数值代数

Homework 3

黄文翀　3200100006

A Course Homework Assignment

2022 年 3 月 13 日

1 书面题解答

Problem 1

证明：若 A 是一个带宽为 $2n+1$ 的对称正定带状矩阵，则其 Cholesky 因子 L 也是带状矩阵. L 的带宽为多少？

Solution.

记 $A=(a_{ij})_{m\times m}$，则 a_{ij} 具有性质：

$$a_{ij}=0 \qquad \forall i=n+1+j,...,m$$

首先断言，L 是一个带宽为 $n+1$ 的下三角带状矩阵，以下按列归纳证明.

考虑第一列，有：

$$l_{11}=\sqrt{a_{11}}$$

$$l_{i1}=\frac{a_{i1}}{l_{11}} \qquad \forall i=2,...,m$$

考虑到 a_{ij} 的性质，即有：

$$l_{i1}=0 \qquad \forall i=n+2,...,m$$

现在假设前 $k-1$ 列的性质已被证明，考虑第 k 列，当 $i\geq n+1+k$ 时，有：

$$l_{ij}=0 \qquad \forall j=1,...,k-1$$

$$a_{ij}=0 \qquad \forall j=1,...,k-1$$

于是由：

$$a_{ik}=\sum_{j=1}^{k} l_{ij}l_{kj}$$

得：

$$0=l_{ik}l_{kk}$$

由正定矩阵的 Cholesky 分解性质知 $l_{kk}\neq 0$，因此必有 $l_{ik}=0$

这样就证明了第 k 列的性质，归纳即得 L 是带宽为 $n+1$ 的下三角带状矩阵.

Problem 2

若 $A = LL^T$ 是 A 的 Cholesky 分解，试证：L 的 i 阶顺序主子阵 L_i 恰好是 A 的 i 阶顺序主子阵 A_i 的 Cholesky 因子.

Solution.
对任一个 i，将 A 与 L 分块如下：

$$A = \begin{bmatrix} A_i & A_i^{(1)} \\ A_i^{(2)} & A_i^{(3)} \end{bmatrix} \qquad L = \begin{bmatrix} L_i & O \\ L_i^{(1)} & L_i^{(2)} \end{bmatrix}$$

于是有：

$$L^T = \begin{bmatrix} L_i^T & \left(L_i^{(1)}\right)^T \\ O & \left(L_i^{(2)}\right)^T \end{bmatrix}$$

注意到：

$$LL^T = \begin{bmatrix} L_iL_i^T & L_i\left(L_i^{(1)}\right)^T \\ L_i^{(1)}L_i^T & L_i^{(1)}\left(L_i^{(1)}\right)^T + L_i^{(2)}\left(L_i^{(2)}\right)^T \end{bmatrix}$$

即有：

$$A_i = L_iL_i^T$$

Problem 3

证明：若 $A \in \mathbb{R}^{n\times n}$ 是对称的，而且其前 $n-1$ 个顺序主子阵均非奇异，则 A 有唯一分解式 $A = LDL^T$，其中 L 是单位下三角阵，D 是对角阵.

Solution.

首先证明**存在性**.

由 LU 分解定理知，存在一个单位下三角矩阵 L 和上三角矩阵 $\tilde{U}$，使得：$A = L\tilde{U}$

令：

$$D = \mathrm{diag}(\tilde{u}_{11}, ..., \tilde{u}_{nn}), \quad U = D^{-1}\tilde{U}$$

则有：

$$U^T DL^T = A^T = A = LDU$$

从而：

$$L^T U^{-1} = D^{-1} U^{-T} LD$$

左式是单位上三角阵，右式是单位下三角阵，因此两边都是单位阵，从而知道 $U = L$，于是：$A = LDL^T$

下面证明**唯一性**.

设：

$$L = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \qquad D = \mathrm{diag}(d_1, ..., d_n)$$

考虑第一列，对比 $A = LDL^T$ 两侧元素，得：

$$d_1 = a_1 \neq 0, \qquad l_{i1} = \frac{a_{i1}}{d_1}(i = 2, ..., n)$$

这就唯一确定了第一列. 现在假设前 $k-1$ 列已经唯一确定，考虑第 k 列，有：

$$a_{kk} = d_k + \sum_{j=1}^{k-1} l_{kj}^2 d_j$$

$$a_{ik} = l_{ik} d_k + \sum_{j=1}^{k-1} l_{ij} l_{kj} d_j \qquad i = k+1, ..., n$$

由此得：

$$d_k = a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2 d_j$$

$$l_{ik} = \frac{1}{d_k}\left(a_{ik} - \sum_{j=1}^{k-1} l_{ij} l_{kj} d_j\right) \qquad i = k+1, ..., n$$

这样第 k 列也唯一确定，按列归纳即可唯一确定 L 与 D.

