Um estudo sobre decomposição matricial
MS512 - Análise Numérica
IMECC - UNICAMP

Gustavo Guedes Faria RA 174112
Inayê Caroline Melo RA 199030
Lucas Hideki Ueda RA 156368

2019/1S

Resumo

Um estudo sobre decomposições matriciais por 3 algoritmos de forma a facilitar a resolução de sistemas lineares. Avaliação do numero de condicionamento e suas propriedades para os metodos.

Palavras-chave: Decomposição, Sistema Linear, Cholesky, Matrizes Ortogonais, Refletores,HouseHolder, Matriz inversa, Gram-Schmidt.

Introdução

Diversas são as formas de se resolver numericamente um sistema linear, decomposições matriciais nos ajudam a deixar essa tarefa mais fácil, especialmente quando for necessário resolver vários sistemas lineares com a mesma matriz de coeficientes. Neste trabalho, estudaremos três dessas decomposições, individualmente e comparando sua eficiência.

1 Os algortimos

1.1 Decomposição de Cholesky

A Decomposição de Cholesky busca decompor a matriz $A = G^tG$. G é uma matriz triangular superior e positiva definida e é chamado fator de Cholesky.

Teorema: Seja $A \in R^{nxn}$ uma matriz simétrica, todas as afirmações são equivalentes:

1. A é positiva definida, isto é, $x^tAx > 0$, para todo $x \neq 0$
2. Os autovalores de A são positivos
3. Os autovalores das submatrizes principais A_k são positivos

4. Existe uma matriz $G \in R^{nxn}$ inversível, com $A = G^t G$

Teorema da decomposição de Cholesky: Seja A uma matriz positiva definida simétrica. Então A pode ser decomposta de forma única no produto $A = G^t G$, onde G é o fator de Cholesky, triangular superior, e possui os elementos da diagonal $g_{i,i} > 0$.

Para obter o fator de Cholesky de uma matriz A, encontra-se as seguintes relações entre as matrizes A e G:

$g_{i,i} = \sqrt{a_{i,i} - \sum_{k=1}^{i-1}(g_{k,i}^2)}$

$g_{i,j} = \frac{a_{i,j} - \sum_{k=1}^{i-1}(g_{k,i}, g_{k,j})}{g_{i,i}}$

Dessa forma existe uma maneira conveniente de se obter os elementos do fator de Cholesky, como é descrito abaixo:

Obter o primeiro elemento: $g_{1,1} = \sqrt{a_{1,1}}$

Obter os elementos da primeira coluna da matriz: $g_{i,1} = \frac{a_{i,1}}{g_{1,1}}$

Para cada coluna seguinte, fazer:

Calcular o primeiro elemento não nulo da coluna (elemento da diagonal);

Obter os elementos abaixo da diagonal;

1.2 Transformação de Householder

Uma matriz Q é dita uma transformação de Householder quando está é uma transformação que mapeia cada vetor x em sua reflexão através do hiperplano:

$$\{v \in R^n :< u, v >= 0\}$$

As demais caracteristicas, sua construção e implementação encontram-se abaixo.

Teorema 1.2.1: Seja $u \in R^n$ unitário, $||u|| = 1$, e defina $Q \in R^{nxn}$ por,

$$Q = I - 2uu^T$$

então:

1. $Qu = -u$
2. $Qv = v$, se v é ortogonal a u então $< u, v >= 0$
3. $Q = Q^T$ (Simetria)
4. $Q^T = Q^{-1}$ (Q é ortogonal)
5. $Q = Q^{-1}$

Proposição 1.2.1: Seja $u \in R^n$ um vetor não nulo em R^n e defina,

$$\gamma = \frac{2}{||u||^2}, \text{ e} = I - \gamma uu^T$$

Então Q é um refletor.

Teorema 1.2.2: Sejam $x, y \in R^n$ com $x \neq y$, mas, $||x|| = ||y||$. Então, existe um único refletor Q tal que

$$Qx = y$$

Corolário 1.2.2: Seja $x \in R^n$ um vetor não nulo. Então existe um refletor Q tal que

$$Q \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \cdots \\ 0 \end{bmatrix} \text{ onde } || * || = ||x||$$

Em Watkins 2ed [3.2.45] é definido o algoritmo abaixo para cálculo da decomposição QR utilizando refletores:

Algorithm 1 Fatoração QR por refletores

---

while *Não percorreu todas as colunas de A* **do**

  Determine um refletor $Q_k = I - \gamma_k.u^{(k)}.u^{(k)t}$ tal que

  $Q_k \left[a_{kk}..a_{nk}\right] = \left[-\tau_k 0...0\right]$

  Guarde $u^{(k)}$ sobre $a_{k:n,k}$

  $a_{k:n,k+1:n} = Q_k a_{k:n,k+1:n}$

  $a_k k = \gamma_k$

end

$\gamma_k = a_{nn}$

Para calcular o refletor Q utilizamos o algoritmo descrito em Watkin 2ed [3.2.37].

