Lecture Notes on LU Decomposition

(Bruno Nachtergaele, 01/24/00)

1 Introduction

Performing a complete reduction to Reduced Row Echelon form is not the most efficient algorithm for solving linear systems. As we will explain below, complete reduction is doing a bit more than necessary. Elementary row operations are, however, at the heart of a large number of algorihtms, especially the ones most frequently used in numerical (i.e., computer) applications. We will see how they can be used to write a matrix as a product of two or more matrices with a simple structure. This is a common procedure in algebra. E.g., factorizing a polynomial (which is equivalent to finding its zeroes, or finding the solutions of a polynomial equation) is also writing an arbitrary polynomial as a product of simple ones.The prime factorization of an integer is another example. Later, we will encounter more applications of factorizations of matrices, but for now our focus is on solving linear systems.

2 Upper and Lower Triangular Matrices

We start with a bit of terminolgy. A square matrix $A = (a_{ij})$ is called *upper triangular* if $a_{ij} = 0$ for all i, j such that $i > j$. I.e., A is of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Similarly, $A = (a_{ij})$ is called *lower triangular* if $a_{ij} = 0$ for all i, j such that $i < j$. I.e., A is of the form

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

A is called a *diagonal* matrix if $a_{ij} = 0$ if $i \neq j$. This is the same as saying that A is both upper and lower triangular, and the only nonvanishing entries of A are on the diagonal, i.e., for $i = j$.

Before we apply these notions to linear systems, we derive some useful algebraic properties of triangular matrices. All matrices in the following are $n \times n$ unless stated otherwise.

Theorem 2.1
i) A is upper triangular if and only if A^T is lower triangular.
ii) If A and B are upper triangular, and $c \in \mathbb{R}$, then
a) cA is upper triangular,
b) $A + B$ is upper triangular, and
b) AB is upper triangular.
iii) All properties of ii) hold with upper triangular *replaced by* lower triangular.

Proof: The proof of i follows directly from the definitions of upper and lower triangular and the transpose of a matrix. The proofs of parts a and b of ii are easy. To prove c, we start from the definition of the matrix product. If $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times n$ matrices, then AB is the $n \times n$ matrix with entries given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

As A and B are assumed to be upper triangular we have that $a_{ik} = 0$ if $i > k$, and $b_{kj} = 0$ of $k > j$. In order to have $a_{ik} b_{kj} \neq 0$, one needs $a_{ik} \neq 0$, implying $i \leq k$, and $b_{kj} \neq 0$, implying $k \leq j$. These two conditions can be satisfied simultaneously only of $i \leq j$. Therefore $(AB)_{ij} = 0$ if $i > j$. Hence, AB is upper triangular.

The proof of *iii* is most easily obtained by considering *ii* with A replaced by A^T and B by B^T, and then using *i*. ∎

As a side remark we mention that the properties *ii.a-b* show that the upper triangular matrices form a subspace of the vector space of all $n \times n$ matrices. With the additional property *ii.c* we have that the upper triangular matrices form a subalgebra of the algebra of all $n \times n$ matrices.

3 Back and Forward Substitution

Consider an $n \times n$ linear system of the form

$$Ax = b$$

with A an upper triangular matrix. To solve a linear system like that it is not necessary to reduce it to Reduced Row Echelon form. Observe that the last equation contains only the unknown x_n and has solution

$$x_n = \frac{b_n}{a_{nn}}$$

if $a_{nn} \neq 0$. If $a_{nn} = 0$, we must distinguish the cases where $b_n = 0$ and $b_n \neq 0$. For concreteness, let is assume that all the diagonal elements of A are nonzero. Next, we substitute the solution for x_n into the second to last equation, which then contains only the unknown x_{n-1}, and which therefore can be solved immediately as well:

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}}.$$

The solutions for x_n and x_{n-1} can now be substituted in the equation # $n-2$, and so on until the complete solution has been found. This process is called *back substitution*. Note that as an intermediate step in our algorithm for reduction to RRE form, we obtain an upper triangular matrix that is row equivalent to A. Back substitution allows one to stop the reduction at that point and solve the linear system.

A similar procedure works when A is lower triangular. In that case the first equation contains only x_1 and can be solved immediately:

$$x_1 = \frac{b_1}{a_{11}}$$

Again, we assume that the diagonal entries of A are nonzero. By substituting the solution for x_1 into the second equation, we get an equation with only x_2 as unknown. The solution is

$$x_2 = \frac{b_2 - a_{21}x_1}{a_{22}}.$$

This process can be continued, now in the *forward* direction, until finally one finds

$$x_n = \frac{b_n - \sum_{k=1}^{n-1} a_{nk}x_k}{a_{nn}}.$$

This procedure is called *forward substitution.*

Next, we consider a linear system $Ax = b$, where now we assume that we know a lower triangular matrix L, and an upper triangular matrix U, such that $A = LU$. This is more general as such A need not be either upper or lower triangular, while both cases are clearly included. Again, we can solve the system without any reduction to RRE form. Define $y = Ux$, where x is the as yet unknown solution of $Ax = b$. Then, y satisfies

$$Ly = b.$$

As L is lower triangular, we can solve for y by forward substitution (as long as the diagonal entries of L are all nonzero). When we have obtained the solution for y we remember that y was defined by

$$Ux = y,$$

with U upper triangular, which means that we can find x from y by back substitution (again under the assumption that all diagonal entries of U are nonzero).

