Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 5, N. 1, 2017.

Trabalho apresentado no CNMAC, Gramado - RS, 2016.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Um Método para o Cálculo da Inversa de Matrizes Simétricas e Positivas Definidas em Bloco

Moisés Ceni[1]
Departamento de Matemática e Desenho, CAp-UERJ, UERJ, Rio de Janeiro, RJ
Programa de Pós-Graduação em Engenharia Mecânica, FEN, UERJ, Rio de Janeiro, RJ
Luiz Mariano Carvalho[2]
Departamento de Matemática Aplicada, IME, UERJ, Rio de Janeiro, RJ
Michael Souza[3]
Departamento de Estatística e Matemática Aplicada, UFC, Ceará, CE

Resumo. Nosso objetivo é propor um novo algoritmo para o cálculo de inversas de matrizes simétricas e positivas definidas (SPD) em bloco. Em [1], os autores propõem um algoritmo baseado no processo de Gram-Schmidt, utilizando um produto interno induzido por uma matriz SPD, para ser usado como precondicionador para o Método de Gradientes Conjugados. Aqui propomos uma generalização desse processo para matrizes SPD em bloco, que também será utilizado, posteriormente, como precondicionador.

Palavras-chave. Matriz SPD, Inversa Aproximada, Matriz em Bloco, Precondicionador.

1 Introdução

Durante todo o texto, as matrizes são reais. Dada uma matriz quadrada $A \in \mathbb{R}^{n\times n}$, o i, j-ésimo bloco de A será denotado por $A_{[i,j]}$. Analogamente, dada uma matriz bloco coluna b, o i-ésimo bloco de b será denotado por $b_{[i]}$.
É natural ainda definir o vetor bloco

$$e_{[i]}^T := [0\,0\ldots I_n\ldots 0],$$

onde o i-ésimo bloco é a identidade.

Definição 1.1. *Dados a matriz $A \in \mathbb{R}^{n\times n}$ e os vetores blocos u, $m\times n$, e v, $m\times n$, dizemos que u e v, com $u \neq v$, são A conjugados se $u^T Av = 0$ (matriz nula de ordem n). Esta relação pode ser escrita como $u^T Av = \delta_{uv}P$, onde δ_{uv} é a função delta de Kronecker e $P \in \mathbb{R}^{n\times n}$.*

Denote por $a_{[i]}^T$ a i-ésima linha bloco de A. No que se segue, A é uma matriz SPD.

[1]moisesceni@gmail.com
[2]luizmc@ime.uerj.br
[3]souza.michael@gmail.com

DOI: 10.5540/03.2017.005.01.0321

2 Regras Gerais

A matriz $A_{m\times m}$ é uma matriz em blocos e cada bloco $A_{[i,j]}$ é $n \times n$, onde $m = n \times k$, $k \in \mathbb{N}$. Se conseguirmos obter uma coleção $\{z_{[i]}\}$, $i = 1, \ldots, k$, de vetores blocos A conjugados, onde cada vetor bloco $z_{[i]}$ é uma matriz $m \times n$, podemos formar a matriz Z, $m \times m$:

$$Z = [z_{[1]}\ z_{[2]}\ \ldots\ z_{[k]}],$$

de modo que

$$Z^T A Z = D,$$

onde D é uma matriz $m \times m$ diagonal por blocos.

Se A e Z forem não singulares, segue que D é não singular e obteremos

$$A^{-1} = Z D^{-1} Z^T.$$

O algoritmo a seguir constrói as matrizes Z e D.

