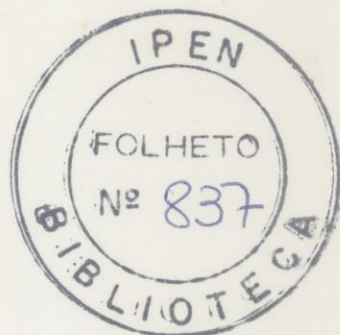


Conferência Nacional



SOBRE A SOLUÇÃO ITERATIVA DE SISTEMAS DE EQUAÇÕES LINEARES EM ENGENHARIA ESTRUTURAL

Paulo de Mattos Pimenta
Doktor-Ingenieur, Universität Stuttgart
Prof. Dr., Escola Politécnica da USP
Engenheiro, IPEN/CNEN-SP

Miguel Mattar Neto
Mestre em Engenharia Nuclear, IPEN/CNEN-SP
Engenheiro, IPEN/CNEN-SP

SUMÁRIO

A solução iterativa de sistemas de equações lineares tem ganho novamente interesse nos últimos anos. Isto justifica-se pela necessidade de tornar análises não lineares mais econômicas, nas quais a solução de vários sistemas é necessária, e pela oportunidade de se utilizar novos recursos computacionais como supercomputadores com cálculo paralelo.

Neste trabalho são apresentados e comparados dois métodos apropriados para sistemas com matrizes simétricas e positivamente definidas: o método dos gradientes conjugados e o método quase-Newton denominado BFGS.

A seguir são apresentadas várias formas de se pré-condicionar o sistema com o objetivo de se acelerar a convergência dos métodos.

O trabalho é concluído por alguns experimentos numéricos.

1. INTRODUÇÃO

Hoje pode-se dizer que quase a totalidade dos programas computacionais de cálculo estrutural utilizam alguma forma do método da eliminação na solução de sistemas de equações lineares. Apesar de toda a pesquisa desenvolvida na implantação eficiente deste método em programas de grande porte os resultados ainda não são satisfatórios. A análise de problemas grandes (tridimensionais por exemplo) e a análise não linear de estruturas (onde um número grande de soluções é requerida) são na maioria das vezes excessivamente caras, embora os computadores tenham se tornado cada vez maiores e mais rápidos. Pode-se afirmar que o método da eliminação tem a sua eficiência prejudicada por duas razões: (i) no decorrer da eliminação a matriz de rigidez, que é originalmente bastante esparsa (grande número de elementos nulos), fica preenchida, requerendo grande capacidade de memória, mesmo considerando a característica dela ser de banda; (ii) a ordem das operações não facilita a utilização de processamento em paralelo, hoje à disposição nos supercomputadores como o CRAY.

Assim justifica-se a pesquisa de algoritmos de solução de sistemas de equações lineares que possam ser mais eficientes que os tradicionais atualmente utilizados em cálculo estrutural. Neste trabalho são apresentados dois métodos iterativos que surgiram no contexto da otimização e que se prestam para cálculo estrutural, pois foram desenvolvidos para matrizes simétricas e positivas definidas. São eles o método dos gradientes conjugados e o método quase-Newton denominado BFGS.

Métodos iterativos têm o inconveniente de requererem frequentemente um número excessivo de iterações. No entanto se o sistema for pré-condicionado, como indicado neste trabalho, o número necessário de iterações até a convergência pode ser bastante reduzido. Além disso os métodos aqui apresentados prestam-se bem para a paralelização do processamento, pois várias operações são efetuadas de forma vetorial. E finalmente, lembrando que a tendência atual dos computadores é de aumentar a velocidade de cálculo em relação a velocidade de acesso a memórias auxiliares, pode-se afirmar que métodos iterativos que consumam bem menos memória central podem compensar um eventual maior número de operações com um menor número de operações de entrada e saída.

Este trabalho procura comparar os dois métodos acima citados como também algumas formas de se pré-condicionar o sistema de equações. Por isso são apresentados alguns resultados numéricos preliminares que servem para ilustrar algumas das afirmações colocadas no texto.