Algorithm 2 Algoritmo que determina o refletor Q

$\beta = max_{1 \leq i \leq n} |x_i|$

if $\beta = 0$ **then**

  $\gamma = 0$

else

  $x_{1:n} = x_{1:n}/\beta$

  $\tau = \sqrt{x_1^2 + ... + x_n^2}$

  if $x_1 < 0$ **then**

    $\tau = -\tau$

  else

    $\tau = \tau$

  end

  $x_1 = \tau + x_1$

  $\gamma = x_1/\tau$

  $x_{2:n} = x_{2:n}/x_1$

  $x_1 = 1$

  $\tau = \tau\beta$

end

1.3 Matrizes Ortogonais: Gram Schmidt

A fatoração QR utilizando decomposição ortogonal por Gram-Schmidt utiliza o processo de Gram-Schmidt, que se consiste em criar uma base ortogonal e eliminar sucessivamente as partes linearmente dependentes para encontrar novos elementos ortogonais.

A ideia central é: utilizando o processo de Gram-Schmidt, encontrar $Q \in R^{nxn}$ ortogonal, tal que $Q^T Ax = Q^T b$ implica $Rx = Q^T b = y$, resolvendo facilmente o sistema por retro-substituição.

Teorema 1.3.1: Sejam $A, Q \in R^{nxn}$ com Q matriz ortogonal, então existe R triangular superior tal que $A = QR$.

Algorithm 3 Processo de Gram-Schmidt

```
for k = 1,2,...,n do
    for i=1,2,...,k-1 do
        r_ik = v_k v_i
        v_k = v_k - v_i r_ik
    end
    r_kk = ||v_k||
    v_k = v_k / r_kk
end
```

2 Inversa de matriz

Para determinarmos a inversa de uma matriz $A \in R^{nxn}$ basta resolvermos o sistema $Ax = b$ com $b = I$, I da matriz identidade nxn. Dessa forma, na presente seção estudaremos a eficiência das diferentes fatorações para resolver esse tipo de sistema.

Todos os testes se deram para matrizes de Hilbert para n = 3, 6 e 15. As matrizes de Hilbert são matrizes onde seus elementos são da forma:

$$\frac{1}{i + j - 1} \tag{1}$$

Como exemplo para n = 3 temos que a matriz de Hilbert tem a forma:

$$\begin{pmatrix} 1 & 0.5 & 0.33 \\ 0.5 & 0.33 & 0.25 \\ 0.33 & 0.25 & 0.2 \end{pmatrix} \tag{2}$$

Por Cholesky

Para Hilbert com n=3, com Cholesky, conseguimos a seguinte fatoração:

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.29 & 0 \\ 0,33 & 0,29 & 0.07 \end{pmatrix}$$

$$\mathrm{G}^t = \begin{pmatrix} 1 & 0.5 & 0.33 \\ 0 & 0.29 & 0.29 \\ 0 & 0 & 0.07 \end{pmatrix}$$

(3)

Dessa forma resolvendo o sistema $Ax = GG^T x = I$ obtemos a inversa A^{-1} como segue:

$$x = A^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix} \tag{4}$$

De fato, $A.A^{-1} = I$:

$$A.A^{-1} = \begin{pmatrix} 1 & 0 & 3e^{-16} \\ -1e^{-15} & 1 & 2e^{-15} \\ -3e^{-15} & 0 & 1 \end{pmatrix} \tag{5}$$

Para n = 6, a precisão continuou satisfatória, como segue nos códigos enviados. Para n>13 a precisão não se manteve.

Por Refletores

Para matrizes ortogonais o sistema linear a ser resolvido fica da forma:

$$Ax = I \longrightarrow QRx = I \longrightarrow Rx = Q^T I \tag{6}$$

Para Hilbert com n=3 conseguimos QR como segue abaixo:

$$Q = \begin{pmatrix} -0.85 & 0.50 & 0.11 \\ -0.42 & -0.56 & -0.70 \\ -0.28 & -0.65 & 0.70 \end{pmatrix}$$

$$\text{R} = \begin{pmatrix} -1.16 & -0.64 & -0.45 \\ -8e^{-17} & -0.10 & -0.10 \\ -5e^{-17} & -6e^{-18} & 0.003 \end{pmatrix} \text{ (7)}$$

É possível notar que R é triangular superior, com precisão de 10^{-17} e Q é ortogonal, pois $Q^T Q = I$ (com precisão de 10^{-16}), como segue abaixo:

$$Q^T Q = \begin{pmatrix} 1 & -1e^{-17} & -4e^{-17} \\ -1e^{-17} & 1 & 1e^{-16} \\ 4e^{-17} & 1e^{-16} & 1 \end{pmatrix} \tag{8}$$

Dessa forma resolvendo o sistema $Ax = I$ obtemos a inversa A^{-1} como segue:

$$x = A^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix} \tag{9}$$

De fato, $A.A^{-1} = I$:

$$A.A^{-1} = \begin{pmatrix} 1 & 8e^{-15} & 1e^{-15} \\ 8e^{-16} & 1 & 0 \\ 1e^{-15} & -1e^{-15} & 1 \end{pmatrix} \tag{10}$$

Para n = 6 e n = 12 manteve-se a precisão, como segue nos codigos enviados. Para n grande a precisão não se manteve e a inversa não foi assertiva.

