da Evolução Diferencial (ED) começa criando uma população de indivíduos amostrados aleatoriamente a partir da região viável usando distribuição de probabilidade uniforme.

A ideia principal da ED é gerar novos indivíduos, denotados vetores doadores, pela adição da diferença ponderada entre dois ou mais indivíduos aleatórios da população a um terceiro indivíduo. Esta operação é chamada *mutação*.

Em seguida, as componentes do indivíduo doador são misturadas com as componentes de um indivíduo escolhido aleatoriamente (denotado vetor alvo), para resultar o chamado vetor experimental. O processo de misturar os parâmetros é referido como *cruzamento*. Se o vetor experimental resultar um valor da função objetivo menor que o vetor alvo, então o vetor experimental substitui o vetor alvo na geração seguinte. Esta última operação é chamada *seleção*. O procedimento é finalizado quando algum critério de parada é obedecido.

A EDMP inicia-se como o algoritmo da ED usual por criar no processador mestre

uma população de indivíduos. Os membros da população são classificados em ordem crescente de valores da função objetivo e distribuídos em k subpopulações ($k \in \mathbb{Z}$ é menor ou igual ao número de processadores) e enviados para, no máximo, k processadores. Cada conjunto independentemente executa a ED. Na etapa da evolução, os conjuntos são obrigados a se misturar e os pontos são realocados para garantir a troca de informações. As subpopulações são então reagrupadas no processador mestre, embaralhadas e divididas novamente em subpopulações. Este processo continua até que algum critério de parada é alcançado.

O algoritmo da EDMP foi implementado em C++ e as comunicações entre os processadores são realizadas por meio da biblioteca Message Passing Interface (MPI). Um detalhamento do método EDMP pode ser encontrado em [1].

3 Problema de Identificação de Forças Dinâmicas

Um processo de identificação de forças pode ser conduzido por meio de um modelo matemático que relacione as respostas medidas de um sistema mecânico com as possíveis forças aplicadas sobre o mesmo.

Os modelos de resposta são construídos a partir de equações diferenciais que relacionam as entradas com as saídas do sistema em observação. Tais modelos podem ser estabelecidos por meio de modelagens nas quais os valores das saídas do sistema são obtidas por meio de experimentos com valores das entradas conhecidas. Assim, não é necessário ter conhecimento sobre a estrutura física interna do sistema.

O problema inverso da identificação de forças consiste em estimar as forças aplicadas sobre um sistema levando em consideração um modelo matemático e as respostas observadas. O modelo matemático inverso é determinado invertendo as equações matemáticas resultantes da modelagem do sistema físico.

Será considerado neste trabalho um problema teórico apresentado por [7] e [1] no qual o sistema é excitado por uma força harmônica.

3.1 Sistema excitado por uma força harmônica

O objetivo é resolver o seguinte sistema linear:

$$X = HF \tag{1}$$

onde

$$X = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{Bmatrix}; \quad F = \begin{Bmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{Bmatrix} \quad e \quad H = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{p-1} & h_{p-2} & \cdots & h_0 \end{bmatrix} \tag{2}$$

sendo que,

$$h(i, j) = A_0 \left(e^{-\xi\omega t(j+1)} \right) \text{sen} \left(\sqrt{1 - \xi^2} \omega t(j + 1) \right) \tag{3}$$

com $i, j = 1, 2, \dots, N-1$, e $j \leq i$.

Considerando $A_0 = e^{-3}$, $\xi = 0,08$, $\omega = 2\pi 500$, $dt = 1/8192$, $t \in [0, 1]$ e $N =$ comprimento(t). A força é dada por:

$$F = 5\text{sen}(2\pi 300t) + 1,5\text{cos}(2\pi 300t) \quad (4)$$

4 Simulações Numéricas

Visto que se conhece a matriz H e o vetor X , o sistema linear (1) será resolvido usando ED, EDMP e o método iterativo Resíduo Mínimo Generalizado (GMRES) para encontrar o vetor F . A função objetivo é dada pelo resíduo $r = |X - HF|$. A modelagem completa deste problema de identificação de forças dinâmicas pode ser visto em [7].

