

Seleção Dinâmica da Dimensão do Subespaço de Krylov no Método GMRES(m) e suas Variantes

T.T. GONÇALEZ, Programa de Pós-Graduação em Matemática Aplicada, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

R.D. DA CUNHA¹, Instituto de Matemática, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

Resumo. Nesse trabalho apresentamos alguns algoritmos adaptativos do Método do Resíduo Mínimo Generalizado (GMRES) [10], um método iterativo para resolver sistemas de equações lineares com matrizes não simétricas e esparsas, o qual baseia-se nos métodos de projeção ortogonal sobre um subespaço de Krylov.

O GMRES apresenta uma versão reinicializada, denotada por GMRES(m), também proposta em [10], com o intuito de permitir a utilização do método para resolver sistemas de equações lineares cuja matriz dos coeficientes apresenta uma grande dimensão. No entanto, escolher um valor apropriado, m , para a dimensão da base do subespaço de Krylov é bastante difícil.

Dessa forma, nesse trabalho, acrescentamos ao GMRES(m) e algumas de suas variantes um critério que tem por objetivo escolher, ao longo das iterações, um m tal que se obtenha a convergência do método, possivelmente de forma mais rápida.

Aproximadamente duas centenas de testes foram realizados utilizando as matrizes da coleção Harwell-Boeing, que foram utilizados para mostrar o comportamento dos algoritmos adaptativos. Foram obtidos resultados muito bons conforme apresentaremos nesse trabalho.

1. Introdução

Nesse trabalho são estudados aprimoramentos ao método iterativo “Generalized Minimum Residual” (GMRES), proposto por Saad e Schultz em 1986 [10], e algumas de suas variantes, utilizados na solução de sistemas de equações lineares da forma

$$Ax = b, \tag{1.1}$$

onde $A \in \mathbb{R}^{n \times n}$ é não-simétrica e esparsa e $x, b \in \mathbb{R}^n$. No GMRES original, conforme proposto por Saad e Schultz, o resíduo inicial $r_0 = b - Ax_0$ é utilizado para gerar uma base ortonormal para um subespaço de Krylov,

$$\mathcal{K}_k(A, r_0) = \text{span}(r_0, Ar_0, \dots, A^{k-1}r_0), \tag{1.2}$$

¹rudnei.cunha@ufrgs.br

onde k é o número de iterações do método GMRES, através do processo de Arnoldi, o qual é uma modificação do processo de ortonormalização modificado de Gram-Schmidt. A solução do sistema acima é tomada como solução de um problema de mínimos quadrados que envolvem uma matriz e um vetor gerados durante o processo de Arnoldi. Dentre algumas características do GMRES, destaca-se a convergência em, no máximo, n iterações. Obviamente, para n muito grande isso não é uma vantagem, pois o processo de Arnoldi requererá o armazenamento de n vetores de dimensão n , além de uma matriz de dimensão $(n+1) \times n$ - o que levará a uma complexidade de ordem $O(n^2)$ - e exigirá n produtos matriz-vetor. Dessa forma em 1986, Saad e Schultz propuseram uma versão chamada *reinicializada*, GMRES(m), onde m é a dimensão da base do subespaço de Krylov, na qual se produz uma base ortonormal para o subespaço de Krylov de dimensão m , $\mathcal{K}_m(A, r_k)$; isto é, a cada iteração k , calculam-se o resíduo $r_k = b - Ax_k$ e os m vetores ortonormais através do processo de Arnoldi e, a partir deles resolve-se um problema de mínimos quadrados, obtendo-se uma aproximação x_{k+1} para a solução do problema. Infelizmente, ao se optar pelo método GMRES(m), perde-se a garantia de convergência apresentada pelo método GMRES, a menos que m seja suficientemente grande. Em [10] foi proposto o seguinte teorema para a escolha de m :

Teorema 1.1. *Suponha que A é diagonalizável. Assuma que existem ν autovalores $\lambda_1, \lambda_2, \dots, \lambda_\nu$ de A com parte real não positiva e tome os outros autovalores sendo fechados em um círculo centrado em C , com $C > 0$ e tendo raio R com $C > R$. Então, o GMRES(m) converge para algum valor inicial de x_0 se*

$$m > \nu \frac{\log\left[\frac{DC}{dR} \kappa(X)^{\frac{1}{\nu}}\right]}{\log\frac{C}{R}}, \quad (1.3)$$

onde

$$D = \max_{i=1, \nu; j=\nu+1, n} |\lambda_i - \lambda_j|, \quad d = \min_{i=1, \nu} |\lambda_i|.$$

Demonstração. Ver [10]. □

Podemos perceber que a determinação desse valor mínimo m depende da determinação da distribuição dos autovalores de A e assim é usado somente em casos especiais, não podendo ser tomada como padrão na escolha do m .

