The LU decomposition

These lecture notes cover a very important algorithm in applied linear algebra, the *LU decomposition*. If an n by n matrix A is non-singular it can be 'factored' into two special matrices, a lower triangular matrix L and an upper triangular matrix U:

$$A = LU$$

The matrix U turns out to be the upper-triangular matrix that results after we do Gauss elimination to eliminate all of the entries in A below the diagonal. In these notes we will work out what appear in the matrix L.

Implementing elementary operations as matrix multiplications

The Gaussian elimination algorithm consisted of a series of *elementary operations*. These operations included switching two rows and substracting a multiple of row i from row j. Our first significant observation is that both of these elementary operations can be implemented as matrix multiplications.

Consider the example of switching rows i and j of an n by n matrix A. If we form the *permutation matrix* $P^{(i,j)}$ that results from switching rows i and j in the n by n identity matrix then we can construct the matrix $A^{(1)}$ that results from switching rows i and j in A by doing

$$A^{(1)} = P^{(i,j)}A$$

How do we construct such a permutation matrix? We begin by assuming that the permutation matrix in question does the row swap on every matrix we apply it to. By making a clever choice for the matrix to apply the permutation to, we can easily see what that permutation matrix should be.

The key is to apply the permutation to an identity matrix. For example, consider the 4 by 4 permutation matrix $P^{(2,4)}$ that swaps rows 2 and 4 of any 4 by 4 matrix A. If we set $A = I_4$, the 4 by 4 identity matrix, we get

$$P^{(2,4)} = P^{(2,4)}\, I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Here we have used the fact that any matrix times an identity matrix is that matrix itself.

Let us check this with an example. Suppose we have

$$A = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

Suppose we wanted to switch rows 2 and 4. We form the permutation matrix

$$P^{(2,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

and then check that

$$P^{(2,4)}A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & -3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & 1 & -3 & 2 \\ -1 & 2 & 1 & -1 \\ 1 & -1 & 2 & 0 \end{bmatrix}$$

This does have the desired effect.

Next, suppose we wanted to form a combination of rows. For example, if we were doing Gaussian elimination we would want to eliminate the element in position 2,1 by forming a multiplier $m_{2,1} = a_{2,1}/a_{1,1} = 2$ and replacing row 2 with ((row 2) - $m_{2,1}$(row 1)). Once again, it turns out that this operation can be implemented as a multiplication by a special matrix $M^{(2,1)}$.

Using the same argument as above, we can find that elimination matrix:

$$M^{(2,1)} = M^{(2,1)}\, I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{2,1} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can check that this matrix does the right thing when applied to A:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

Full Gauss Elimination with Matrices

Now that we know how to construct elimination matrices, we can use them to do a complete Gauss elimination.

The three matrices needed to eliminate the elements in column one are

$$M^{(2,1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{2,1} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M^{(3,1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -m_{3,1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M^{(4,1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -m_{4,1} & 0 & 0 & 1 \end{bmatrix}$$

Here is what happens when we construct these elemination matrices and apply them in sequence to the original A:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 2 & 3/2 & 1/2 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

In general, we will be applying successive multiplications by elementary operation matrices to $A^{(1)} = A$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -m_{4,1} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -m_{3,1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{2,1} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{(1)} = A^{(2)}$$

An interesting thing happens when we take the product of the three elementary operation matrices on the left:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -m_{4,1} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -m_{3,1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{2,1} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -m_{2,1} & 1 & 0 & 0 \\ -m_{3,1} & 0 & 1 & 0 \\ -m_{4,1} & 0 & 0 & 1 \end{bmatrix}$$

This product has a particularly pleasing and simple form. It looks like the identity matrix with the elements below the diagonal in the first column replaced by $-m_{2,1}$, $-m_{3,1}$, and $-m_{4,1}$. The author calls this a *column elimination matrix* $M^{(1)}$ and writes

$$M^{(1)}A^{(1)} = A^{(2)}$$

Next, we would construct a column elimination matrix $M^{(2)}$ that contains the multipliers needed to eliminate the elements below the diagonal in the second column:

$$M^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$M^{(2)}\,A^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 2 & 3/2 & 1/2 \\ 0 & 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 0 & 9/2 & -5/2 \\ 0 & 0 & 9/2 & 1/2 \end{bmatrix} = A^{(3)}$$

We now have

$$M^{(2)}M^{(1)}A^{(1)} = A^{(3)}$$

Finally, we construct a column elimination matrix $M^{(3)}$ to eliminate down the third column.

$$M^{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

$$M^{(3)}\,A^{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 0 & 9/2 & -5/2 \\ 0 & 0 & 9/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 0 & 9/2 & -5/2 \\ 0 & 0 & 0 & 3 \end{bmatrix} = A^{(4)}$$

This leads to

$$M^{(3)}M^{(2)}M^{(1)}A = A^{(4)} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 0 & 9/2 & -5/2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Note that $A^{(4)}$ is an upper triangular matrix. We will call this U.

$$M^{(3)}M^{(2)}M^{(1)}A = U \tag{1}$$

Recall that our goal originally was to construct

$$A = L\,U$$

It looks like we can construct L by multiplying both sides of (1) by the inverses of the M matrices in the appropriate sequence:

$$A = \left(M^{(1)}\right)^{-1} \left(M^{(2)}\right)^{-1} \left(M^{(3)}\right)^{-1} U$$

We now have two remaining problems. The first is computing the inverses of the column elimination matrices. The second is multiplying those inverses together and hoping that the result is a lower triangular matrix we can call L.

