Chapter 10. Solving Large Systems

Note. In this chapter, Fraleigh and Beauregard consider numerical algorithms, the results of computational techniques, and “consideration of time” (and round-off error). In Section 10.1, a consideration of “flops” (i.e., a single arithmetic operation) is given. We skip this part of Chapter 10 and go straight into the more theoretical topic of $LU$ factorization.

10.2 The $LU$-Factorization

Note. In this section, we first consider a system of equations of the form $A\vec{x} = \vec{b}$ with a unique solution which can be solved using Gauss-Jordan elimination but without using the row operation of interchanging rows. We will consider the more general case which allows interchanging rows later in the section.

Note. If coefficient matrix $A$ can be put in row echelon form without row interchanges (so the only needed elementary row operation is adding a multiple of of one row to another; there is no need to scale any rows since we do not need the pivots to be 1), then there is an upper triangular matrix $U$ and a sequence of elementary matrices $E_i$ such that $E_hE_{h-1}\cdots E_2E_1A = U$. This is the key observation to showing the existence of an $LU$-factorization of such a matrix.
Theorem 10.A. If $A$ is an $n \times n$ matrix which can be put in row echelon form without interchanging rows then there is a lower triangular matrix $L$ and an upper triangular matrix $U$ such that $A = LU$.

Example 1. The proof of Theorem 10.A gives the algorithm by which the $LU$-factorization of an appropriate matrix can be found. We row reduce $A$ to row echelon form $U$ and, as each elementary row operation is performed in the reduction of $A$, perform the inverse of that operation to a matrix starting with the identity matrix. This will build up matrix $L$ from the identity by filling in the entries in the first column from top to bottom (this is backwards from the order given in the proof). With

$$A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 8 & 4 \\ -1 & 3 & 4 \end{bmatrix} \quad \text{and} \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{array}{c}
A \\
R_2 \rightarrow R_2 - 2R_1 \\
R_3 \rightarrow R_3 + R_1 \\
R_3 \rightarrow R_3 - 3R_2 \\
\end{array} \quad \begin{array}{c}
I \\
R_2 \rightarrow R_2 + 2R_1 \\
R_3 \rightarrow R_3 - R_1 \\
R_3 \rightarrow R_3 + 3R_1 \\
\end{array}$$
So we have
\[
U = \begin{bmatrix}
1 & 3 & -1 \\
0 & 2 & 6 \\
0 & 0 & -15
\end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix}
1 & 0 & 0 \\
2 & 1 & 0 \\
-1 & 3 & 1
\end{bmatrix}.
\]
Notice that in fact \( A = LU \).

**Example 2.** Solve the system
\[
\begin{align*}
x_1 + 3x_2 - x_3 &= -4 \\
2x_1 + 8x_2 + 4x_3 &= 2 \\
-x_1 + 3x_2 + 4x_3 &= 4.
\end{align*}
\]
Notice that the coefficient matrix is the same as given in example 1. Also, matrix \( L \) records the row operations used in the reduction of \( A \) (but indirectly).

**Solution.** We minimize the number of arithmetic operations using the information from Example 1:
\[
\vec{b} = \begin{bmatrix}
-4 \\
2 \\
4
\end{bmatrix} \quad R_2 \rightarrow R_2 - 2R_1 \quad \begin{bmatrix}
-4 \\
10 \\
4
\end{bmatrix} \quad R_3 \rightarrow R_3 + R_1 \quad \begin{bmatrix}
-4 \\
10 \\
0
\end{bmatrix} \quad R_3 \rightarrow R_3 - 3R_2 \quad \begin{bmatrix}
-4 \\
10 \\
30
\end{bmatrix} = \vec{c}.
\]
Then the augmented matrix \([A \mid \vec{b}]\) is row equivalent to \([U \mid \vec{c}]\). Since \([U \mid \vec{c}]\) has an upper triangular coefficient matrix then it is easy to solve with back substitution:
\[
[U \mid \vec{c}] = \begin{bmatrix}
1 & 3 & -1 & -4 \\
0 & 2 & 6 & 10 \\
0 & 0 & -15 & -30
\end{bmatrix} \quad \text{implies} \quad \begin{align*}
x_1 + 3x_2 - x_3 &= -4 \\
2x_2 + 6x_3 &= 10 \\
-15x_3 &= -30.
\end{align*}
\]
So \( x_3 = 2 \), \( x_2 = -1 \), and \( x_1 = 1 \).
Definition. Let $A$ be an $n \times n$ matrix which can be put in row echelon form without interchanging rows. If $A = LU$ where the diagonal entries of $L$ are all $q$, then the combined $L \setminus U$ display for $A$ is an $n \times n$ matrix with the same diagonal entries as matrix $U$, entries above the diagonal the same as the corresponding entries of $U$, and the entries below the diagonal the same as the corresponding entries of $L$.