Segundo [8], o GMRES é um método iterativo utilizado para resolver sistemas lineares que tem a propriedade de minimizar a cada iteração a norma do vetor residual sobre o subespaço de Krylov. O algoritmo é derivado a partir do processo de Arnoldi por construção da base orthogonal L_2 do subespaços de Krylov. O GMRES pode ser considerado como uma generalização do algoritmo MINRES de Paige e Saunders ([6]) e é teoricamente equivalente ao método do Resíduo Conjugado Generalizado (GCR). Uma descrição mais detalhada do GMRES pode ser vista em [7].

A fim de avaliar o algoritmo EDMP foram realizados testes com diferentes quantidades de indivíduos na população. Foi adotado os seguintes parâmetros para ED e EDMP: fator de diferença $f = 0,7$ e probabilidade de cruzamento $CR = 0,9$. Nos testes com EDMP foram utilizados 80 processadores do cluster e o critério de parada é o número máximo de gerações da população e a verificação de sua estagnação. Em seguida, as soluções numéricas encontradas com ED e EDMP são comparadas com as soluções analíticas e as obtidas pelo GMRES no MATLAB[®].

As soluções analíticas do problema proposto, dada pela Eq. (4), foram calculadas no MATLAB. O custo computacional para obtê-las foi muito baixo uma vez que a média do tempo de execução é de aproximadamente 0,5 segundos.

O código computacional do algoritmo EDMP foi implementado de forma a permitir o operador escolher o número n de execuções do algoritmo. Embora a solução final seja a melhor solução encontrada entre as n execuções, também é apresentado como resposta a média e o desvio padrão do valor da função objetivo e da solução ótima nas n execuções do algoritmo. No problema proposto, EDMP foi executado 10 vezes.

As simulações com EDMP foram executadas no cluster do Laboratório de Computação Científica Aplicada e Tecnologia de Informação - LCCATI da FACIP/UFU que possui as seguintes características:

- Frontend: 16 núcleos de processamento, 64GB de memória RAM, 1 adaptador Infiniband QDR X4 (40 Gbps) e duas unidades RAID (3TB home e 12TB tmp1) que são exportadas para todo o cluster;
- Oito (08) nós computacionais, cada um com 64 núcleos de processamento, 128 GB de memória RAM, 2 adaptadores Infiniband QDR X4 (40 Gbps cada adaptador).

Performance de pico de 8.8 GFlops por núcleo (4.5 TFLOPS total + 0.25 TFLOPS no frontend disponíveis para pequenas tarefas);

- Um nó de serviço com 8 núcleos, 64GB de memória RAM, 1 adaptador Infiniband QDR X4 (40 Gbps);
- Um servidor adicional com 16 núcleos de processamento, 64GB de memória RAM, 1 adaptador Infiniband QDR X4 (40 Gbps) e uma unidade de storage com 28 TB (Por enquanto este servidor está inativo);
- Switch infiniband 36 portas QDR X4 (40 Gbps por porta) e um Switch ethernet 1Gb 48 portas.

Teste 1

No primeiro teste foram adotados $t \in [0; 0,01]$, 1600 indivíduos na população inicial o que equivale à 20 indivíduos por processador e, no máximo, 100 iterações para o critério de parada. Neste caso a matriz H tem 82×82 entradas. A Fig. 1 apresenta uma comparação entre a solução analítica e a numérica encontrada com ED, EDMP e GMRES. O valor do resíduo na solução ótima encontrada por EDMP foi $r = |X - HF| = 9,91 \times 10^{-5}$.