Problem 4

设:

$$A = \begin{pmatrix} 16 & 4 & 8 & 4 \\ 4 & 10 & 8 & 4 \\ 8 & 8 & 12 & 10 \\ 4 & 4 & 10 & 12 \end{pmatrix} \qquad b = \begin{pmatrix} 32 \\ 26 \\ 38 \\ 30 \end{pmatrix}$$

用平方根法证明 A 是正定的，并给出方程组 $AX = b$ 的解.

Solution.

使用 Cholesky 分解法可求得下三角矩阵 L，使得 $A = LL^T$，

$$L = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 2 & 2 & 2 & 0 \\ 1 & 1 & 3 & 1 \end{pmatrix}$$

故 A 是正定的. 根据 $Ly = b$ 解得:

$$y = \begin{pmatrix} 8 \\ 6 \\ 5 \\ 1 \end{pmatrix}$$

再由 $L^T x = y$ 得:

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Problem 5

给出按行计算 Cholesky 因子 L 的详细算法.

Solution.
设 $A = LL^T$，考虑第一行，有：

$$l_{11} = \sqrt{a_{11}}$$

现在假设前 $k-1$ 行已经完成计算，对于第 k 行，有：

$$a_{kj} = \sum_{s=1}^{j} l_{ks} l_{js} \qquad j = 1, ..., k$$

于是得到递推计算式：

$$l_{kj} = \frac{1}{l_{jj}} \left(a_{kj} - \sum_{s=1}^{j-1} l_{ks} l_{js} \right) \qquad j = 1, ..., k-1$$

$$l_{kk} = \sqrt{a_{kk} - \sum_{s=1}^{k-1} l_{ks}^2}$$

按行递推即可计算出 L.

Problem 6

设 $H = A + iB$ 是一个正定 Hermite 矩阵，其中 $A, B \in \mathbb{R}^{n\times n}$.

（1）证明：矩阵

$$C = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

是正定对称的;

（2）试给出一种仅用实数运算的算法来求解线性方程组

$$(A + iB)(x + iy) = (b + ic), \qquad x, y, b, c \in \mathbb{R}^n.$$

Solution.

（1）由 Hermite 矩阵的定义，$A + iB = H = H^H = A^T - iB^T$，即:

$$A = A^T \qquad B = -B^T$$

于是有:

$$C^T = \begin{pmatrix} A^T & B^T \\ -B^T & A^T \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = C$$

故 C 是对称矩阵.

由正定矩阵的定义，$\forall x \in \mathbb{R}^n$，$0 < x^T H x = x^T A x + i x^T B x$，即:

$$x^T A x > 0 \qquad x^T B x = 0 \qquad \forall x \in \mathbb{R}^n$$

$\forall y \in \mathbb{R}^{2n}$，记 $y = (y_1^T\ y_2^T)^T$，其中 $y_1, y_2 \in \mathbb{R}^n$，有:

$$y^T C y = \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1^T A y_1 + y_2^T B y_2 - y_1^T B y_1 + y_2 A y_2 = (y_1 + y_2)^T A (y_1 + y_2) > 0$$

故 C 也是正定矩阵.

（2）展开得：$(Ax - By) + i(Bx + Ay) = b + ic$

即求解:

$$Ax - By = b \qquad Bx + Ay = c$$

设 $z = (x^T\ y^T)^T$，注意到:

$$Cz = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} Ax - By \\ Bx + Ay \end{pmatrix}$$

因此只需求解方程:

$$Cz = \begin{pmatrix} b \\ c \end{pmatrix}$$

再取 z 的前 n 个分量作为 x，后 n 个分量作为 y 即可.

2 上机题报告

2.1 朴素 Cholesky 分解法求解正定对称方程的通用子程序

直接使用按列递推公式计算 L 的全部元素，利用一些矩阵和向量运算，可以简洁地编写如下程序:

```
function L = cholesky(A)
    n = size(A,1);
    L = zeros(n,n);
    for k = 1:n
        L(k,k) = sqrt(A(k,k) - sum(L(k,1:k-1).^2));
        L(k+1:n,k) = (A(k+1:n,k) - L(k+1:n,1:k-1)*(L(k,1:k-1).'))/
            L(k,k);
    end
end

function x = solveEquationWithCholesky(A, b)
    A = cholesky(A);
    y = solveLowerTriangularEquation(A, b);
    x = solveUpperTriangularEquation(A.', y);
end
```

