MA60053 - Computational Linear Algebra Direct methods for linear system of equations (To be updated)

M. Rajesh Kannan

Department of Mathematics, Indian Institute of Technology Kharagpur, email: rajeshkannan1.m@gmail.com, rajeshkannan@maths.iitkgp.ac.in



・ロット (雪) (き) (き) (き)

1/20

Definition

For $1 \le k \le n-1$, let $m \in \mathbb{R}^n$ be a vector with $e_j^T m = 0$ for $1 \le j \le k$, meaning that m is of the form

$$m = (0, 0, \ldots, 0, m_{k+1}, \ldots, m_n)^T.$$

An elementary lower triangular matrix is a lower triangular matrix of the specific form

$$L_{k}(m) = I - me_{k}^{T} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & -m_{k+1} & 1 & \\ & \vdots & \ddots & \\ & & -m_{n} & & 1 \end{bmatrix}$$

An elementary lower triangular matrix $L_k(m)$ has the following properties:

- **1** det $L_k(m) = 1$.
- 2 $L_k(m)^{-1} = L_k(-m).$
- Multiplying a matrix A with L_k(m) from the left leaves the first k rows unchanged and, starting from row j = k + 1, subtracts the row m_j(a_{k1},..., a_{kn}) from row j of A.

(日) (雪) (日) (日) (日)

Gaussian elimination method

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix, and $b \in \mathbb{R}^n$. Set $A^{(1)} = A$, $b^{(1)} = b$,

Gaussian elimination method

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix, and $b \in \mathbb{R}^n$. Set $A^{(1)} = A$, $b^{(1)} = b$, and then iteratively

$$A^{(k+1)} = L_k A^{(k)}, b^{(k+1)} = L_k b^{(k)},$$

where

$$L_k = I - m_k e_k^T$$

and

$$m_k = \left(0, 0, \dots, 0, \frac{a_{k+1,k}^{(k)}}{a_{kk}^{(k)}}, \frac{a_{k+2,k}^{(k)}}{a_{kk}^{(k)}}, \dots, \frac{a_{n,k}^{(k)}}{a_{kk}^{(k)}}\right),$$

provided $a_{kk}^{(k)} \neq 0$.

Gaussian elimination method

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be a non-singular matrix, and $b \in \mathbb{R}^n$. Set $A^{(1)} = A$, $b^{(1)} = b$, and then iteratively

$$A^{(k+1)} = L_k A^{(k)}, b^{(k+1)} = L_k b^{(k)},$$

where

$$L_k = I - m_k e_k^T$$

and

$$m_k = \left(0, 0, \dots, 0, \frac{a_{k+1,k}^{(k)}}{a_{kk}^{(k)}}, \frac{a_{k+2,k}^{(k)}}{a_{kk}^{(k)}}, \dots, \frac{a_{n,k}^{(k)}}{a_{kk}^{(k)}}\right),$$

provided $a_{kk}^{(k)} \neq 0$. Assuming that, the process does not end prematurely, it stops to the linear system $A^{(n)}x = b^{(n)}$, where $A^{(n)}$ is upper triangular, and it has the same solution as Ax = b.

Definition

A matrix $L = (I_{ij}) \in \mathbb{R}^{n \times n}$ is called a normalized lower triangular matrix if $I_{ij} = 0$ for i < j and $I_{ii} = 1$.

Definition

A matrix $L = (I_{ij}) \in \mathbb{R}^{n \times n}$ is called a normalized lower triangular matrix if $I_{ij} = 0$ for i < j and $I_{ii} = 1$.

Definition

The LU factorization of a matrix A is the decomposition A = LU in to the product of a normalized lower triangular matrix L, and an upper triangular matrix U.

Definition

A matrix $L = (I_{ij}) \in \mathbb{R}^{n \times n}$ is called a normalized lower triangular matrix if $I_{ij} = 0$ for i < j and $I_{ii} = 1$.

Definition

The LU factorization of a matrix A is the decomposition A = LU in to the product of a normalized lower triangular matrix L, and an upper triangular matrix U.