Para comprovar que nem todas as matrizes satisfazem tal teorema, geramos uma matriz utilizando o programa MATLAB 5.3, a qual satisfaz as condições do teorema apresentado acima. Apresentamos os gráficos da estrutura e dos autovalores de tal matriz, na Figura 1. Para $n = 100$ obtivemos os seguintes valores para os parâmetros do teorema: $D = 192,5682$; $d = 54,0479$; $C = 3,0004$; $R = 1,4140$; $\nu = 17$, tal que $m > 84,8287$. Utilizando a mesma matriz no método GMRES(m) obtivemos $m = 20$. Observe que o valor de m proposto pelo teorema é muito superior ao necessário para a resolução do sistema através do GMRES(m). Isso mostra que tal teorema não é adequado e, particularmente para as matrizes utilizadas nos testes que serão apresentados nesse trabalho, a distribuição dos autovalores das mesmas não satisfaz a hipótese do teorema. Em [5] são apresentados gráficos da

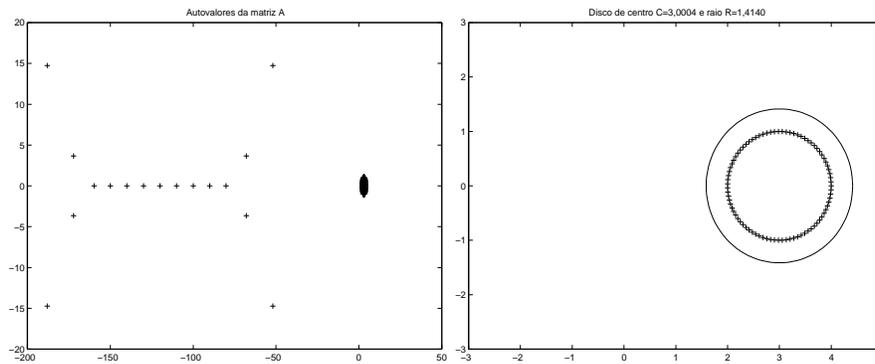


Figura 1: a) Autovalores da matriz, b) Detalhe ampliado.

distribuição de autovalores de várias matrizes armazenadas em [8], e para nenhuma delas o teorema é satisfeito. Em vista disso, tal teorema torna-se muito limitado, tornando-se necessário uma estratégia mais abrangente e que não apresente tantas limitações quanto às características da matriz.

Outros resultados referentes à convergência do GMRES foram propostas por Joubert [7] e Embree [4]. Joubert propõe um algoritmo adaptativo para o GMRES(m), no sentido de efetuar a reinicialização apenas quando a convergência não é considerada adequada, de acordo com um critério envolvendo a taxa de convergência ([7, p. 440]), porém sem alterar o valor de m ao longo das iterações. Já os resultados propostos por Embree utilizam quantidades como o número de condição e “field of values”, cujo cálculo aumentaria consideravelmente o custo computacional de se resolver um sistema linear; porém, esses resultados não apresentam um procedimento para a seleção de um valor mais adequado para m . Dessa forma, concluímos que nem o teorema proposto em [10] nem as condições propostas em [4] são adequados para a escolha e/ou correção do valor de m .

Nesse trabalho apresentamos estratégias para determinar dinamicamente, ao longo do processo de solução do sistema $Ax = b$, um valor adequado para m , de forma a garantir a convergência do método GMRES(m) e de algumas de suas variantes, bem como obter a solução com o menor gasto computacional possível, em termos de número de operações de ponto-flutuante.

2. Descrição dos Métodos

Os experimentos foram realizados utilizando-se o método GMRES(m); uma variante do GMRES(m) utilizando transformações de Householder, ao invés do processo de Arnoldi, proposta por Walker [12], a qual chamaremos de GMRESH(m); o “Simpler GMRES” e sua variante usando transformações de Householder [13], os quais chamaremos de SGMRES(m) e SGMRESH(m); o “LooseGMRES”, proposto Baker et al. [1], chamados de LGMRES(m) e LGMRESH(m); o GMRES adaptativo

proposto em [3], chamado de A-GMRES(m); e o BC-GMRES(m), proposto em [9]. Apresentamos agora uma breve explicação das variantes do GMRES, com suas principais características e vantagens.

