Lecture Note on Linear Algebra

10. The LU Factorisation

Wei-Shi Zheng,
wszheng@ieee.org, 2011

1 What Do You Learn from This Note

This lecture note is to connect the following two questions?

- Remember that why we need to introduce ROW REDUCTION ALGORITHM for any matrix A?
- We know a matrix A can be reduced to a echelon form matrix U, that is there is a linear transform L such that $A = LU$. This is also called a factorization of matrix A. How does this factorization help solving the matrix equation $A\vec{x} = \vec{b}$?

In this lecture, we focus on a particular type of factorization called LU factorization of matrix, which provides an alternative method for solving matrix equation.

Basic concept： lower triangular matrix(下三角矩阵), upper triangular matrix (上三角矩阵), LU factorization (LU 分解)

2 Triangular Matrix

We first introduce the following:

Definition 1 (lower(upper) triangular matrix，下三角阵/ 上三角阵). A square matrix A is called a lower triangular matrix if $[A]_{ij} = 0$ whenever $i < j$ where $[A]_{ij}$ is the entry of A at i^{th} row and j^{th} column. Reversely, a square matrix A is called a upper triangular matrix if $[A]_{ij} = 0$ whenever $i > j$. Furthermore, a unit lower(upper) triangular matrix is a lower(upper) triangular matrix A such that $[A]_{ii} = 1$ for all i.

注意：无论是下三角阵还是上三角阵，它们都一定是方阵。课本上132页的定义与一般的三角阵定义不同，课本允许非方阵的形式。然而，一般来说三角阵都指方阵，见wiki: http://en.wikipedia.org/wiki/Triangular_matrix

We denote $\mathrm{L}_n(\mathrm{U}_n)$ for the set of lower(upper) triangular matrices of size n and $\mathrm{L}_n^1(\mathrm{U}_n^1)$ the set of unit lower(upper) triangular matrices of size n.

Examples:

(1) Lower triangular matrix：

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{pmatrix} \tag{1}$$

(2) Upper triangular matrix：

$$\begin{pmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2}$$

Theorem 2. Let A and B are lower triangular matrices. Then
1. $A+B$ is a lower triangular matrix;
2. rA is lower triangular matrix;
3. AB is lower triangular matrix;
4. A^T is upper triangular matrix;

Proof. 1., 2. and 4. are obvious. We will focus on the proof of statement 3
3. For $i<j$, we have

$$[AB]_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{j-1} a_{ik}b_{kj} + \sum_{k=j}^{n} a_{ik}b_{kj} = 0$$

□

3 The LU Factorization (LU分解)

3.1 Definition

Definition 3 (LU Factorization of a matrix A). *If a $m \times n$ matrix A is factorized as follows:*

$$A = LU, \tag{3}$$

where L is a $m \times m$ lower triangular matrix with 1 on the diagonal and U is a $m \times n$ echelon form of A, then 3 is the LU factorization of matrix A.

3.2 Why is LU factorization useful?

To answer the question in the title of this section, we first investigate the following two cases first for solving a matrix equation $A\vec{x} = b$.

A simple case: Now let A be L which a $m \times m$ lower triangular matrix with 1 on the diagonal (现在我们假设L是$m \times m$的下三角阵，其中对角线元素是1). Consider the matrix equation $L\vec{x} = \vec{b}$. This equation has a unique solution and can be solved readily. The solution $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ can be computed as follows:

$$\begin{aligned} x_1 &= b_1 \\ x_2 &= b_2 - l_{21}x_1 \\ &\cdots \\ x_m &= b_m - (l_{m1}x_1 + \cdots + l_{m,m-1}x_{m-1}). \end{aligned}$$

where l_{ij} is the entry value of L at (i, j).

A more complex case: Moreover if we assume that (注意这里我们先假设A能分解成下面的形式，我们后面再回头看怎么做到) matrix for any matrix equation $A\vec{x} = \vec{b}$, $A \in \mathbb{R}^{m \times n}$, can be factorized by LU factorization: $A = LU$. Denote $\vec{y} = U\vec{x}$. Then, the matrix equation is equivalent to the following pair of equations:

$$L\vec{y} = \vec{b} \tag{4}$$

$$U\vec{x} = \vec{y} \tag{5}$$

Recall the first simple case, solving equation $L\vec{y} = \vec{b}$ is easy. Since U is the echelon form of A, so it is also very easy to solve the equation $U\vec{x} = \vec{y}$

3.3 How to perform LU factorization?

From the last section, if LU factorization is applicable to matrix A, we see that solving $A\vec{x} = \vec{b}$ boils down to computing a factorization $(L,\ U)$, which is called an LU factorization of A. If LU factorization is applicable to matrix A, we now demonstrate how to compute such a factorization.

Recall that we can transform any matrix $m \times n$ matrix A into an equivalent matrix U in echelon form using a series of elementary row operations. Let us think about this procedure more carefully. We observe that the elementary row operations used in the procedure are of TYPE 1 ($r_i \leftrightarrow r_j$) and TYPE 3 ($r_i := r_i + \lambda r_j$ where $i > j$) only.

Suppose that we DO NOT need to use any TYPE 1 operations in this procedure(注意：即这里我们假设我们只用行倍加变换，row replacement). Thus $U = E_l E_{l-1} \cdots E_1 A$, where for any k, E_k is a TYPE 3 elementary matrix $E_m(i, j; \lambda)$ with $i > j$. As we have mentioned in the preceding examples, for all k, E_k are lower triangular matrices, so is $(E_l \cdots E_1)^{-1}$ by THEOREM 3. Now let $L = (E_l \cdots E_1)^{-1} = E_1^{-1} \cdots E_l^{-1}$, we have

$$A = (E_l \cdots E_1)^{-1} U = (E_1^{-1} \cdots E_l^{-1}) U = LU.$$

In practice, $L = E_1^{-1} \cdots E_l^{-1}$ can be computed simultaneously as transforming A into U:
1. Initially let $L := I_m$;
2. If $(r_i := r_i + \lambda r_j)$ is applied in the transformation, then replace the (i, j)–th entry of L by $-\lambda$, i.e. $l_{ij} := -\lambda$.

Remark: Again, we emphasize that NOT ALL matrices can be LU–factorised. LU–factorising a matrix A is possible iff A can be transformed into echelon form using only TYPE 3 elementary row operations.

Example: Textbook P145.

VIEW OF DELFT, *by Vermeer*