Decomposição LU e Cholesky

Prof. Doherty Andrade - DMA-UEM

Sumário

1 Introdução **1**

2 Método de Eliminação de Gauss **1**

3 Decomposição LU **2**

4 O método de Cholesky **5**

5 O Algoritmo para a decomposição Cholesky **7**

1 Introdução

A decmposição LU e a decomposição de Cholesky são métodos diretos de resolução de sistemas de equações lineares $Ax = b$. O objetivo é decompor a matriz A como $A = LU$ no caso da decomposição LU, $A = LL^T$ no caso da decomposição de Cholesky.

Antes de falarmos sobre a decomposição LU e Cholesky precisamos relembrar o método de eliminação de Gauss. A decomposição LU é uma conseqüência do método de eliminação de Gauss.Isto é, a decomposição de Cholesky é uma especialização da decomposição LU. Os métodos de decomposição são importantes na resolução de sistemas de equações lineares, como veremos a seguir. Veja (3.2) e (3.3).

2 Método de Eliminação de Gauss

Este método é um método direto para resolver sistemas de equações lineares, consiste em tranformar a matriz ampliada do sistema em uma outra equivalente que seja triangular superior. Para isto usa apenas operações elementares sobre linhas da matriz.

Uma vez tendo o sistema na forma triangular superior, a solução é obtida pela substituição inversa.

Dado o sistema $Bx = b$, seja A a matriz ampliada do sistema. Em cada passo k, vamos denotar por $A^{(k)} = (a_{ij}^{(k)})$ a matriz obtida após este passo.

Os passos são os seguintes:

Passo 1: eliminar $a_{k1}, k = 2, 3, \ldots, n$, obtendo $A^{(1)}$

Passo 2: eliminar na matriz $A^{(1)}$, $a_{k2}, k = 3, 4, \ldots, n$.

Prosseguindo até atingir o passo $(n-1)$ cujo objetivo é eliminar $a_{n,n-1}$

Os elementos $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{nn}^{(n-1)}$são chamados de pivot e os números

$$m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, i = k+1, \ldots, n \tag{2.1}$$

são chamados de multiplicadores.

Existem dois problemas neste método que é instável:

(1i) o pivot nulo e

(2i) pivot próximo de zero.

Nestes casos deve-se adotar uma estratégia de pivoteamento que reduz erros de arredondamento. A estratégia de pivoteamento parcial consiste em tomar para pivot o maior elemento (em módulo) da coluna. No pivoteamento completo ou total, toma-se, em cada passo, como pivot o maior elemento (em módulo) **dentre os elementos da matriz que ainda atuam no processo** de eliminação de Gauss. O pivoteamento completo é pouco utilizado, pois envolvendo muitas comparações torna-se trabalhoso e consome tempo e memória de máquina.

Pode-se provar que as estratégias de pivoteamento minimizam erros. O pivoteamento parcial e o pivoteamento completo minimizam erros igualmente.

3 Decomposição LU

A decomposição LU é uma variante do método de eliminação de Gauss.

Dizemos que uma matriz $M_{n\times n}$ admite uma decomposição LU, se existem matrizes L e U, triangulares inferior e superior, respectivamente, tais que $M = LU$.

Dizemos que uma matriz $A_{n\times n}$ é diagonalmente dominante se

$$|a_{ii}| \geq \sum_{j=1, j\neq i}^{n} |a_{ij}|, i = 1, 2, .., n.$$

A matriz A é estritamente diagonalmente dominante se

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|, i = 1, 2, .., n.$$

Teorema 1 *Toda matriz estritamente diagonalmente dominante tem inversa. Além disso, o processo de eliminação de Gauss pode ser realizado sem a necessidade de permutação de linhas ou colunas.*

Conhecer a decomposição em LU de uma matriz A é útil e facilita na resolução de um sistema de equações lineares $Ax = b$. Esta é a grande utilidade dos métodos de decomposição ou fatoração.

Vejamos como isso acontece, é muito simples.

Suponha que desejamos obter a solução do sistema de equações lineares $Ax = b$, mas sabendo que A se fatora como $A = LU$. Então resolvemos dois sistemas de equações lineares mais simples:

$$Ly = b \tag{3.2}$$
$$Ux = y. \tag{3.3}$$

Note que o sistema (3.2) é simples (pois é um sistema triangular inferior) e o sistema (3.3) é também simples (pois é um sistema triangular superior). Para a solução do sistema (3.2) usamos substituição inversa, enquanto que para a solução do sistema (3.3) usamos substituição direta.

Lembramos que a cada operação elementar sobre as linhas de uma matriz A, corresponde uma matriz chamada elementar E tal que EA é o resultado desta operação. A matriz elementar E é obtida da matriz identidade realizando sobre ela a mesma operação elementar realizada sobre a matriz A.