O GMRESH(m) apresenta melhores propriedades numéricas que o GMRES(m), especialmente nas iterações finais, quando os resíduos tornam-se cada vez mais próximos, e utiliza menos armazenamento, apesar de requerer duas vezes mais operações aritméticas do que o GMRES(m). Segundo [12], em GMRESH(m) os vetores de Householder têm dimensão menor e, junto com a matriz triangular superior gerada pelo algoritmo, podem ser armazenados em uma matriz de ordem $n \times (\text{maxit} + 1)$, onde maxit é o número máximo de iterações. Isto é um pouco menos armazenamento do que o requerido pelo GMRES(m), o qual armazena toda a base de vetores ortonormal, bem como as matrizes triangulares superiores. Assim, como em geral temos que $\text{maxit} \ll n$, então essa seria a maior vantagem entre esses dois algoritmos.

O Simpler GMRES surgiu em 1994 com Homer F. Walker e Lu Zhou [13], sendo apresentado com uma modificação no processo de Arnoldi original do GMRES. Segundo [13], na implementação usual do GMRES, o processo de Arnoldi é aplicado com $v_1 = \frac{r_0}{\|r_0\|_2}$. A ortogonalização em cada passo do algoritmo pode ser feita usando o processo modificado de Gram-Schmidt ou transformações de Householder. Note que o SGMRES é resultante apenas da modificação feita no v_1 , tornando assim o problema de mínimos quadrados em um problema de mínimos quadrados triangular superior, eliminando a fase de triangularização da matriz de Hessenberg.

O Loose GMRES surgiu em 2003 com Baker, Jessup e Manteuffel; veja [1]; sendo apresentado com uma técnica para acelerar a convergência do GMRES. O LGMRES(m, k) é muito semelhante ao método do gradiente conjugado completo (CG) com condicionador polinomial, e sua implementação não exige muitas modificações no GMRES(m). Segundo [1], uma motivação para propor o LGMRES se deu pelo fato de que o GMRES(m) não mantém a ortogonalidade entre espaços aproximados generalizados em sucessivas reinicializações, tornando assim sua convergência vagarosa ou até mesmo levando a divergência. Também é sabido, conforme [2], que métodos aumentados formam uma classe de técnicas de aceleração; esses métodos buscam evitar a divergência melhorando assim as informações obtidas com o GMRES em cada reinicialização. Tipicamente, um subespaço A-invariante é agrupado ao subespaço aproximado de Krylov, resultando assim em um subespaço de Krylov aumentado. O subespaço A-invariante associado com os menores autovalores é comumente usado. A idéia então consiste de adicionar vetores ao subespaço aproximado de Krylov que representaria o espaço aproximado para cada ciclo de reinicialização. O LGMRES(m, k) tipicamente não requer mais iterações do que o GMRES(m), e o fato do mesmo ser um acelerador de uma possível convergência do GMRES(m) não significa que o mesmo não possa ser aplicado sem o uso de pré-condicionadores.

O método GMRES-Adaptativo foi proposto por Maria Sosonkina Driver em 1997 [3], e o denotamos por A-GMRES. A diferença básica em relação ao GMRESH(m) consiste no fato de que o A-GMRES apresenta, se necessário, alguns testes para incrementar o m e então reinicializar o método. De acordo com [3] a essência do A-GMRES é obter um parâmetro m adequado para a reinicialização do problema.

O BC-GMRES foi proposto por Takashi Nodera e Kentaro Moriya em 2003 [9], e é um método que escolhe o ciclo de reinicialização do m baseado no teste da norma residual e na distribuição dos zeros do polinômio residual do GMRES(m). Essa estratégia foi baseada em algumas conclusões apresentadas em [11]. O BC-GMRES executa a reinicialização quando a distribuição dos zeros do polinômio residual do GMRES(m) tornam-se ideais, ou seja, pode-se dizer que quanto mais espalhados estão os zeros do polinômio residual mais provável que a distribuição dos zeros aproximados seja similar aos zeros ideais. No BC-GMRES a decisão da reinicialização é tomada através da análise de duas condições, a saber: a distribuição dos zeros aproximados e o teste de convergência da norma residual.