Concluindo esta introdução apresenta-se a notação empregada neste trabalho. O sistema de equações é representado por

$$A x = b \quad (1)$$

onde A é a matriz de rigidez, simétrica e positiva definida, x é o vetor dos deslocamentos nodais e b o vetor dos carregamentos nodais. Para matrizes serão portanto utilizadas letras latinas maiúsculas, para vetores letras latinas minúsculas e para escalares letras gregas minúsculas. A transposição de uma matriz ou vetor será representada por um asterisco.

2. O MÉTODO DOS GRADIENTE CONJUGADOS

Este algoritmo é uma generalização do método clássico dos gradientes conjugados apresentado em [1]. Nele o sistema (1) é pré-condicionado por uma matriz B , que é uma aproximação de A , mas de fatoração ou inversão mais fácil. Introduzindo-se o resíduo r na iteração j como

$$r_j = b - Ax_j \quad (2)$$

o algoritmo é apresentado esquematicamente a seguir

Passo	Operações
1	$j=0; r_0 = b - Ax_0; p_0 = z_0 = B^{-1} r_0$
2	$\alpha_j = r_j^* z_j / p_j^* A p_j$
3	$x_{j+1} = x_j + \alpha_j p_j$
4	$r_{j+1} = r_j - \alpha_j A p_j$
5	se $\ r_{j+1}\ < \delta$ convergência é atingida
6	$z_{j+1} = B^{-1} r_{j+1}$
7	$\beta_j = r_{j+1}^* z_{j+1} / r_j^* z_j$
8	$p_{j+1} = z_{j+1} + \beta_j p_j$
9	$j = j+1$, retorna ao passo 2

Sobre o algoritmo acima pode-se fazer as seguintes afirmações:

- (i) Se $B = I$ (matriz identidade) obtém-se o método clássico
- (ii) Se $B = A$ o método converge imediatamente
- (iii) O método requer o armazenamento de apenas cinco vetores
- (iv) Aplicações deste método podem ser vistas em [2], [3] e [4]
- (v) O pré-condicionamento com B equivale a resolver o sistema.

$$\tilde{A} \tilde{x} = \tilde{b} \quad (3)$$

sendo

$$\tilde{A} = C^{-1} A C^{-1*} \quad (3a)$$

$$\tilde{x} = C^* x \quad (3b)$$

$$\tilde{b} = C^{-1} b \quad (3c)$$

$$B = C C^* \quad (3d)$$

com o método clássico

- (vi) x_0 é uma estimativa inicial da solução. Normalmente adota-se $x_0 = 0$. Em problemas não lineares pode ser adotada a solução do incremento anterior.
- (vii) δ é adotado normalmente igual a $10^{-5} \|b\|$, onde as barras duplas indicam a norma euclídeana do vetor.
- (viii) pode-se demonstrar que o método converge em número finito de iterações menor ou igual a ordem do sistema, se não houver erros de arredondamento.

3. MÉTODO QUASE-NEWTON TIPO BFGS

O método BFGS de atualização foi introduzido em [5], [6], [7] e [8] no contexto da otimização. Um bom texto sobre métodos quase-Newton pode ser encontrado em [9]. A primeira aplicação em problemas estruturais pode ser encontrada em [10]. Em [11] pode ser vista uma aplicação em problemas acoplados complexos. Em [5] tem-se a primeira aplicação na solução de sistemas de equações lineares.

O algoritmo com pré-condicionamento através da matriz B é apresentado esquematicamente a seguir.

Passo	Operações
1	$j=0; r_0 = b - Ax_0, p = B^{-1} r_0$
2	$v_i = w_i = 0 \quad (i = 1, \dots, n)$
3	$\alpha = p^* r_j / p^* A p$
4	$x_{j+1} = x_j + \alpha p$
5	$r_{j+1} = r_j - \alpha A p$
6	se $\ r_{j+1}\ < \delta$ convergência é atingida
7	$v_{i-1} = v_i; w_{i-1} = w_i \quad (i=2, \dots, n)$
8	$v_n = p / p^* r_j$ $w_n = r_{j+1} - (1 - \sqrt{\alpha}) r_j$
9	$p = r_{j+1}$ $p = p + (v_i^* p) w_i \quad (i = n, \dots, 1)$ $p = B^{-1} p$ $p = p + (w_i^* p) v_i \quad (i = 1, \dots, n)$
10	$j = j + 1$ retorna ao passo 3

Algumas observações fazem-se aqui necessárias.