In conclusion, to solve $Ax = b$, it is suffcient to find L and U, which are lower and upper triangular matrices respectively, such that $A = LU$. If the diagonal entries of L and U are all nonzero, the solution of the linear system is unique and can be found by simple foward and back substitution.

The simple procedures of back and forward substitution can also be regarded as an straightforward way to compute the inverses of lower and upper triangular matrices. The only condition is that their diagonal entries are all non-zero. The latter is indeed a necessary and sufficient condition for a triangular matrix to be non-singular, as we will see later. Once the inverse of

L and U have been obtained the inverse of A itself can be found by the rule for the inverse of a product:

$$A^{-1} = (LU)^{-1} = U^{-1}L^{-1}.$$

4 LU factorization

From the foregoing it is clear that knowing a factorization of a matrix A as a product of a lower and an upper triangular matrix, i.e., $A = LU$, is indeed very useful. Such a factorization is called LU factorization or LU decomposition. We will later prove that such a factorization, with L and U satisfying the condition that they have no zero elements on their diagonals, is equivalent to A, or a permutation of A, being non-singular. For simplicity, we will now explain how such an LU factorization can be obtained in the most frequent case where the reduction of A to RRE form does not require any exchanges of rows. Row-exchanges can be included with a small additional effort.

Not surprisingly, the factorization procedure itself involves elementary row operations. Suppose $A = (a_{ij})$ is an $n \times n$ matrix with $a_{11} \neq 0$. The first step in our algorithm for reduction to RRE form is then the ERO which replaces the 2nd row by the 2nd minus the first multiplied by a_{21}/a_{11}. This leaves us with the matrix

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} - a_{12}a_{21}/a_{11} & a_{23} - a_{13}a_{21}/a_{11} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \end{pmatrix}.$$

The new trick comes now. The ERO we just performed on A can be viewed as the multiplication of A from the left with the lower triangular matrix E_{12} defined by

$$E_{12} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a_{21}/a_{11} & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

You can easily verify that

$$A_1 = E_{12}A$$

Next, the ERO which replaces the 3rd row by the 3rd row minus a multiple of the first such that the first element of the 3rd row vanishes, can be obtained by left multiplication of A_1 by

$$E_{13} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ -a_{31}/a_{11} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

which is also lower triangular. $A_2 := E_{13}A_1$ then has a 0 in the position 13. It is now clear how to define matrices E_{1j} such that

$$A_{n-1} = E_{1n}E_{1,n-1}\cdots E_{13}E_{12}A$$

is the result of $n-1$ ERO's applied to A, and has all zeroes in the first column except for the top element. Next, we proceed with more ERO's to make the elements of the 2nd column under the diagonal vanish. This can be achieved by left multiplication with lower triangular matrices of the form

$$E_{2j} = \begin{pmatrix} 1 & 0 & 0 & \cdots & & & 0 \\ 0 & 1 & 0 & \cdots & & & 0 \\ 0 & 0 & 1 & \cdots & & & 0 \\ \vdots & \vdots & \vdots & \ddots & & & \vdots \\ 0 & -a_{j2}/a_{22}^{(1)} & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & & & 1 \end{pmatrix},$$

where $a_{22}^{(1)}$ is the 22 element of A_1, i.e.,

$$a_{22}^{(1)} = a_{22} - a_{12}a_{21}/a_{11}.$$

We assume again that $a_{22}^{(1)} \neq 0$. One can continue in the same way with ERO's to make the elements under the diagonal in the third column vanish, and the 4th, and so on until eventually we obtain an upper triangular matrix that is row equivalent to A. Denote this matrix by U. The above procedure then shows that

$$U = E_{n-1,n}E_{n-2,n-1}E_{n-2,n}\cdots E_{23}E_{1n}\cdots E_{12}A, \tag{4.1}$$

where all the E_{ij} are lower triangular matrix matrices.

The next step is to notice that the inverses of the E_{ij} are lower triangular matrices of the same form. Simply change the sign of the only non-vanishing off-diagonal element. E.g.,

$$E_{12}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21}/a_{11} & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

By using the inverses of all E_{ij} we can solve back for A from (4.1) and obtain

$$A = LU$$

with

$$L = E_{12}^{-1} E_{13}^{-1} \cdots E_{1n}^{-1} E_{23}^{-1} \cdots E_{n-1,n}^{-1}.$$

From the theorem in Section 2 it follows that L is lower triangular as it is a product of lower triangular matrices. We have thus obtained a factorization of A as a product of lower triangular and an upper triangular matrix.

When this procedure is implemented in a computer program, one does not actually perform this large number of matrix multiplications with matrices that are mostly zeroes anyway. Each time only one row of the matrix is modified and this is easily programmed directly. By writing the procedure as a matrix product, however, we were able to see that it indeed leads to an LU factorization of A.

A procedure to obtain the LU factorization manually is described in Section 9.3 of Kolman.