3. Critério de Escolha do Valor de m

Apresentamos agora o critério desenvolvido para a escolha de m adequado, sendo que os objetivos para a elaboração do tal critério foram a redução da norma residual e a variação de m , ou seja, buscamos um procedimento que possa determinar o aumento ou a redução de m . Inicialmente tentamos desenvolver um critério que apresentasse resultados satisfatórios quanto ao número de convergência dos métodos adaptativos e ao número de não convergências.

Foram elaborados cinco critérios ao longo de todo o trabalho, sendo que os três primeiros utilizavam o número de dígitos significativos exatos (DIGSE) como ferramenta para a comparação da variação do m . A idéia aqui era determinar a quase estagnação da convergência através das normas de dois resíduos sucessivos. Conforme pode-se observar em [5], os critérios elaborados com o DIGSE não apresentaram resultados satisfatórios, visto que o m praticamente não apresentou variação. Optou-se então por alterar o valor de m de acordo com o logaritmo (base 10) das normas dos resíduos, e estabelecendo faixas de variação dentro das quais alteraríamos o valor de m de forma adequada. Por exemplo, se $\log \|r\|_2 > 0$, então o valor de m é duplicado (até um certo limite pré-especificado m_{max} , o qual pode ser escolhido em função de fatores como a memória disponível, por exemplo). Se $\log \|r_i\|_2 < 0$ e $\log \|r_i\|_2 > \frac{2 \log \epsilon}{3}$, então o processo já está convergindo, e cabe agora verificar se está convergindo lentamente, caso em que adiciona-se a m o valor de m_{orig} (o valor inicialmente estipulado para m); caso contrário, podemos reduzir o valor de m , subtraindo dele $\frac{m_{orig}}{3}$ (desde que m nunca seja menor do que m_{orig}). Se $\log \|r_i\|_2 < 0$ e $\log \|r_i\|_2 < \frac{2 \log \epsilon}{3}$, então o método já está quase que alcançando a tolerância especificada, e pode-se ainda alterar o valor de m , porém com quantidades menores, pois pode ser dispendioso aumentar por demais m , já que a convergência quase foi alcançada. Por isso, quando a norma do resíduo encontra-se nessa faixa, porém a taxa de convergência é lenta, então adiciona-se a m o valor de $\frac{m_{orig}}{2}$; caso contrário, reduzimos o valor de m , subtraindo dele $\frac{m_{orig}}{4}$.

Em nosso critério, utilizamos como fator de determinação da taxa de convergência do método a relação $\frac{\|r_{i-5}\|_2}{\|r_i\|_2}$ e dizemos que ela é “rápida” se tal relação for superior a 2. Além disso, a fim de evitar muitas mudanças no valor de m , optamos por alterá-lo, se necessário, a cada cinco iterações. Tal idéia é baseada no algoritmo adaptativo para escolha do parâmetro de relaxação do método SOR conforme pro-

posto por Hageman e Young em 1981 [6]. O critério de escolha do valor de m pode ser apresentado de forma algorítmica como segue:

Algoritmo 3.1. *Critério de escolha de m :*

1. Se $\text{mod}(i, 5) = 0$
2. Se $\log \|r\|_2 > 0$
3. $m = \min(2m, m_{max})$
4. senão, se $\log \|r\|_2 \leq 0$ e $\log \|r\|_2 > \frac{2 \log tol}{3}$
5. Se $i > 5$ e $\frac{\|r_{i-5}\|}{\|r_i\|} > 2$
6. $m = \max(m_{orig}, m - \frac{m_{orig}}{3})$
7. senão
8. $m = \min(m + m_{orig}, m_{max})$
9. fim se
10. senão
11. Se $i > 5$ e $\frac{\|r_{i-5}\|}{\|r_i\|} > 2$
12. $m = \max(m_{orig}, m - \frac{m_{orig}}{4})$
13. senão
14. $m = \min(m + \frac{m_{orig}}{2}, m_{max})$
15. fim se
16. fim se
17. $\|r_{i-5}\|_2 = \|r\|_2$
18. fim se

Esse critério foi embutido em todas as variantes do GMRES apresentadas na Seção 2, a cujas abreviações foi adicionada a letra **A** para indicar o uso do critério. Os métodos foram comparados como segue: SGMRES-A(m) com SGMRES(m); SGMRESH-A(m) com SGMRESH(m); BC-GMRES(m) com BC-GMRES-A(m); LGMRES-A(m, k) com LGMRES(m, k) e, por fim, o A-GMRES(m) com o método GMRESH(m), já que estes dois utilizam a formulação do GMRES(m) com base nas transformações de Householder.