- (i) n é o número máximo de atualizações tipo BFGS
- (ii) α é um parâmetro que é obtido no passo 3 pela minimização da energia potencial na direção do vetor p.
- (iii) o método requer que $4 + 2n$ vetores sejam armazenados, o que representa uma desvantagem em relação ao método dos gradientes conjugados.
- (iv) se $B = I$, obtém-se o método de Jacobi acelerado por atualizações do tipo BFGS; com $n=0$ obtém-se o método de Jacobi.
- (v) as observações (vi) e (vii) da seção anterior são aqui também válidas.

4. PRÉ-CONDICIONAMENTO

4.1 Introdução

A matriz B que aproxima A nos dois métodos iterativos apresentados deve ser uma matriz fácil e eficientemente fatorável ou inversível. Para a descrição das possibilidades de criação de B, é necessária a introdução da seguinte notação. Seja X uma matriz qualquer, simétrica e positiva definida. X pode ser decomposta em fatores de Crout conforme

$$X = U_{CR}^*(X) D_{CR}(X) U_{CR}(X) \quad (4a)$$

onde U_{CR} é uma matriz superiormente triangular com elementos unitários na diagonal e D_{CR} é uma matriz diagonal. X admite ainda uma decomposição em fatores de Cholesky de acordo com

$$X = U_{CH}^*(X) U_{CH}(X) \quad (4b)$$

sendo que

$$U_{CH} = D_{CR}^{1/2} U_{CR} \quad (4c)$$

Em analogia com a decomposição em fatores pode-se definir decomposições em parcelas de Gauss-Seidel segundo

$$X = U_{GS}^*(X) + D_{GS}(X) + U_{GS}(X) \quad (5a)$$

ou de Trujillo na seguinte forma

$$X = U_{TR}^*(X) + U_{TR}(X) \quad (5b)$$

com

$$U_{TR} = U_{GS} + \frac{1}{2} D_{GS} \quad (5c)$$

U_{GS} é uma matriz superiormente triangular com elementos nulos na diagonal. $D_{GS}(X)$ é a diagonal de X.

4.2 Pré-condicionamento de Marchuk e de Hughes

Para apresentar as possibilidades de pré-condicionamento segundo Marchuk e Hughes define-se a matriz

$$\bar{A} = D_{GS}^{-1/2}(A) \left[A - D_{GS}(A) \right] D_{GS}^{-1/2}(A) \quad (6)$$

e portanto tem-se a igualdade abaixo.

$$A = D_{GS}^{1/2}(A) \left[I + \bar{A} \right] D_{GS}^{1/2}(A) \quad (7)$$

Em problemas estruturais A tem normalmente elementos diagonais maiores em valor absoluto que os demais de modo que \bar{A} tem elementos nulos na diagonal e elementos inferiores a unidade nas demais posições.

Suponha-se agora que \bar{A} possa ser decomposto em parcelas conforme

$$\bar{A} = \sum_{i=1}^n \bar{A}_i \quad (8)$$

\bar{A}_i podem ser por exemplo as contribuições de cada elemento ou de cada subestrutura para \bar{A} . Mas \bar{A}_i podem ser também quaisquer outras parcelas de \bar{A} , como por exemplo as dadas por (5a) ou (5b), ou mesmo linhas e colunas de \bar{A} .

Uma forma de se fatorar aproximadamente A através de (7) é adotar a aproximação

$$\bar{A} \approx I + \bar{A} \quad (9)$$

dada por

$$\bar{A} = \prod_{i=1}^n (I + \bar{A}_i) \quad (10)$$

Desenvolvendo-se o produtório em (10) chega-se a

$$\bar{A} = I + \bar{A} + \dots \quad (11)$$

onde os termos não indicados são resultantes dos produtos dos elementos de \bar{A}_i . Se estes elementos forem inferiores a unidade

a aproximação (10) fica justificada. Utilizando (5b) por exemplo chega-se a

$$\tilde{A} = \left[I + U_{TR}^*(A) \right] \left[I + U_{TR}(\tilde{A}) \right] \quad (12)$$

\tilde{A} dada por (12) é simétrica, positiva definida e requer a mesma memória que a esparsa matriz A.