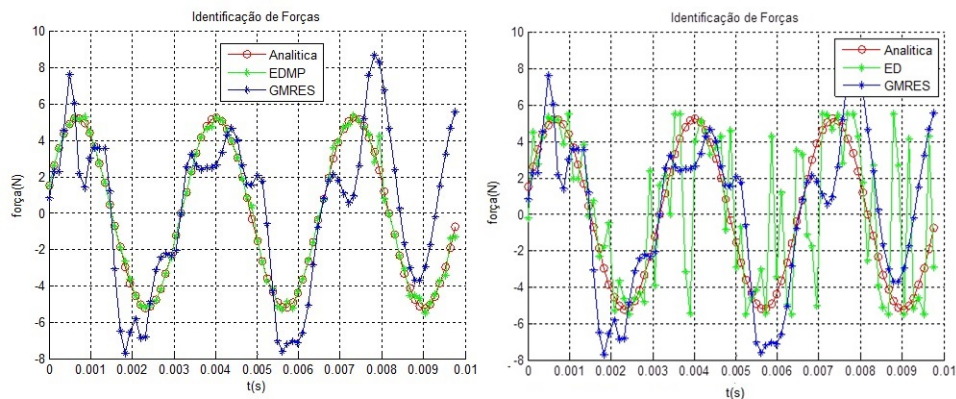


Figura 1: Solução analítica e numérica encontrada com ED, EDMP e GMRES - Teste 1.

O algoritmo da EDMP, após 6,95 minutos, convergiu para a solução ótima obtendo os seguintes valores para média e desvio padrão do valor da função objetivo na solução ótima entre as 10 execuções do algoritmo: média do resíduo $\bar{r}(X^*) = 0,0014$ e desvio padrão do resíduo $\sigma(r(X^*)) = 0,0021$. O algoritmo da ED foi interrompido depois de 15 minutos de execução sem convergir para a solução ótima.

Teste 2

Neste teste o objetivo é comparar as soluções entre ED, EDMP e GMRES em simulações onde o sistema linear é “grande”. Para tanto, foi utilizado $t \in [0; 0,03]$, 3200 indivíduos da população inicial, o que equivale à 40 indivíduos por processador e, no máximo, 100 iterações para o critério de parada. A matriz H tem 246×246 entradas. A Fig. 2 apresenta a comparação entre as soluções obtidas e a solução analítica. O valor do resíduo na solução ótima encontrada por EDMP foi $r = |X - HF| = 9,82 \times 10^{-5}$.

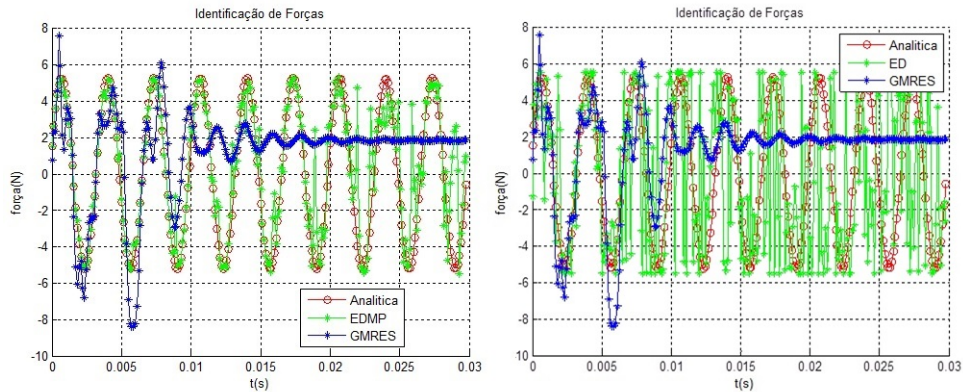


Figura 2: Solução analítica e numérica encontrada com ED, EDMP e GMRES - Teste 2.