Os métodos utilizados para a realização dos testes foram empregados sem pré-condicionadores e utilizou-se como critério de parada $\|r_i\|_2 < \epsilon$, sendo $\epsilon = 10^{-10}$. Os valores m_{orig} e m_{max} foram estipulados como 10 e 50, respectivamente (o que representa um intervalo de valores típicos para m encontrados na literatura).

Os testes foram realizados utilizando o software MATLAB versões 5.3 e 6.2 em um computador PC com processador AMD Athlon XP 2.400 e 512 MBytes de memória RAM. Ao todo foram realizados 114 testes, com as seguintes matrizes: ADD32, ARC_130, CAVITY05, CDDE1, E05R0000, E05R0300, FIDAPM05, FS_183_1, FS_760_1, GRE115, IMPCOL_A, JPWH991, NOS3, ORSIRR1, PDE900, SAYLR3, SHERMAN1 e STEAM1, presentes na coleção “Matrix Market” [8]. Para fins de explanação, apresentamos os gráficos comparativos da evolução da convergência do método adaptativo com o seu equivalente não-adaptativo, para as matrizes CDDE1 e E05R0300; além disso, num outro gráfico apresentamos a variação do valor de m juntamente com a curva de $\log \|r_i\|_2$. Salientamos que o eixo horizontal, o qual representa as iterações, apresenta escalas diferentes em cada

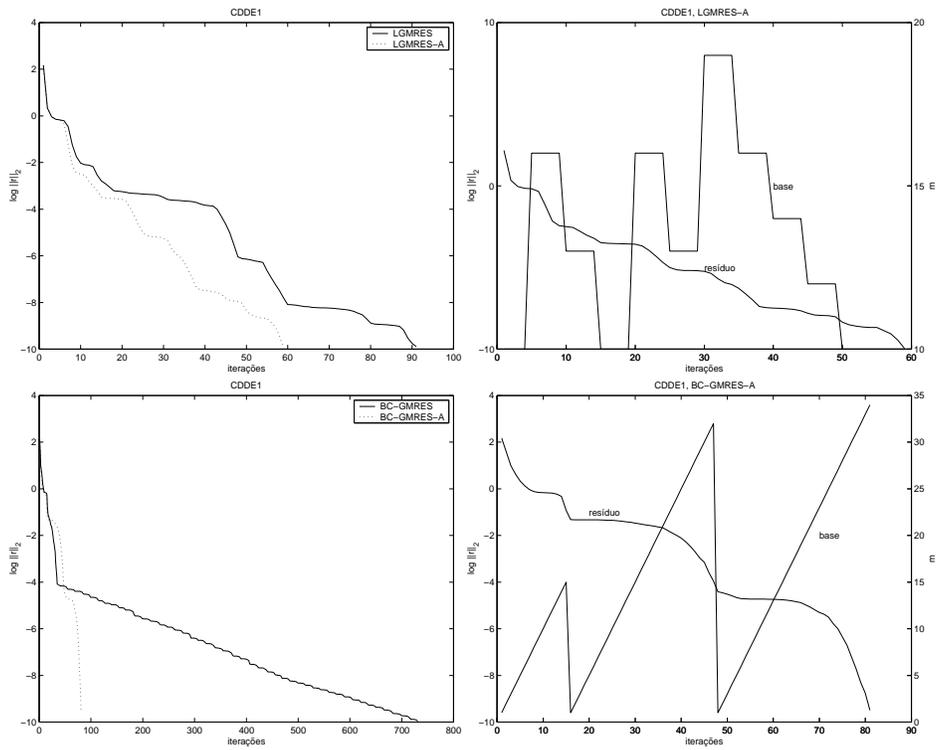


Figura 2: Curvas residuais para o problema CDDE1: acima, usando LGMRES($m, 2$) e variação de m e a curva residual para o método adaptativo LGMRES-A($m, 2$); abaixo, usando BC-GMRES(m) e variação de m e a curva residual para o método adaptativo BC-GMRES-A(m).

método. As Figuras 2 e 3 ilustram o comportamento dos métodos usando o critério adaptativo.