No método de **eliminação de Gauss**, realizamos operações elementares sobre a matriz A com o objetivo de triangularizá-la, obtendo em cada passo as matrizes

$$A = A^{(0)}, A^{(1)}, ..., A^{(n-1)}$$

sendo $A^{(n-1)}$triangular superior. Logo, podemos notar do processo de eliminação de Gauss que a matriz triangular superior $A^{(n-1)}$ é obtida por meio de um produto de matrizes elementares:

$$A^{(n-1)} = E_{n-1}E_{n-2}...E_2E_1A.$$

Ou seja,

$$A = (E_{n-1}E_{n-2}...E_2E_1)^{-1}A^{(n-1)} = (E_1^{-1}E_2^{-1}...E_{n-1}^{-1})A^{(n-1)}.$$

Note que $A^{(n-1)}$ é triangular superior. Se evitarmos permutações de linhas ou colunas sobre A **(não pode realizar permutações)** a matriz dada por $(E_1^{-1}E_2^{-1}...E_{n-1}^{-1})$ é triangular inferior com 1 na diagonal principal (produto de triangulares inferiores com 1 na diagonal é uma matriz triangular inferior com 1 na diagonal).

Estas observações são a base para a demonstração do seguinte resultado.

Teorema 2 *Se $A_{n\times n}$ é uma matriz com todos os determinantes de suas submatrizes principais diferentes de zero, então, existe uma única matriz triangular inferior $L = (m_{ij})$ (estes são os multiplicadores dados em (2.1)) com $m_{ii} = 1$, e uma matriz triangular superior $U = (u_{ij})$ tal que $A = LU$.*

O teorema acima na verdade dá um processo para construir as matrizes L e U: a matriz L é obtida a partir dos multiplicadores m_{ij} que você armazena durante o processo de eliminação de Gauss, sendo $m_{ii} = 1$; a matriz U é a matriz triangular superior obtida ao final do processo de eliminação de Gauss.

Uma versão mais simples do teorema acima é a seguinte:

Teorema 3 *Se o método de eliminação de Gauss pode ser realizado sobre $A_{n\times n}$ sem trocas de linhas ou colunas, então A pode ser escrita como $A = LU$, onde $L = (m_{ij})$ é triangular inferior com $m_{ii} = 1$, e $U = (u_{ij})$ matriz triangular superior.*

Reforçando: (i) Nas matrizes estritamente diagonalmente dominantes A, os sistemas $Ax = b$ podem ser resolvidos pelo método de eliminação de Gauss sem a necessidade de permutar linhas ou colunas.

(2i) Se $A_{n\times n}$ é uma matriz com todos os determinantes de suas submatrizes principais diferentes de zero, então o método de eliminação de Gauss pode ser aplicado sem a necessidade de permutar linhas ou colunas.

4 O método de Cholesky

A decomposição de Cholesky procura decompor uma matriz A na forma $A = LL^T$, onde L é uma matriz triangular inferior com elementos da diagonal principal estritamente positivos. Para tanto, exige-se muito mais da matriz A.

Uma matriz A é dita definida positiva se A é simétrica e se $x^T Ax > 0$, para todo $x \neq 0$.

Por exemplo, a matriz $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ é simétrica e além disso,

$$[xyz] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2x^2 - 2xy + 2y^2 - 2yz + 2z^2 =$$

$$= x^2 + (x-y)^2 + (y-z)^2 + z^2 > 0.$$

Teorema 4 *(a) Se $A_{n\times n}$ é uma matriz simétrica, então A possui n autovalores reais e n autovetores ortonormais.*
(b) Uma matriz simétrica A é definida positiva se, e somente se, seus autovalores são reais e positivos.
(c) Uma matriz simétrica A é definida positiva se, e somente se, cada uma de suas submatrizes principais tem determinante positivo.
(d) Uma matriz simétrica A é definida positiva se, e somente se, o processo de eliminação de Gauss pode ser realizado sem permutação de linhas ou colunas e tem todos os elementos pivots positivos.

Teorema 5 (Cholesky) *Uma matriz simétrica A é definida positiva se, e somente se, pode ser fatorada como LL^T, onde L é uma matriz triangular inferior com elementos positivos na diagonal.*

Como obter a matriz L do método de Cholesky?

$$\begin{bmatrix} a_{11} & a_{21} & ... & a_{n1} \\ a_{21} & a_{22} & ... & a_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & ... & a_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 & 0 & ... & 0 \\ l_{21} & l_{22} & 0 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & 0 & \vdots & 0 \\ l_{k1} & l_{k2} & ... & l_{kk} & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ l_{n1} & l_{n2} & .. & l_{nk} & ... & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & ... & l_{k1} & ... & l_{n1} \\ 0 & l_{22} & ... & l_{k2}... & ... & l_{n2} \\ 0 & 0 & ... & ... & ... & ... \\ 0. & 0 & ... & l_{kk}... & ... & l_{nk} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & ... & 0 & 0 & l_{nn} \end{bmatrix}$$

A maneira mais prática para obter os coeficientes (l_{ij}) é começando pela primeira coluna, depois para a a segunda coluna e assim por diante.