Our first concern is easy to overcome. What is special about column elimination matrices is that they represent operations that are both easy to understand *and easy to reverse*. A typical column elimination matrix seeks to eliminate items down column k by combining rows with multiples of row k. To reverse the process, we just combine the rows with multiples of row k again, but this time we use multipliers that are the negatives of the originals. This neatly reverses the original operations. Here is an example using $M^{(1)}$ from above.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 2 & 3/2 & 1/2 \\ 0 & 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ -1 & -1 & 1 & -3 \\ 1 & 2 & 2 & 2 \\ 0 & 1 & 3 & 2 \end{bmatrix} = A^{(1)}$$

In other words,

$$(M^{(1)})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Using this trick we can quickly construct the inverses of $M^{(2)}$ and $M^{(3)}$.

$$(M^{(2)})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$(M^{(3)})^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Something truly amazing happens when we form the products of the three inverse matrices.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & -2 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix}$$

Notice that the resulting matrix looks like a matrix formed by simply combining the columns of the three column elimination matrices! As a further bonus, this matrix is also lower triangular, and is the L matrix we wanted.

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & -2 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix}$$

This can now be summarized by the following remarkable theorem.

Theorem (LU decomposition) If A is a non-singular matrix that can be reduced to an upper triangular matrix U by Gauss elimination without requiring any row swaps and using a set of multipliers $m_{j,i}$, then the matrix L formed by placing $m_{j,i}$ in each of the below diagonal spaces of an n by n identity matrix satisfies

$$A = L\,U$$

We can now confirm that the L and U matrices we found above are correct:

$$L\,U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ -1/2 & -2 & 1 & 0 \\ 0 & -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & -1 & 3/2 & -3/2 \\ 0 & 0 & 9/2 & -5/2 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix} = A$$

Dealing with row swaps

Note that in the discussion above we have been avoiding any discussion of row swaps. In ordinary Gaussian elimination we occasionally have to do row swaps. For example, consider the matrix

$$A = \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 1 & 1 & -1 \\ 2 & 0 & 1 & 3 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

After the first round of elimination which eliminates down column 1 we come to

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & 0 \\ -1 & 1 & 1 & -1 \\ 2 & 0 & 1 & 3 \\ 0 & 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 2 & -3 & 3 \\ 0 & 1 & 3 & 2 \end{bmatrix} = A^{(2)}$$

This $A^{(2)}$ is going to require a row swap to get the 0 out of the diagonal position in row 2. For example, we can proceed by swapping rows 2 and 4.

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 2 & -3 & 3 \\ 0 & 0 & 3 & -1 \end{bmatrix}$$

We then continue by constructing appropriate $M^{(2)}$ and $M^{(3)}$ matrices.

How will the row swap affect the LU decomposition? Here is a clever trick designed to work around the problem caused by the row swap. What we do is carry out Gaussian elimination as usual collecting the multipliers as we go along and also making a record of all the row swaps that were needed. We then replay the elimination a second time, except this time around we start by applying all of the row swaps we accumulated to A before doing any elimination steps. The result will be a matrix that can be reduced *using exactly the set of multipliers we discovered in the first round*, and this time without the need for further row swaps.

Since we have been using matrices and matrix multiplication to represent our elimination steps, we should be consistent and develop a permutation matrix that does the preliminary row swaps. This matrix P is easy to develop: we just do the exact same set of row swaps on an identity matrix and we have P.

We can now summarize our results in a theorem.

Theorem (LUP decomposition) Let P be a permutation matrix that converts the n by n nonsingular matrix A to a form that can be reduced by Gaussian elimination to an upper triangular U without requiring any row swaps. If the elimination uses a set of multipliers $m_{j,i}$, then the matrix L formed by placing $m_{j,i}$ in each of the below diagonal spaces of an n by n identity matrix satisfies

$$PA = LU$$

Note Most frequently in applications we will want to use this result to factor the original A into three matrices. To do this, it is more useful to rewrite the last result as

$$A = P^{-1}LU$$

A further simplification comes from a nice fact about permutation matrices. The inverse of a permutation matrix is just the transpose of the matrix. Thus, we have

$$A = P^t LU$$

This provides the factorization of A we wanted.