Marchuk em [12] introduziu no lugar de (10) a seguinte fatoração aproximada,

$$\tilde{A} = \prod_{i=1}^n \left(I + \frac{1}{2} \bar{A}_i \right) \prod_{i=n}^1 \left(I + \frac{1}{2} A_i \right) \quad (13)$$

que garante que \tilde{A} seja simétrica e positiva definida.

Hughes em 13 introduziu as seguintes aproximações, onde \bar{A}_i são obrigatoriamente as contribuições dos elementos ou das subestruturas para \tilde{A}

$$\tilde{A} = \prod_{i=1}^n \left(I + \frac{1}{2} \bar{A}_i \right) \prod_{i=n}^1 \left(I + \frac{1}{2} \bar{A}_i \right) \quad (14a)$$

$$\tilde{A} = \prod_{i=1}^n \left[I + U_{CR}^*(\bar{A}_i) \right] \prod_{i=1}^{n^*} D_{CR}(\bar{A}_i) \prod_{i=n}^1 \left[I + U_{CR}(\bar{A}_i) \right] \quad (14b)$$

$$\tilde{A} = \prod_{i=1}^n \left[I + U_{CH}^*(\bar{A}_i) \right] \prod_{i=n}^1 \left[I + U_{CH}(\bar{A}_i) \right] \quad (14c)$$

$$\tilde{A} = \prod_{i=1}^n \left[I + U_{TR}^*(\bar{A}_i) \right] \prod_{i=n}^1 \left[I + U_{TR}(\bar{A}_i) \right] \quad (14d)$$

As seguintes observações cabem neste ponto do trabalho: (i) as fatorações (14) têm a vantagem de reduzir substancialmente a necessidade de memória, uma vez que cada elemento ou subestrutura é processado de cada vez. Como elementos sem graus de liberdade em comum não se influenciam em (14) o processamento pode ser feito em paralelo para estes elementos [13]. (ii) as fatorações (14) garantem que \tilde{A} seja simétrica e positiva definida. (iii) (14a) é um caso particular de (13). (iv) (12) é um caso particular de (14d) se houver apenas uma subestrutura igual a própria estrutura. (v) a formação de subestruturas em (14) não é necessa

riamente vantajosa. (vi) as fatorações (14) foram testadas em [13] e [14].

4.2 Prê-condicionamento por aproximações diretas

Em [2], [3] e [4] são apresentadas outras formas de se criar a matriz B aproximante de A. A mais simples é adotar

$$B = D_{GS}(A) \quad (15)$$

Outra possibilidade é adotar uma matriz diagonal de blocos retirada de A

$$B = \left[\begin{array}{cccc} A_1 & & & \\ & \dots & & \\ & & A_i & \\ & & & \dots \\ & & & & A_n \end{array} \right] \quad (16)$$

onde os semi-colchetes indicam a formação em diagonal das matrizes A_i . Estas são submatrizes de A que contem a sua diagonal. Pode-se demonstrar que B é simétrica e positiva definida. (16) tem a vantagem de que os blocos A_i podem ser processados em paralelo e podem ser tão grandes quanto a memória disponível permita.

A terceira possibilidade é adotar

$$B = \tilde{U}^*(A) \tilde{U}(A) \quad (17)$$

onde $\tilde{U}(A)$ é uma matriz triangular obtida pela fatoração de A, considerando-se neste processo apenas os elementos que tenham a posição de elementos não nulos de A. B tem portanto a mesma necessidade de memória que A. (17) no entanto é difícil de ser utilizada em conexão com cálculo paralelo e não é necessariamente positiva definida.

A quarta possibilidade é utilizar uma fatoração do tipo

$$B = \tilde{\tilde{U}}^*(A) \tilde{\tilde{U}}(A) \quad (18)$$

onde $\tilde{\tilde{U}}$ é obtida pela fatoração incompleta de A. Isto é obtido fazendo elementos de uma linha de $\tilde{\tilde{U}}$ menores que uma certa porcentagem (ψ) do respectivo elemento diagonal iguais a zero. Isto pode reduzir substancialmente a necessidade de memória para o armazenamento de B.