Coluna 1 (a_{j1})

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = \begin{bmatrix} l_{11}^2 \\ l_{21}l_{11} \\ \vdots \\ l_{n1}l_{11} \end{bmatrix}$$

então

$$\begin{aligned} l_{11} &= \sqrt{a_{11}} \\ l_{j1} &= \frac{a_{j1}}{l_{11}}, j = 2, 3, ..., n \end{aligned}$$

Coluna 2 (a_{j2})

$$\begin{bmatrix} a_{21} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{n2} \end{bmatrix} = \begin{bmatrix} l_{11}l_{21} \\ l_{21}^2 + l_{22}^2 \\ l_{31}l_{21} + l_{32}l_{22} \\ \vdots \\ l_{n1}l_{21} + l_{n2}l_{22} \end{bmatrix}$$

então obtemos (já conhecemos l_{21} e l_{11})

$$\begin{aligned} l_{22} &= \sqrt{a_{22} - l_{21}^2} \\ l_{j2} &= \frac{a_{j2} - l_{j1}l_{21}}{l_{22}}, j = 3, 4, ..., n \end{aligned}$$

Coluna k (a_{jk})

$$\begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kk} \\ \vdots \\ a_{kj} \\ \vdots \\ a_{nk} \end{bmatrix} = \begin{bmatrix} l_{11}l_{21} \\ l_{21}l_{k1} + l_{22}l_{k2} \\ \vdots \\ l_{k1}^2 + l_{k2}^2 + l_{k3}^2 + ... + l_{kk}^2 \\ \vdots \\ l_{j1}l_{k1} + l_{j2}l_{k2} + ... + l_{jk}l_{kk} \\ \vdots \\ l_{n1}l_{k1} + l_{n2}l_{k2} + ... + l_{nk}l_{kk} \end{bmatrix}$$

donde obtemos

$$\begin{aligned} a_{kk} &= l_{k1}^2 + l_{k2}^2 + l_{k3}^2 + ... + l_{kk}^2 \\ a_{kj} &= l_{j1}l_{k1} + l_{j2}l_{k2} + ... + l_{jk}l_{kk} \end{aligned}$$

ou ainda,

$$\begin{aligned} l_{kk} &= \sqrt{a_{kk} - (l_{k1}^2 + l_{k2}^2 + l_{k3}^2 + ... + l_{k,k-1}^2)} \\ l_{kj} &= \frac{\left[a_{jk-} \sum_{i=1}^{k-1} l_{ji}l_{ki}\right]}{l_{kk}}, j = k+1, k+2, ...n. \end{aligned}$$

5 O Algoritmo para a decomposição Cholesky

Apresentamos a seguir o algoritmo para a decomposição Cholesky.

Algoritmo 6

Dados: dimensão n, entradas $a_{i,j}, 1 \leq i \leq n,\ 1 \leq j \leq n$.

Passo 1: Seja $l_{11} = \sqrt{a_{11}}$

Passo 2: Para $j = 2, \ldots, n$ seja $l_{j1} = \frac{a_{j1}}{l_{11}}$.

Passo 3: Para $i = 2, \ldots, n-1$ faça passo 4 e passo 5.

Passo 4: Seja

$$l_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)^{\frac{1}{2}}$$

Passo 5: Para $j = i+1, \ldots, n$ seja

$$l_{ji} = \frac{\left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik} \right)}{l_{ii}}$$

Passo 6: Seja

$$l_{nn} = \left(a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2 \right)^{\frac{1}{2}}.$$

Passo 7: Imprima $l_{ij}, i = 1, \ldots, n; j = 1, \ldots, n.$

Exercício 7 *1. Determine se a matriz admite a decomposição de Cholesky, em caso afirmativo determine esta decomposição.*

$$A = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 10 & 4 \\ -4 & 4 & 9 \end{bmatrix}$$

2. Determine se a matriz admite a decomposição de Cholesky, em caso afirmativo determine esta decomposição.

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

3. Use decomposição de Cholesky para obter as soluções dos sistemas $Ax = b$ e $Bx = b$, onde

$$b = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}.$$

Referências

[1] S. D. Conte, Elementary Numerical Analysis. MacGraw-Hill, 1965.