Foram executadas 557.680 avaliações da função objetivo em 1,7 horas com EDMP. A média e desvio padrão do valor da função objetivo na solução ótima entre as 10 execuções do algoritmo da EDMP foram: média do resíduo $\bar{r}(X^*) = 0,0011$ e desvio padrão do resíduo $\sigma(r(X^*)) = 0,0018$. Novamente, depois de 16 horas de processamento, o algoritmo ED foi interrompido sem encontrar a solução ótima do problema, como pode ser visto na Fig. 2.

Este teste exemplifica uma das dificuldades dos algoritmos ED e GMRES que é obter a solução de “grandes” sistemas lineares. Enquanto o algoritmo da ED é muito lento e pode não encontrar a solução ótima do problema, o GMRES, a partir de determinado momento, começa a divergir.

A solução obtida por EDMP no teste 2 é muito mais precisa que a do GMRES e o valor do resíduo encontrado mostrou-se baixo. Uma desvantagem é o tempo de execução do algoritmo. Para esta simulação o GMRES levou 0,7 segundos para encontrar a solução do sistema porém, tal solução é bastante diferente da solução analítica. Embora o esforço computacional do EDMP seja grande, acarretando em muito tempo de execução, a solução por ele obtida está próxima da solução analítica.

5 Conclusões

Os resultados obtidos com EDMP foram muito satisfatórios já que o erro cometido é pequeno e aceitável. A maior vantagem de se usar EDMP é o fato de não se preocupar com condicionamento de matrizes e o da matriz não precisar ser quadrada. Além disso, o problema de identificação de forças dinâmicas foi resolvido sem usar nenhuma técnica de

regularização da matriz de coeficientes, visto que esta é uma matriz retangular formada por três blocos triangulares. Os testes realizados com EDMP permitem concluir que o método proposto se apresenta como uma alternativa para resolução de sistemas lineares de grande porte.

Visto que EDMP trabalha com avaliação da função objetivo, uma desvantagem do EDMP ao solucionar sistemas lineares de grande porte é o elevado tempo computacional. Além disso, visto que há um grande volume de dados para armazenar, a solução destes problemas precisa de um cluster com memória distribuída ou supercomputadores com grande memória compartilhada.

Embora a solução de sistemas lineares seja conceitualmente simples, na prática podem surgir muitos desafios. Acredita-se, devido as simulações realizadas nesta pesquisa, que o EDMP seja uma alternativa promissora, principalmente devido a praticidade de não se preocupar com as características da matriz dos coeficientes do sistema linear.

Referências

- [1] M. A. L. Brandão, Evolução diferencial melhorada implementada em processamento paralelo. Tese (doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Mecânica, (2014).
- [2] S-L. Cheng, C. Hwang, Optimal approximation of linear systems by a differential evolution algorithm, *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on , vol. 31, 698–707, (2001).
- [3] J. Duda, I. Skalna, Differential Evolution Applied to Large Scale Parametric Interval Linear Systems. *Lecture Notes in Computer Science: Large-Scale Scientific Computing*, vol. 7116, 206–213, (2012).
- [4] S. J. Leon, *Algebra Linear: com aplicações*. 8^aed, Rio de Janeiro: LTC, (2011).
- [5] S. Y. Liong, M. Atiquzzaman, Optimal Design ow Water Distribution Network using Shuffled Complex Evolution, *Journal of The Institution of Engineers*, vol. 44, 1, (2004).
- [6] C. C. Paige, M. A. Saunders, Solution of sparse indefinite systems of linear equations, *SIAM J. Numerical Analysis*, vol. 12, 617–629, (1975).
- [7] L. A. Purcina, *Técnicas de Otimização Evolutivas aplicadas à Solução de Grandes Sistemas Lineares*. Tese de doutorado. Universidade Federal de Uberlândia, (2010).
- [8] Y. Saad and M. H. Schultz, Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing*, vol. 7, 856–869, (1986).
- [9] R. M. Storn , K. V. Price, Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization Over Continuous Spaces, *International Computer Science Institute*, vol. 12, 1–16, (1995).