MATH-UA 140 - Linear Algebra

Lecture 7: LU Factorization

This lecture is short, but covers a topic which is very important in applied mathematics and computational science in general. The main idea is to provide a useful and convenient way to represent Gaussian elimination in matrix form. As we will see, this is tantamount to finding a particular factorization for a matrix A.

1) A simple example

Let us start with a fairly simple example.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

We already saw how to transform A into an upper triangular matrix, by elimination:

$$E_{21} A = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} = U$$

Now, if we want to return back to A from U, we reverse the elimination step:

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} U = E_{21}^{-1} U = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} = A$$

This is the factorization we are interested in in this lecture. we have $A = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$

The first matrix on the right hand side is <u>lower triangular</u>. We call it L. U is upper triangular, and we have

$A = LU$

This is called the LU factorization of A. Let us now see how/why it works in general.

2) General idea of LU factorization

In the following, we focus on the case where Gaussian elimination does not involve row exchanges. This case is simpler, and also the most common. We will include row exchanges later in this course.

As we said, LU factorization is a nice way of representing Gaussian elimination with matrices. Let us remind ourselves how Gaussian elimination works in matrix form, say for a 3-by-3 matrix.

The key idea of Gaussian elimination is to put 0s below the pivots of A in each columns. This is done by multiplying A with elimination matrices. For the first column of A, we thus first put a zero in the 21 entry by multiplying by the elimination matrix E_{21}, and then put a zero in the 31 entry by multiplying by E_{31}. These steps are:

$$E_{31} E_{21} A$$

Then, we go to the second column of the resulting matrix, and put a zero in the 32 entry by multiplying by E_{32}. At that point, we have an upper triangular matrix:

$$E_{32} E_{31} E_{21} A = U \qquad (*)$$

with $E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -l_{31} & 0 & 1 \end{bmatrix}$, $E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -l_{32} & 1 \end{bmatrix}$

and l_{21}, l_{31}, and l_{32} are the multipliers.
We know that E_{21}, E_{31}, and E_{32} have inverses (see previous lecture) so $E_{32} E_{31} E_{21}$ is invertible, and we can rewrite $(*)$ as:

$$A = (E_{32} E_{31} E_{21})^{-1} U$$

$$\Leftrightarrow \quad A = (E_{21}^{-1} E_{31}^{-1} E_{32}^{-1}) U$$

with $E_{21}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_{31}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_{31} & 0 & 1 \end{bmatrix}$, and $E_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{bmatrix}$

Now, it is important to observe that the product of the elimination matrices is a lower triangular matrix. To see this, think about what elimination does. For example, E_{21}^{-1} applied to E_{31}^{-1} reproduces E_{31}^{-1}, except that the 2nd row of E_{31}^{-1} is replaced with the 2nd row of E_{31}^{-1} plus l_{21} times its first row. Since the only nonzero entry of the first row is in the first column, the resulting matrix remains lower triangular. The zero in the entry 23 remains a zero.

So $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$ is a lower triangular matrix we call L, and we can write

$$A = (E_{21}^{-1}E_{31}^{-1}E_{32}^{-1})U = LU$$

This is the expected LU factorization of A.

3) Why is this convenient?

We saw in lectures 3 and 4 that the "direct" representation of Gaussian elimination, given by (*), is satisfying for solving systems of equations. Why bother with the "inverse" version and LU factorization? To find the answer, let us first evaluate $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_{31} & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$$

L has all 1's on the diagonal, and the entries ij below the diagonal are the multipliers l_{ij}.

This is a general result, which allows us to construct L quickly as we apply Gaussian elimination to a matrix A.

Example: Returning to the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{bmatrix}$ we have looked at several times to illustrate Gaussian elimination, we can use the elimination steps seen in Lecture 5 to immediately write:

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 1 & 3 \\ 0 & \frac{3}{2} & -\frac{5}{2} \\ 0 & 0 & \frac{4}{3} \end{bmatrix}}_{U}$$

So we see that LU factorization is a way to store all the key pieces of information that were involved in the Gaussian elimination process: the multipliers in L and the pivots in U.

4) Why is this convenient for solving systems?

In many applications in science, we have to solve systems of the form $A\vec{x} = \vec{b}$ for several right-hand sides $\vec{b}_1, \vec{b}_2, \ldots$
LU factorization does not really save time if one has to solve

the system once, for one $\vec{b}_1$ but is convenient if we want to solve the system repeatedly, for $\vec{b}_1, \vec{b}_2, \ldots \vec{b}_n$

Indeed, the factorization $A = LU$ does not depend on the right-hand side of the system. So we can use it to solve the system no matter what $\vec{b}$ is, and once the factorization is known, it is easier to use it than use A directly. This is why:

Let us say we want to solve $A\vec{x} = \vec{b} \Leftrightarrow LU\vec{x} = \vec{b}$. The idea is to solve two triangular systems:

(1) $L\vec{c} = \vec{b}$ for some undetermined $\vec{c}$
(2) $U\vec{x} = \vec{c}$

The combination of the two systems is indeed equivalent to the starting system:

$$\vec{b} = L\vec{c} = LU\vec{x} = A\vec{x}$$

but the point is that because (1) and (2) are triangular systems, they can be easily solved by forward (for (1)) and backward (for (2)) substitution.

QUESTION: Solve the system $\begin{cases} 2x + y = 3 \\ x + 3y = 5 \end{cases}$

using LU factorization

5) A final comment on notation

Some people dislike the lack of "symmetry" between L and U: L has 1's on the diagonal, while U has the pivots on the diagonal. LU can be "symmetrized" by factoring U as follows:

$$U = \begin{bmatrix} d_1 & u_{12} & \cdots & u_{1n} \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & u_{n-1\,n} \\ 0 & \cdots & 0 & d_n \end{bmatrix} = \underbrace{\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{bmatrix}}_{\text{D, diagonal matrix with pivots of U}} \underbrace{\begin{bmatrix} 1 & \frac{u_{12}}{d_1} & \cdots & \frac{u_{1n}}{d_1} \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \frac{u_{n-1\,n}}{d_{n-1}} \\ 0 & \cdots & 0 & 1 \end{bmatrix}}_{\text{New upper triangular matrix U' with 1's on the diagonal}}$$

In that case, the triangular factorization of A may be written as

$$A = LDU'$$

and is sometimes called the symmetric LU factorization of A.