(18) apresenta no entanto algumas desvantagens. (i) é de difícil programação. (ii) é difícil de ser utilizada em conexão com cálculo paralelo. (iii) a previsão de necessidade de memória para B é difícil. (iv) B pode ser não positiva definida impossibilitando a fatoração (18).

Pode-se no entanto antever modos de se contornar estas desvantagens. Se (18) for aplicada em conexão com (16) as desvantagens (i) a (iii) podem desaparecer. (iv) pode ser evitada por um artifício sugerido em [2]. Desta forma (18) pode ser uma das maneiras mais promissoras de se pré-condicionar o sistema.

5. EXEMPLOS NUMÉRICOS

Alguns experimentos numéricos foram feitos com os dois métodos iterativos apresentados neste trabalho e com algumas das possibilidades de pré-condicionamento apresentadas na seção 4.

O primeiro exemplo analisado foi um pórtico plano quadrilado com retículo de 1,00m, sendo no total 4 retículos em cada direção. O módulo de elasticidade adotado foi 300 MN/m², a área da seção transversal das barras 0,5m² e o momento de inércia 0,05m⁴. O número de equações é igual a 60, sendo que a matriz de rigidez apresenta uma banda, cuja semi-largura é 14. A matriz é esparsa com coeficiente de esparsidade 0,304 (número de elementos não nulos/número de elementos dentro da banda). Na tabela 1 abaixo estão apresentados o número de iterações necessários até a convergência para diversas estratégias. O programa com o método dos gradientes conjugados recebeu o nome de GRACON e o com o método quase-Newton o nome de BFGS. A opção de pré-condicionamento é indicada na primeira coluna através do número da expressão no texto. No uso do programa BFGS foi adotado um número ilimitado de atualizações.

Tabela 1. Exemplo 1

Matriz B	GRACON	BFGS
I	43	43
(15)	37	37
(16) nº de blocos = 15	27	27
(16) nº de blocos = 10	29	29
(16) nº de blocos = 5	27	27
(16) nº de blocos = 2	17	17
(12)	37	37
(17)	16	16
(18) $\psi=5\%$	9	9
(18) $\psi=10\%$	12	12
(18) $\psi=5\%$ + (16) com 2 blocos	12	12

Pode-se concluir da tabela 1 que as opções do tipo (18) são as mais eficientes. A última linha mostra a eficiência da combinação de (18) com (16). Os dois métodos iterativos foram totalmente equivalentes como pode se aferir da mesma tabela. O coeficiente de esparsidade para as três últimas linhas foi igual a 0,494, 0,394 e 0,411, respectivamente.

A seguir renumerou-se os nós do pórtico em forma de caracol, obtendo-se uma semilargura de banda igual a 41. Os resultados são apresentados na tabela 2.

Tabela 2. Exemplo 2

Matriz B	GRACON	BFGS
I	43	43
(15)	37	37
(16) nº de blocos = 10	30	30
(16) nº de blocos = 5	27	27
(16) nº de blocos = 2	23	23
(12)	42	56
(17)	17	17
(18) $\psi = 5\%$	8	8

A menos da opção (12) os métodos iterativos foram equivalentes. Novamente a opção (18) mostrou-se a mais eficiente. A tabela 2 mostra que as opções não são muito sensíveis a renumerações, ao contrário dos métodos de eliminação.

A opção (12) teve um péssimo desempenho mostrando que para aproximações do tipo (14) a subestrutura não é vantajosa. Conclusão semelhante foi apresentada em [14].

Finalmente foi analisada uma viga em balanço de 25,00m de comprimento, dividida em 25 elementos iguais ao dos exemplos 1 e 2. Este é um problema que não pode ser resolvido em precisão simples em computadores do tipo IBM, Burroughs e UNIVAC. O número de equações é 50 (os graus de liberdade horizontais foram suprimidos) e a semilargura da banda é 4. Os resultados são apresentados na tabela 3.