Theorem

Let $A \in \mathbb{R}^{n \times n}$. Let $A_p \in \mathbb{R}^{p \times p}$ be the p-th principal submatrix of A, that is,

If $det(A_p) \neq 0$ for $1 \leq p \leq n$, then A has an LU factorisation.

Definition

A matrix $L = (I_{ij}) \in \mathbb{R}^{n \times n}$ is called a normalized lower triangular matrix if $I_{ij} = 0$ for i < j and $I_{ii} = 1$.

Definition

The LU factorization of a matrix A is the decomposition A = LU in to the product of a normalized lower triangular matrix L, and an upper triangular matrix U.

Theorem

Let $A \in \mathbb{R}^{n \times n}$. Let $A_p \in \mathbb{R}^{p \times p}$ be the p-th principal submatrix of A, that is,

If $det(A_p) \neq 0$ for $1 \leq p \leq n$, then A has an LU factorisation. In particular, every symmetric, positive definite matrix possesses an LU factorisation.

Definition

A matrix A is called strictly row diagonally dominant if

$$\sum_{k=1,k\neq i}^n |a_{ik}| < |a_{ii}|, \qquad 1 \le i \le n.$$

Theorem

A strictly row diagonally dominant matrix $A \in \mathbb{R}^{n \times n}$ is invertible and possesses an LU factorisation.

Theorem

The set of non-singular (normalised) lower (or upper) triangular matrices is a group with respect to matrix multiplication.

Theorem

If the invertible matrix A has an LU factorisation then it is unique.

Remark

If A = LU then the linear system b = Ax = LUx can be solved in two steps. First, we solve Ly = b by forward substitution and the Ux = y by back substitution. Both are possible in $O(n^2)$ time.

Remark

A numerically reasonable way of calculating the determinant of a matrix is to first compute the LU factorisation and then use the fact that $det(A) = det(LU) = det(L)det(U) = det(U) = u_{11}u_{22} \dots u_{nn}$.

In the case of a tridiagonal matrix, there is an efficient way of constructing the LU factorisation, at least under certain additional assumptions. Let the tridiagonal matrix be given by

$$A = \begin{bmatrix} a_{1} & c_{1} & & 0 \\ b_{2} & a_{2} & c_{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & b_{n} & a_{n} \end{bmatrix}$$
(1)

Assume A is a tridiagonal matrix of the form (1) with

$$egin{array}{rcl} |a_1| &> & |c_1| > 0, \ |a_i| &\geq & |b_i| + |c_i|, & b_i, c_i
eq 0, \ 2 \leq i \leq n-1, \ |a_n| &\geq & |b_n| > 0. \end{array}$$

Then, A is invertible and has an LU factorisation of the form

$$A = \begin{bmatrix} 1 & & 0 \\ l_2 & 1 & & \\ & \ddots & \ddots & \\ 0 & & l_n & 1 \end{bmatrix} \begin{bmatrix} u_1 & c_1 & 0 \\ & u_2 & \ddots & \\ & & \ddots & c_{n-1} \\ 0 & & & u_n \end{bmatrix}$$

The vectors $I \in \mathbb{R}^{n-1}$ and $u \in \mathbb{R}^n$ can be computed as follows:

$$u_1 = a_1$$
 and $l_i = \frac{b_i}{u_{i-1}}$ and $u_i = a_i - l_i c_{i-1}$ for $2 \le i \le n$.

Definition

A permutation matrix $P_{ij} \in \mathbb{R}^{n \times n}$ is called an elementary permutation matrix if it is of the form

$$P_{ij} = I - (e_i - e_j)(e_i - e_j)^T.$$

This means that the matrix P_{ij} is an identity matrix with rows (columns) i and j exchanged.

Remark

An elementary permutation matrix has the properties

$$P_{ij}^{-1} = P_{ij} = P_{ji} = P_{ij}^7$$

and det(P_{ij}) = -1, for $i \neq j$ and $P_{ii} = I$. Pre-multiplication of a matrix A by P_{ij} exchanges rows *i* and *j* of A. Similarly post-multiplication exchanges columns *i* and *j* of A.