Por Gram Schmidt

Para Hilbert com n=3 conseguimos QR como segue abaixo:

$$Q = \begin{pmatrix} 0.85 & -0.50 & 0.11 \\ 0.42 & 0.56 & -0.70 \\ 0.28 & 0.65 & 0.70 \end{pmatrix}$$

$$\text{R} = \begin{pmatrix} 1.16 & 0.64 & 0.45 \\ 7e^{-16} & 0.10 & 0.10 \\ -2e^{-14} & -2e^{-14} & 0.004 \end{pmatrix} \quad (11)$$

É possível notar que R é triangular superior, com precisão de 10^{-14} e Q é ortogonal, pois $Q^TQ = I$ (com precisão de 10^{-14}), como segue abaixo:

$$Q^TQ = \begin{pmatrix} 1 & -1e^{-14} & 4e^{-15} \\ -1e^{-14} & 1 & 2e^{-15} \\ 4e^{-15} & 2e^{-15} & 1 \end{pmatrix} \tag{12}$$

Dessa forma resolvendo o sistema $Ax = I$ obtemos a inversa A^{-1} como segue:

$$x = A^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix} \tag{13}$$

De fato, $A.A^{-1} = I$:

$$A.A^{-1} = \begin{pmatrix} 1 & 8e^{-15} & 1e^{-15} \\ 8e^{-16} & 1 & 0 \\ 1e^{-15} & -1e^{-15} & 1 \end{pmatrix} \tag{14}$$

Para n=6 e n=12 foram replicados os estudos e o resultado foi muito similar para n=6. Com n $\geq$ 14, os erros se tornam muito mais significativos, divergindo bastante dos valores esperados na matriz identidade.

3 Número de condicionamento

O número de condicionamento é uma medida indicando se o problema tem "boas condições"para ser tratado numericamente. Um problema com um número de condição pequeno é chamado de bem condicionado, enquanto os problemas que possuem um número de condição elevado são denominados mal condicionados.

Utilizando a 1-norma matricial (Equação x) temos que sua definição é dada por:

$$||A||_1 = max_{i \leq j \leq n} \sum_{i=1}^{n} |a_{ij}| \tag{15}$$

$$K_1(A) = ||A^{-1}||_1.||A||_1 \tag{16}$$

No presente trabalho checaremos as condições mostradas na equação (k) e (j) sobre o número de condicionamento.

$$K_1(A) \geq \frac{||a_i||}{||a_j||} \text{ , } para\ qualquer\ coluna\ i\ e\ j\ de\ A. \tag{17}$$

$$K_1(A) \geq \frac{||A||_1.||A^{-1}.w||_1.}{||w||_1} \text{ , } w \in R^n \tag{18}$$

É notável para testes realizados em uma matriz de Hilbert de dimensão 3, usando relfetores Householder, que essas condições são satisfeitas, como mostram as tabelas v e s.

(Coluna i Coluna j)	$\frac{\|\|a_i\|\|}{\|\|a_j\|\|}$
(1,2)	1.69
(1,3)	2.34
(2,1)	0.59
(2,3)	1.38
(3,1)	0.42
(3,2)	0.72

Tabela 1 – Combinações possíveis para avaliação da relação da equação 2.

É notável que nenhum valor chega perto do número de condicionamento $K1 = 748$ de uma matriz de Hilbert de ordem 3. O mesmo foi testado para Gram-Schmidt e Cholesky e o resultado foi o mesmo, inclusive com $K1 = 748$ em ambos.

Novamente o número de condicionamento satisfaz sua propriedade.

Testes foram realizados para matrizes de Hilbert de ordem maiores, apresentando os mesmo resultados positivos, e optou-se pela apresentação para ordem n=3 pela visuabilidade.

Conclusão

Através desse trabalho foi possível verificar a eficácia de algoritmos de fatoração na resolução de sistemas lineares e compreender o erro computacional envolvido. Em especial, foi possível notar o efeito cumulativo de tal erro, sendo cada vez mais perturbador ao desempenho do programa à medida que n aumenta.

Vetor aleatório	$\frac{\|\|A\|\|_1.\|\|A^{-1}.w\|\|_1.}{\|\|w\|\|_1}$
(44,5,80)	461
(48,25,5)	109
(86,86,95)	57
(90,17,59)	252
(92,61,0)	215
(4,75,68)	47
(93,90,40)	116
(1,79,48)	192
(8,20,1)	453
(63,95,96)	24

Tabela 2 – 10 vetores aleatórios para checar condição da equação 3

Referências

1. Notas de aula
2. Watkins 2ed