Algoritmo 1 - Método da Inversa

1: $z_{[i]}^{(0)} = e_{[i]}, 1 \leq i \leq k$
2: **for** $i = 1, \ldots, k$ **do**
3: **for** $j = i, \ldots, k$ **do**
4:
$$P_j^{(i-1)} := \begin{cases} (z_{[i]}^{(i-1)})^T A z_{[j]}^{(i-1)} & i \neq j \\ a_{[i]}^T z_{[i]}^{(i-1)} & i = j \end{cases}$$
5: **end for**
6: **if** $i = k$ **then**
7: vá para 13
8: **end if**
9: **for** $j = i+1, \ldots, k$ **do**
10:
$$z_{[j]}^{(i)} := z_{[j]}^{(i-1)} - z_{[i]}^{(i-1)} (P_i^{(i-1)})^{-1} P_j^{(i-1)}$$
11: **end for**
12: **end for**
13: $z_{[i]} := z_{[i]}^{(i-1)}$ and $P_i := P_i^{(i-1)}, 1 \leq i \leq k$
14: **return** $Z = [z_{[1]}\ z_{[2]}\ \ldots\ z_{[k]}]$, and $D = \begin{pmatrix} P_1 & 0_{n\times n} & \ldots & 0_{n\times n} \\ 0_{n\times n} & P_2 & \ldots & 0_{n\times n} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{n\times n} & 0_{n\times n} & \ldots & P_k \end{pmatrix}$

Lema 2.1. *Se $A \in \mathbb{R}^{n\times n}$ é SPD, então A é não singular.*

Demonstração. Veja [4], Fact 3.24. $\square$

DOI: 10.5540/03.2017.005.01.0321

Lema 2.2. *Se $A \in \mathbb{R}^{n\times n}$ é SPD, então todas as submatrizes líderes principais e o complemento de Schur são matrizes SPD.*

Demonstração. Veja [4], Fact 3.27. □

Teorema 2.1. *O Algoritmo 1 produz uma matriz D não singular e uma matriz Z triangular superior (por blocos), cujas entradas diagonais são a identidade e cujas colunas são A conjugadas.*

Demonstração. A prova seguirá por indução: mostraremos que a matriz Z é triangular com entradas diagonais iguais à identidade fazendo a indução sobre as colunas de Z, ou seja, mostraremos que $z_{[i]}$ só tem preenchimento não nulo até o i-ésimo bloco, o qual é a identidade, com indução em i. Juntamente com essa indução, chegaremos à conclusão de que as colunas $z_{[i]}$ de Z são A conjugadas. Finalmente, demonstraremos que cada P_i gerado é não singular.

ETAPA 1 - Primeiro caso.
A primeira coluna $z_{[1]}$ de Z nada mais é que

$$z_{[1]} = \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Portanto, $z_{[1]}$ só tem preenchimento não nulo até o primeiro bloco, o qual é a identidade. Agora, pelos Lemas 2.1 e 2.2, $P_1 = A_{[1,1]}$ é não singular. Segue ainda que $z_{[1]}$ A conjuga com ele mesmo.

ETAPA 2
Já provamos que vale o primeiro passo da indução. Provaremos agora que os $(l-1)$-ésimos primeiros passos implicam no l-ésimo passo. Mas o que dizem os $(l-1)$-ésimos primeiros passos? São três informações:

- Se $z_{[i]} \in \{z_{[1]}, \dots, z_{[l-1]}\}$, então $z_{[i]}$ só tem preenchimento não nulo até o i-ésimo bloco, o qual é a identidade;
- $\{z_{[1]}, \dots, z_{[l-1]}\}$ são A conjugados;
- $P_1, \dots, P_{l-1}$ são não singulares.

Nosso intuito é concluir que:

- $z_{[l]}$ só tem preenchimento não nulo até o l-ésimo bloco, o qual é a identidade;
- $z_{[l]}, A$ conjuga com $z_{[i]} \in \{z_{[1]}, \dots, z_{[l-1]}\}$;
- P_l é não singular.