Let A be an $n \times n$ matrix. There exists elementary lower triangular matrices $L_i = L_i(m_i)$ and elementary permutation matrices $P_i = P_{r_i i}$ with $r_i \ge i$, i = 1, 2, ..., n - 1, such that

$$U = L_{n-1}P_{n-1}L_{n-2}P_{n-2}\ldots L_2P_2L_1P_1A$$

is upper triangular.

Let A be an $n \times n$ matrix. There exists elementary lower triangular matrices $L_i = L_i(m_i)$ and elementary permutation matrices $P_i = P_{r_i i}$ with $r_i \ge i$, i = 1, 2, ..., n - 1, such that

$$U = L_{n-1}P_{n-1}L_{n-2}P_{n-2}\ldots L_2P_2L_1P_1A$$

is upper triangular.

Theorem

For every non-singular matrix $A \in \mathbb{R}^{n \times n}$ there is a permutation matrix P such that PA possesses an LU factorisation PA = LU.

(日) (雪) (日) (日) (日)

Let $A \in \mathbb{R}^{n \times n}$ be symmetric and positive definite, and let $C \in \mathbb{R}^{n \times m}$. Then,

- 2 rank($C^T A C$) = rank(C).
- **(3)** $C^{T}AC$ is positive definite if and only if rank(C) = m.

Cholesky decomposition

Definition

A decomposition of a matrix A of the form $A = LL^T$ with a lower triangular matrix L is called a Cholesky factorisation of A.

Cholesky decomposition

Definition

A decomposition of a matrix A of the form $A = LL^T$ with a lower triangular matrix L is called a Cholesky factorisation of A.

Theorem

Suppose $A = A^T$ is positive definite. Then, A possesses a Cholesky factorisation.

Cholesky decomposition

Definition

A decomposition of a matrix A of the form $A = LL^T$ with a lower triangular matrix L is called a Cholesky factorisation of A.

Theorem

Suppose $A = A^T$ is positive definite. Then, A possesses a Cholesky factorisation.

Theorem

If A is positive definite, then there exists a unique lower triangular matrix L such that the diagonal entries of L are positive, and $A = LL^{T}$.

If Q is an orthogonal matrix, then

$$\bigcirc \langle Qx, Qy \rangle = \langle x, y \rangle,$$

$$||Qx||_2 = ||x||_2,$$

$$|| QA ||_2 = || A ||_2,$$

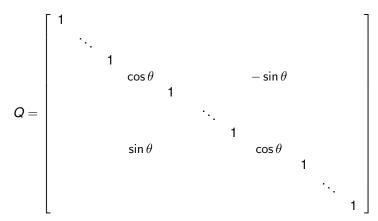
If Q_1 and Q_2 are orthogonal, then so is Q_1Q_2 ,

▲□▶ ▲□▶ ▲目▶ ▲目▶ 三目 のへで

5
$$||Q||_2^2 = ||Q^{-1}||_2^2 = 1 = k_2(Q).$$

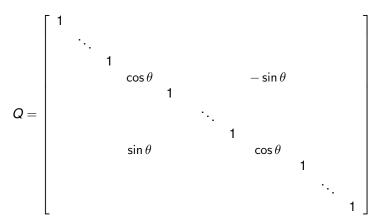
Givens rotator

A Given rotator (or Jacobi rotators, or plane rotator) is a matrix of the form:



Givens rotator

A Given rotator (or Jacobi rotators, or plane rotator) is a matrix of the form:



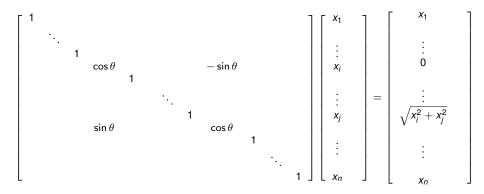
・ロト ・ 四ト ・ モト ・ モト

æ

Theorem

The matrix Q is orthogonal.