Observe que o método adaptativo obteve a convergência em menos iterações e como, sempre que m aumenta, há uma correspondente diminuição no valor do resíduo. Além disso, na Figura 2 observa-se uma rápida diminuição no valor da norma do resíduo, para a variante adaptativa do método BC-GMRES(m) (cabe ressaltar que o perfil de dente-de-serra para a dimensão da base é típica do método BC-GMRES, já que ao realizar uma reinicialização no GMRES(m) a base inicia em 1). Na Figura 3, referente ao problema E05R0300, observa-se um comportamento atípico, dentre todos os testes realizados, pois não se obteve convergência com qualquer dos métodos, nem tampouco foi o critério adaptativo capaz de garantir convergência.

A Tabela 1 apresenta os resultados obtidos com todos os métodos estudados, para esse problema, onde pode-se observar que as versões adaptativas foram sempre melhores do que as não-adaptativas, garantindo a convergência e com menor custo

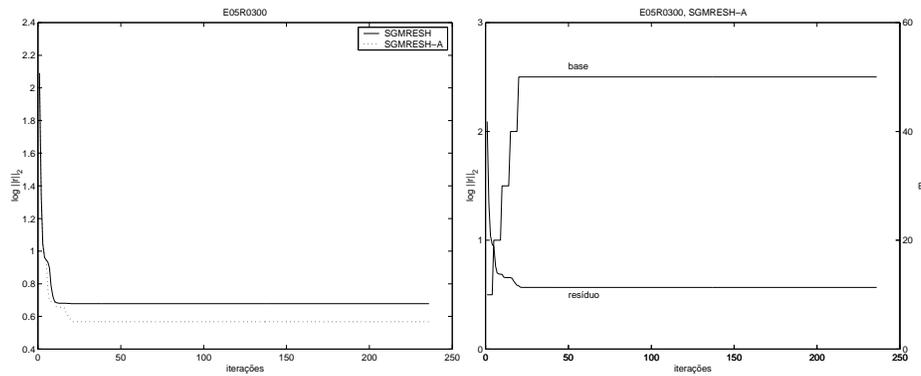


Figura 3: Curvas residuais para o problema E05R0300, usando SGMRESH(m) e variação de m e a curva residual para o método adaptativo SGMRESH-A(m).

computacional.

Podemos resumir os resultados dos experimentos, tabelados em [5, pp.63-163], da seguinte forma: 76 testes convergiram para os métodos adaptativos, ou seja, para os métodos acrescidos do critério desenvolvido nesse trabalho para a escolha do m adequado ao sistema estudado, apresentando menor número de flops e de iterações do que os demais; 24 métodos não adaptativos, ou seja, métodos que não apresentam nenhum critério para a escolha do m , convergiram apresentando menor número de flops do que as versões adaptativas e 12 testes não apresentaram convergência dentro da tolerância pré-especificada, em nenhum dos casos.

Dentre os valores apresentados devemos destacar que, para algumas matrizes, as versões adaptativas mantiveram-se absolutas, ou seja, todos os métodos adaptativos convergiram e apresentaram menor número de flops do que os demais, para as matrizes CDDE1, GRE115, NOS3, ORSIRR1, SAYLR3 e SHERMAN1. Em outros, pode-se notar que apenas as versões adaptativas obtiveram a convergência dentro da tolerância pré-especificada, como para FS_183_1; GRE115; LNS_131 e STEAM1, o que indica que o uso das versões adaptativas é indicado. Podemos citar ainda que entre 114 testes realizados, 5 apresentaram convergência apenas para as versões adaptativas, nenhuma apenas para as versões não-adaptativas, e dentre todos os testes, somente 2 sistemas não obtiveram a convergência em nenhum dos casos. Salientamos que, para os sistemas utilizando as matrizes E05R0300 e a IMPCOL_A, não se obteve convergência em nenhuma das versões; e, nos testes realizados com a matriz ARC_130, alguns métodos apresentaram o mesmo número de flops para as versões adaptativas e não adaptativas, fazendo com que não pudéssemos concluir nada a respeito das diferenças apresentadas por essas versões.

Um fato interessante foi que o número de iterações necessárias para obter a convergência das versões adaptativas sempre foram menores do que as demais; isso, independentemente, do número de flops apresentados pelos métodos.