Note. The $L \setminus U$ display for $A$ is equal to $L + U - I$. In Example 1, we have

$$U = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 6 \\ 0 & 0 & -15 \end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix}.$$ 

So the $L \setminus U$ display is

$$\begin{bmatrix} 1 & 3 & -1 \\ 2 & 2 & 6 \\ -1 & 3 & -15 \end{bmatrix}.$$ 

Note. When matrix $A$ is of the proper type to have an $LU$-factorization, the factorization is not unique. If $A = LU$ then $A = (rL)((1/r)U)$ for any $r \neq 0$. The technique given above produces a matrix $L$ with entries of 1 along the diagonal, but this is not necessary.

Note. If $A = LU$ where the diagonal entries of $L$ are all 1, then we can multiply row $i$ of matrix $U$ by $1/u_{ii}$ and produce an upper triangular matrix $U^*$ with diagonal entries of a. We then create $n \times n$ diagonal matrix $D$ with $d_{ii} = u_{ii}$. This gives $U = DU^*$. We then have a factorization of $A$ as $A = LDU^*$ where the diagonal entries of $L$ and $U^*$ are all 1. The next theorem tells us that when such a factorization of a matrix exists, it is unique.
Theorem 10.1. Unique Factorization.

Let $A$ be an $n \times n$ matrix. When a factorization $A = LDU$ exists where

1. $L$ is lower triangular with all main diagonal entries 1,

2. $U$ is upper triangular with all main diagonal entries 1, and

3. $D$ is a diagonal matrix with all main diagonal entries nonzero,

it is unique.

Definition. An $n \times n$ matrix which is a product of elementary matrices which represents row interchanges is a permutation matrix.

Note. By using a permutation matrix, we can finally address the $LU$-factorization of a matrix which cannot be put in row echelon form without the use of row interchanges.

Theorem 10.2. $LU$-Factorization.

Let $A$ be an invertible square matrix. Then there exists a permutation matrix $P$, a lower triangular matrix $L$, and an upper triangular matrix $U$ such that $PA = LU$. 
Example 7. Consider \[
\begin{bmatrix}
1 & 3 & 2 \\
-2 & -6 & 1 \\
2 & 5 & 7
\end{bmatrix}
\] We have
\[
A \xrightarrow{R_2 \rightarrow R_2 + 2R_1} \begin{bmatrix}
1 & 3 & 2 \\
0 & 0 & 5 \\
2 & 5 & 7
\end{bmatrix} \xrightarrow{R_3 \rightarrow R_3 - 2R_1} \begin{bmatrix}
1 & 3 & 2 \\
0 & 0 & 5 \\
0 & -1 & 3
\end{bmatrix}.
\]

So let \( P = \begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0
\end{bmatrix} \). Then \( U = \begin{bmatrix}
1 & 3 & 2 \\
0 & -1 & 3 \\
0 & 0 & 5
\end{bmatrix} \). The two operations on \( A \) which produce \( U \) give \( L = \begin{bmatrix}
1 & 0 & 0 \\
2 & 1 & 0 \\
-2 & 0 & 1
\end{bmatrix} \) (notice that the second and third rows of \( A \) and \( U \) involve a row interchange). We then have
\[
PA = \begin{bmatrix}
1 & 3 & 2 \\
2 & 5 & 7 \\
-2 & -6 & 1
\end{bmatrix} \begin{bmatrix}
1 & 0 & 0 \\
2 & 1 & 0 \\
-2 & 0 & 1
\end{bmatrix} \begin{bmatrix}
1 & 3 & 2 \\
0 & -1 & 3 \\
0 & 0 & 5
\end{bmatrix} = LU.
\]