ETAPA 3 - $z_{[l]}$ só tem preenchimento não nulo até o l-ésimo bloco, o qual é a identidade.
Podemos escrever, no l-ésimo passo do algoritmo:

$$z_{[l]} = z_{[l]}^{(l-1)} = z_{[l]}^{(l-2)} - z_{[l-1]} P_{l-1}^{-1} P_l^{(l-2)}.$$

Por sua vez,

$$z_{[l]}^{(l-2)} = z_{[l]}^{(l-3)} - z_{[l-2]} P_{l-2}^{-1} P_l^{(l-3)},$$

e, portanto,

$$z_{[l]} = z_{[l]}^{(l-3)} - z_{[l-2]} P_{l-2}^{-1} P_l^{(l-3)} - z_{[l-1]} P_{l-1}^{-1} P_l^{(l-2)}.$$

Aplicando este mesmo raciocínio, concluímos que:

$$z_{[l]} = z_{[l]}^{(0)} - z_{[1]} P_1^{-1} P_l^{(0)} - z_{[2]} P_2^{-1} P_l^{(1)} - \ldots - z_{[l-2]} P_{l-2}^{-1} P_l^{(l-3)} - z_{[l-1]} P_{l-1}^{-1} P_l^{(l-2)}.$$

Agora, $z_{[l]}^{(0)}$ só tem preenchimento não nulo no l-ésimo bloco, que é a identidade. Por outro lado, os $(l-1)$ passos anteriores implicam que $z_{[i]}$ só tem preenchimento não nulo até o i-ésimo bloco, para $i = 1, \ldots, l-1$.
Segue que $z_{[l]} = z_{[l]}^{(l-1)}$ não tem preenchimento depois do l-ésimo bloco e é igual à identidade no l-ésimo bloco.

ETAPA 4 - $z_{[l]}, A$ conjuga com $z_{[i]} \in \{z_{[1]}, \ldots, z_{[l-1]}\}$.
Seja $z_{[i]} \in \{z_{[1]}, z_{[2]}, \ldots, z_{[l-1]}\}$. Então,

$$\begin{aligned} z_{[i]}^T A z_{[l]} &= z_{[i]}^T A (z_{[l]}^{(l-2)} - z_{[l-1]} P_{l-1}^{-1} P_l^{(l-2)}) \\ &= z_{[i]}^T A z_{[l]}^{(l-2)} - z_{[i]}^T A z_{[l-1]} P_{l-1}^{-1} P_l^{(l-2)}. \end{aligned}$$

Agora, se $i = l-1$, então $z_{[i]}^T A z_{[l-1]} = P_{l-1}$ e portanto,

$$z_{[i]}^T A z_{[l]} = z_{[i]}^T A z_{[l]}^{(l-2)} - P_l^{(l-2)} = 0.$$

Por outro lado, se $i \neq l-1$, então $z_{[i]}^T A z_{[l-1]} = 0$ e

$$\begin{aligned} z_{[i]}^T A z_{[l]} &= z_{[i]}^T A z_{[l]}^{(l-2)} \\ &= z_{[i]}^T A (z_{[l]}^{(l-3)} - z_{[l-2]} P_{l-2}^{-1} P_l^{(l-3)}) \\ &= z_{[i]}^T A z_{[l]}^{(l-3)} - z_{[i]}^T A z_{[l-2]} P_{l-2}^{-1} P_l^{(l-3)}. \end{aligned}$$

Novamente, se $i = l-2$, então $z_{[i]}^T A z_{[l-2]} = P_{l-2}$ e, assim

$$z_{[i]}^T A z_{[l]} = z_{[i]}^T A z_{[l]}^{(l-3)} - P_l^{(l-3)} = 0.$$

Se, por outro lado, $i \neq l-2$, $z_{[i]}^T A z_{[l-2]} = 0$ e teremos:

$$z_{[i]}^T A z_{[l]} = z_{[i]}^T A z_{[l]}^{(l-3)}.$$

Repetindo este mesmo argumento quantas vezes for necessário, chegaremos no último caso possível, ou seja:

$$z_{[i]}^T A z_{[l]} = z_{[i]}^T A z_{[l]}^{(0)} - P_l^{(0)} = 0.$$

ETAPA 5 - P_l é não singular.
Seja B_l o $(l \cdot n) \times (l \cdot n)$ bloco líder principal de A. Então, B_l é não singular pelo Lema 2.1.
Se $z_{[i]} \in \{z_{[1]}, z_{[2]}, \dots, z_{[l]}\}$ então, pela etapa 3, $z_{[i]}$ tem o i-ésimo bloco igual à identidade e só possui blocos não nulos antes do i-ésimo bloco, inclusive este. Logo,

$$z_{[l]}^T A$$

anula $a_{[j]}^T$ para $j > l$.
Reciprocamente,

$$A z_{[l]}$$

anula a coluna $a_{[j]}$ de A para $j > l$.
Portanto, $z_{[l]}^T A z_{[l]}$ é equivalente a:

$$\tilde{z}_{[l]}^T B_l \tilde{z}_{[l]},$$

onde $\tilde{z}_{[l]}$ é a restrição de $z_{[l]}$ aos l primeiros blocos.
Pela etapa 4 e pelos $(l-1)$ passos anteriores, $z_{[l]}$ A conjuga com $z_{[i]} \in \{z_{[1]}, \dots, z_{[l-1]}\}$ e $\{z_{[1]}, \dots, z_{[l-1]}\}$ são A conjugados. Como $\tilde{Z} = [\tilde{z}_{[1]} \ \tilde{z}_{[2]} \ \dots \ \tilde{z}_{[l]}]$ é não singular e B_l também é não singular. Segue que

$$\tilde{Z}^T B_l \tilde{Z} = \tilde{D} = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_l \end{pmatrix}$$

tem posto completo. Então P_l é não singular.

$\square$

Corolário 2.1. *Sejam B_i o bloco líder principal $(i \cdot n) \times (i \cdot n)$ de A e $\Delta_i = \det(B_i)$. Vale a relação $\det(P_i) = \dfrac{\Delta_i}{\Delta_{i-1}}$, onde $\Delta_0 = 1$, $i = 1, \dots, k$.*

Demonstração. A relação $\tilde{Z}^T B_l \tilde{Z} = \tilde{D}$ nos diz que $det(B_l) = \det(\tilde{D}) = \det(P_1) \cdot \ldots \cdot \det(P_l)$, pois $\det(\tilde{Z}) = 1$. Assim, $\det(P_1) = \det(B_1) = \dfrac{\Delta_1}{\Delta_0}$.
Agora, se $\det(P_k) = \dfrac{\Delta_k}{\Delta_{k-1}}$, então

$$\Delta_{k+1} = \det(B_{k+1}) = \det(P_1) \cdot \det(P_{k+1}) = \frac{\Delta_1}{\Delta_0} \cdot \frac{\Delta_2}{\Delta_1} \cdot \ldots \cdot \frac{\Delta_k}{\Delta_{k-1}} \det(P_{k+1}) = \Delta_k \det(P_{k+1}).$$

Portanto, o resultado segue imediatamente. $\square$

No artigo [1] é provado que, se implantarmos uma estrutura arbitrária de zeros na parte estritamente superior de Z ou eliminarmos os termos que tenham uma magnitude menor que um parâmetro previamente estabelecido, o Algoritmo 1 (bloco escalar) não quebrará (isto é, $P_i \neq 0$). Para tanto é suficiente que A seja uma matriz H [4]. Lembramos que qualquer matriz diagonalmente dominante é uma matriz H.