Tabela 3. Exemplo 3

Matriz B	GRACON	BFGS
I	81	50
(15)	79	51
(16) nº de blocos = 13	57	48
(12)	71	73
(18) $\psi = 5\%$	9	8
(18) $\psi = 5\%$ + (16) com 13 blocos	57	48

A aplicação dos métodos iterativos num problema mal condicionado nos revela que o método quase-Newton BFGS tem em geral um melhor desempenho e que a opção (18) é sensivelmente mais eficiente. É interessante notar que as outras opções testadas pouco trouxeram em relação a opção $B = I$. Finalmente deve-se salientar que os erros de arredondamento fizeram que neste exemplo o número de iterações necessárias até a convergência fosse superior ao número de equações em várias opções.

Observação final: todos os exemplos aqui mostrados foram calculados no computador CDC-CYBER 170/750 do CTA.

6. CONCLUSÕES

O trabalho apresenta e compara dois métodos iterativos para a solução de sistemas de equações lineares em engenharia estrutural, o método dos gradientes conjugados e o método quase-Newton, BFGS. Experimentos numéricos preliminares mostram a equivalência destes métodos, com a exceção de um exemplo mal condicionado onde o método quase-Newton é superior.

Além disso várias formas de se pré-condicionar o sistema são apresentados e discutidos. Nos experimentos numéricos a fatoração de Choleski incompleta combinada com blocos diagonais mostra-se como uma alternativa eficiente e que combina o uso de toda capacidade da memória central combinada com processamento paralelo.

BIBLIOGRAFIA

- [1] Hestenes, M.R. e Steifel, E., "Methods of Conjugate Gradients for Solving Linear Systems", Journal of Research of National Bureau of Standards, Vol. 49, nº 6, December 1952.
- [2] Jennings, A. e Malik, G.M., "The solution of sparse linear equations by the conjugate gradient method", International Journal for Numerical Methods in Engineering, Vol. 12, 1978.
- [3] Glowinski, R., Periaux, J. e Dinh, Q.V., "Domain Decomposition Methods for Nonlinear Problems in Fluid Dynamics", INRIA Report nº 147, July 1982.
- [4] Glowinski, R., "Numerical Methods for Nonlinear Variational Problems", Springer Series in Computational Physics, Springer Verlag, 1983.
- [5] Broyden, C.G., "A new Double-Rank Minimization Algorithm", Notices American Mathematic Society, Vol. 16, 1969.
- [6] Fletcher, R., "A new Approach to Variable Metric Algorithms", Computation Journal, Vol. 13, 1970.

- [7] Goldfarb, D., "A Family of Variable Metric Methods derived by Variational Means", Math. Comp., Vol. 24 , 1969.
- [8] Shanno, D.F., "Conditioning of Quasi-Newton Methods for Function Minimization, Math. Comp. Vol. 24, 1969.
- [9] Dennis, J.E. e More, J.J., "Quasi-Newton Methods: Motivation and Theory", SIAM Review, Vol. 19, nº 1, January 1977.
- [10] Matties, H. e Strang, G., "The Solution of Nonlinear Finite Element Equations", International Journal of Numerical Methods in Engineering, Vol. 14, nº 11, 1979.
- [11] Argyris, J.H., Doltsinis, J., Pimenta, P.M. e Wustenberg H. "Thermomechanical Response of Solids at High Strains - Natural Approach", Computer Methods in Applied Mechanics and Engineering, Vol. 31/32, 1982.
- [12] Marchuk, G.I., "Methods of Numerical Mathematics", Springer-Verlag, 1975.
- [13] Hughes, T.J.R., Winget, J., Levit, I. e Tezduyar, T.E. , "New alternating direction procedures in finite element analysis based upon EBE approximate factorizations" , Recent Developments in Computer Methods for Nonlinear Solid and Structural Mechanics, (eds. S.N. Atluri and N. Perrone), ASME Applied Mechanics Symposia Series , June 1983.
- [14] Argyris, J.H., Doltrinis, J., Pimenta, P.M. e Wustenberg, H. "Natural Finite Element Techniques for Viscous Flow Problems", 5th International Symposium on Finite Element Methods in Flow Problems, University of Texas at Austin, January, 1984.