2.2 改进的 Cholesky 分解法求解正定对称方程的通用子程序

直接使用按列递推公式计算 L,D 的全部元素，利用一些矩阵和向量运算，可以简洁地编写如下程序:

```
function [L, D] = choleskyImproved(A)
    n = size(A,1);
    L = eye(n);
    D = zeros(1,n);
    for j = 1:n
        v = D(1:j-1).*L(j,1:j-1);
        D(j) = A(j,j) - dot(L(j,1:j-1),v);
        L(j+1:n,j) = (A(j+1:n,j)-L(j+1:n,1:j-1)*(v.'))/D(j);
    end
end

function x = solveEquationWithCholeskyImproved(A, b)
    [A, D] = choleskyImproved(A);
    y = solveLowerTriangularEquation(A, b)./(D.');
    x = solveUpperTriangularEquation(A.', y);
end
```

2.3 上机习题 1.2(1) 与 1.3(1) 报告

本题中，我令 b 中的每个元素在 $[0, 20]$ 区间内均匀随机. 生成 A 与 b 后，分别使用上述两个子程序以及 LU 分解法、列主元 LU 分解法的通用子程序求解方程，由于不知道方程的真解，只能通过计算 $\|Ax-b\|_2$ 来确定误差的大小. 测试使用的代码如下（文件 ex_1_2_1.m）:

```
A = zeros(100,100);
A(1,1) = 10;
for i = 2:100
    A(i,i) = 10;
    A(i-1,i) = 1;
    A(i,i-1) = 1;
end
b = unifrnd(0,20,100,1);
x1 = solveEquationWithCholesky(A,b);
x2 = solveEquationWithCholeskyImproved(A,b);
x3 = solveEquationWithLU(A,b);
x4 = solveEquationWithPLU(A,b);
disp([vecnorm(A*x1-b),vecnorm(A*x2-b),vecnorm(A*x3-b),vecnorm(A*x4
    -b)]);
```

进行了多次测试，其结果如下:

算法	误差计算式	误差（$\times 10^{-13}$）				
朴素 Cholesky 分解	$\|Ax_1-b\|_2$	0.1930	0.2187	0.1800	0.2206	0.2166
改进 Cholesky 分解	$\|Ax_2-b\|_2$	0.1085	0.1302	0.1407	0.1499	0.1450
朴素 LU 分解	$\|Ax_3-b\|_2$	0.1000	0.1119	0.0903	0.1366	0.1171
列主元 LU 分解	$\|Ax_4-b\|_2$	0.1000	0.1119	0.0903	0.1366	0.1171

朴素 Cholesky 分解的误差高于改进 Cholesky 分解，这是意料之中的，但是 LU 分解与列主元 LU 分解的结果令人相当惊讶：朴素 LU 分解的计算精度竟高于列主元法 LU 分解，也高于改进 Cholesky 算法.

我仔细检查了四个通用子程序，并没有发现任何错误。因此**我推测改进 Cholesky 分解法仅仅是一个时间上占优的算法，在计算精度上并不具有优越性.**

另外，本测试中朴素 LU 分解与列主元 LU 分解之间没有体现出任何差异，这是由矩阵 A 的特殊性质决定的，在第 k 次消元过程中主元永远就是第 k 行那个元素本身，因此不需要选主元. 本题测试列主元法显然是多余的.

2.4 上机习题 1.2(2) 与 1.3(2) 报告

生成 A 与 b 后，分别使用上述两个子程序以及 LU 分解法、列主元 LU 分解法的通用子程序求解方程。方程的真解显然是 $s = (1, 1, ..., 1)^T_{40}$，通过计算 $\|x-s\|_2$ 来确定误差的大小. 测试使用的代码如下（文件 ex_1_2_2.m）:

```
n = 40;
A = zeros(n,n);
b = zeros(n,1);
```

```
s = ones(n,1);
for i = 1:n
    for j = 1:n
        A(i,j) = 1.0/(i+j-1);
    end
    b(i) = sum(A(i,1:n));
end
x1 = solveEquationWithCholesky(A,b);
x2 = solveEquationWithCholeskyImproved(A,b);
x3 = solveEquationWithLU(A,b);
x4 = solveEquationWithPLU(A,b);
disp([vecnorm(x1-s),vecnorm(x2-s),vecnorm(x3-s),vecnorm(x4-s)]);
```

测试结果为:

$$||x_1 - s|| = 2539$$

$$||x_2 - s|| = 605.1$$

$$||x_3 - s|| = 467.7$$

$$||x_4 - s|| = 677.1$$

这一测试的结果与上一项测试类似. 也印证了上一项测试的结论. 但在这项测试中, 朴素 Cholesky 分解与改进 Cholesky 分解的计算精度差距明显了许多.