Let $x \in \mathbb{R}^n$ be such that the coordinates x_i and x_j are non-zero. Then,



QR factorisation

Theorem

For every $A \in \mathbb{R}^{m \times n}$, $m \ge n$, there exists an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ such that A = QR.

QR factorisation

Theorem

For every $A \in \mathbb{R}^{m \times n}$, $m \ge n$, there exists an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ such that A = QR.

Theorem

The factorisation A = QR of a non-singular matrix $A \in \mathbb{R}^{n \times n}$ into the product of an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and an upper triangular matrix $R \in \mathbb{R}^{n \times n}$ is unique, if the signs of the diagonal elements of R are prescribed.

QR factorisation

Theorem

For every $A \in \mathbb{R}^{m \times n}$, $m \ge n$, there exists an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ such that A = QR.

Theorem

The factorisation A = QR of a non-singular matrix $A \in \mathbb{R}^{n \times n}$ into the product of an orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and an upper triangular matrix $R \in \mathbb{R}^{n \times n}$ is unique, if the signs of the diagonal elements of R are prescribed.

Householder transformation(Reflectors)

Definition

A Householder matrix $H = H(w) \in \mathbb{R}^{m \times m}$ is a matrix of the form

$$H(w) = I - 2ww^T \in \mathbb{R}^{m \times m},$$

where $w \in \mathbb{R}^m$ satisfies either $||w||_2 = 1$ or w = 0.

Householder transformation(Reflectors)

Definition

A Householder matrix $H = H(w) \in \mathbb{R}^{m \times m}$ is a matrix of the form

$$H(w) = I - 2ww^T \in \mathbb{R}^{m \times m},$$

where $w \in \mathbb{R}^m$ satisfies either $||w||_2 = 1$ or w = 0.

A more general form of a Householder matrix is given by

$$H(w) = I - 2\frac{ww^{T}}{w^{T}w},$$

for an arbitrary vector $w \neq 0$.

Let $H = H(w) \in \mathbb{R}^{m \times m}$ be a Householder matrix.

- H is symmetric.
- 2 HH = I so that H is orthogonal.
- **3** *det*(*H*(*w*)) = −1 *if* $w \neq 0$.
- Storing H(w) only requires storing the m elements of w.
- The computation of the product Ha of H = H(w) with a vector a ∈ ℝ^m requires only O(m) time.

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三 のQ@

Let $H = H(w) \in \mathbb{R}^{m \times m}$ be a Householder matrix.

- H is symmetric.
- 2 HH = I so that H is orthogonal.
- 3 det(H(w)) = -1 if $w \neq 0$.
- Storing H(w) only requires storing the m elements of w.

Solution of the product Ha of H = H(w) with a vector a ∈ ℝ^m requires only O(m) time.

Theorem

For every vector $a \in \mathbb{R}^m$, there is a vector $w \in \mathbb{R}^m$ with w = 0 or $||w||_2 = 1$ such that $H(w)a = ||a||_2e_1$.

For every vector $a \in \mathbb{C}^m$, there is a vector $w \in \mathbb{C}^m$ with w = 0 or $||w||_2 = 1$ such that $H(w)a = ||a||_2e_1$.

Theorem (Schur factorisation)

Let $A \in \mathbb{C}^{n \times n}$. There exists a unitary matrix, $U \in \mathbb{C}^{n \times n}$, such that $R = U^* AU$ is an upper triangular matrix.

<ロト < 同ト < 回ト < 回ト = 三

Theorem (Real Schur factorisation)

To $A \in \mathbb{C}^{n \times n}$ there is an orthogonal matrix $Q \in \mathbb{C}^{n \times n}$ such that

$$Q^{T}AQ = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ 0 & R_{22} & \dots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_{mm} \end{bmatrix}$$

where the diagonal blocks R_{ii} are either 1 × 1 or 2 × 2 matrices. A 1 × 1block corresponds to a real eigenvalue, a 2 × 2 block corresponds to a pair of complex conjugate eigenvalues. If *A* has only real eigenvalues then it is orthogonal similar to an upper triangular matrix.