Considere a matriz

$$A = \begin{pmatrix} 2,00 & 0,40 & 0,10 & 0,00 \\ 0,40 & 1,08 & 2,00 & 0,00 \\ 0,10 & 2,00 & 3,96 & 0,00 \\ 0,00 & 0,00 & 0,00 & 1,00 \end{pmatrix},$$

e uma tolerância para exclusão de $Tol = 0,06$ para as entradas da parte estritamente superior de Z. Aplicando o Algoritmo 1 para o caso onde o bloco tem ordem um (que coincide com o algoritmo proposto por [1]), há uma quebra em P_3, produzindo:

$$D = \begin{pmatrix} 2,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 1,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & NaN \end{pmatrix}.$$

Considere novamente a matriz A com uma tolerância para exclusão de $Tol = 0,06$ para as entradas da parte estritamente superior de Z. Vamos aplicar o Algoritmo 1 para o caso onde os blocos tenham ordem 2, ou seja, vamos repartir a matriz A em

$$A = \begin{pmatrix} A_{[1,1]} & A_{[1,2]} \\ A_{[2,1]} & A_{[2,2]} \end{pmatrix},$$

onde

$$A_{[1,1]} = \begin{pmatrix} 2,00 & 0,40 \\ 0,40 & 1,08 \end{pmatrix}, A_{[1,2]} = \begin{pmatrix} 0,10 & 0,00 \\ 2,00 & 0,00 \end{pmatrix}, A_{[2,1]} = \begin{pmatrix} 0,10 & 2,00 \\ 0,00 & 0,00 \end{pmatrix},$$

$$e \quad A_{[2,2]} = \begin{pmatrix} 3,96 & 0,00 \\ 0,00 & 1,00 \end{pmatrix}.$$

Neste caso, as matrizes produzidas pelo Algoritmo 1 são:

$$D = \begin{pmatrix} 2,0000 & 0,4000 & 0,0000 & 0,0000 \\ 0,4000 & 1,0800 & 0,0000 & 0,0000 \\ 0,0000 & 0,0000 & 0,0346 & 0,0000 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \end{pmatrix}$$

[4] A é uma matriz H se $\tilde{A} = \tilde{a}_{ij}$ é uma matriz M, onde

$$\tilde{a}_{ij} := \begin{cases} -|a_{ij}| & i \neq j, \\ a_{ii} & i = j. \end{cases}$$

e

$$Z = \begin{pmatrix} 1,0000 & 0,0000 & 0,0346 & 0,0000 \\ 0,0000 & 1,0000 & -1,9800 & 0,0000 \\ 0,0000 & 0,0000 & 1,0000 & 0,0000 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \end{pmatrix}.$$

Repare que neste caso não houve quebra. Essa diferença é justificada pelo Corolário 2.1, pois quanto maior for a ordem do bloco, menor é a quantidade de menores principais não nulos que são necessários para que o algoritmo funcione.

3 Conclusões

O método proposto estende o resultado apresentado em [1] para matrizes SPD com uma estrutura em blocos. A família de matrizes para a qual o Algoritmo 1 não quebra é maior do que a família de matrizes para a qual o algoritmo proposto em [1] funciona.

Agradecimentos

O primeiro autor agradece à CAPES pelo apoio ao projeto de pesquisa. Agradecemos ainda a Univesidade do Estado do Rio de Janeiro e, em especial, ao Programa de Pós-Graduação em Engenharia Mecânica. Agradecemos ainda aos revisores pelo cuidadoso trabalho feito.

Referências

[1] M. Benzi, C. D. Meyer and M. Tuma. A Sparse Approximate Inverse Preconditioner for the Conjugate Gradiente Method, *Soc. Ind. Appl. Math.*, 17:1135-1149, 1996.

[2] M. Benzi and M. Tuma. A Sparse Approximate Inverse Preconditioner for Nonsymmetric Linear Systems, *Soc. Ind. Appl. Math.*, 19:968-994, 1998.

[3] R. A. Horn and C. R. Johnson. *Matrix Analysis*, Cambridge University Press, New York, 2013.

[4] I. C. F. Ipsen. *Numerical Matrix Analysis, Linear Systems an Least Squares*, Soc. Ind. Appl. Math., Philadelphia, 2009.

[5] M. J. Tsatsomeros. Principal Pivot Transforms: properties and applications, *Lin. Alg. Appl.*, 307:151–165, 1999.

DOI: 10.5540/03.2017.005.01.0321