Método	i	m mínimo	m máximo	Flops
A-GMRES	400	3	10	$5,0780 \times 10^8$
GMRESH-A	30	2	20	$9,5554 \times 10^7$
SGMRES	1000	7	10	$3,5790 \times 10^8$
SGMRES-A	30	10	40	$4,2707 \times 10^7$
SGMRESH	400	2	10	$3,2589 \times 10^8$
SGMRESH-A	25	10	20	$4,4876 \times 10^7$
BC-GMRES	1000	10	10	$1,8685 \times 10^8$
BC-GMRES-A	70	10	60	$3,8236 \times 10^7$
LGMRES($m,2$)	35	10	10	$1,6624 \times 10^7$
LGMRES-A($m,2$)	14	10	20	$1,4284 \times 10^7$
LGMRES($m,3$)	45	10	10	$3,9552 \times 10^7$
LGMRES-A($m,3$)	40	10	20	$2,3118 \times 10^7$

Tabela 1: Resultados dos métodos para o problema CDDE1.

4. Conclusão

Em virtude dos dados apresentados podemos concluir que as versões adaptativas formam uma excelente alternativa para a resolução de sistemas não-simétricos esparsos, visto que tendem a diminuir o custo computacional e o tempo de execução dos mesmos.

Cabe ressaltar que os testes foram realizados sem o uso de pré-condicionadores, visto que o nosso objetivo nesse trabalho era de elaborar um critério adequado para a escolha da dimensão do subespaço de Krylov, em relação ao método GMRES em si. Não obstante, testes realizados posteriormente demonstraram que pré-condicionadores (como fatorações incompletas, polinomiais e outros) podem ser usados em combinação com o critério aqui proposto, com resultados semelhantes aos aqui relatados.

Agradecimentos

Os autores agradecem aos revisores pelas observações feitas, as quais contribuíram para a melhoria deste artigo.

Referências

- [1] A. Baker, E. Jessup, T. Manteuffel, A Technique for Accelerating the Convergence of Restarted GMRES. Technical Report CU-CS-945-03, Dept. of Computer Science, College of Engineering and Applied Science, University of Colorado at Boulder, 2003.
- [2] A. Chapman, Y. Saad, Deflated and augmented Krylov subspace techniques. *Numerical Linear Algebra with Applications*, **4** (1997), 15–41.

- [3] M.S. Driver, *Parallel sparse linear algebra for homotopy methods*. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, 1997.
- [4] E. Embree, How descriptive are GMRES convergence bounds? Research Report NA-99/08, Numerical Analysis Group, Oxford University Computing Laboratory, 1999.
- [5] T.T. Gonçalez, Algoritmos adaptativos para o método GMRES(m). Dissertação de mestrado, Programa de Pós-Graduação em Matemática Aplicada, UFRGS, 2005.
- [6] L.A. Hageman, D.M. Young, “Applied Iterative Methods”, Academic Press, New York, 1981.
- [7] W. Joubert, On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems. *Numerical Linear Algebra with Applications*, **1** No. 5 (1994), 427–447.
- [8] Mathematical and Computational Sciences Division, Information Technology Laboratory, National Institute of Standards and Technology, Matrix Market: A visual repository of test data for use in comparative studies of algorithms for numerical linear algebra. <http://math.nist.gov/MatrixMarket/>, 2003.
- [9] K. Moriya, T. Nodera, New adaptive GMRES(m) method with choosing suitable restart cycle m . In R. Wyrzykowski et al., editor, *Parallel Processing and Applied Mathematics 2003*, number 3019 in Lecture Notes in Computer Science, pp 1105–1113, Berlin, 2003. Springer-Verlag.
- [10] Y. Saad, M.H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal of Scientific and Statistical Computing*, **7** (1986), 856–869.
- [11] N. Tsuno, T. Nodera, The speedup of the GMRES(m) method using the early restarting procedure. *Journal of the Information Processing Society of Japan*, **40**, No. 4 (1999), 1760–1773.
- [12] H.F. Walker, Implementation of the GMRES method using Householder transformations. *SIAM Journal on Scientific and Statistical Computing*, **9**, No. 1 (1989), 152–163.
- [13] H.F. Walker, L. Zhou, A simpler GMRES. *Linear Algebra and its Applications*, **1** (1994), 574–581.