2.4 Solving Linear Systems

\[
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0.25 & 1 & 0 \\
1 & 0 & 1
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0.5
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}
= \begin{bmatrix}
0.25 & 0.5 & 1 \\
1 & 0 & 0 \\
1 & 1 & 0
\end{bmatrix},
\]

and hence

\[
A = \begin{bmatrix}
1 & 2 & 2 \\
4 & 4 & 2 \\
4 & 6 & 4
\end{bmatrix}
= \begin{bmatrix}
0.25 & 0.5 & 1 \\
1 & 0 & 0 \\
1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
4 & 4 & 2 \\
0 & 2 & 2 \\
0 & 0 & 0.5
\end{bmatrix}
= LU.
\]

Note that \( L \) is not lower triangular, but it is triangular in the more general sense (it is a permutation of a lower triangular matrix). Alternatively, we can take

\[
P = P_2 P_1 =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
= \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{bmatrix},
\]

and

\[
L =
\begin{bmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
0.25 & 0.5 & 1
\end{bmatrix},
\]

so that

\[
P A = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
1 & 2 & 2 \\
4 & 4 & 2 \\
4 & 6 & 4
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
0.25 & 0.5 & 1
\end{bmatrix}
\begin{bmatrix}
4 & 4 & 2 \\
0 & 2 & 2 \\
0 & 0 & 0.5
\end{bmatrix}
= LU,
\]

where \( L \) now really is lower triangular but \( A \) is permuted.

The name “partial” pivoting comes from the fact that only the current column is searched for a suitable pivot. A more exhaustive pivoting strategy is \textit{complete pivoting}, in which the entire remaining unreduced submatrix is searched for the largest entry, which is then permuted into the diagonal pivot position. Note that this requires interchanging columns as well as rows, and hence it leads to a factorization of the form

\[
P AQ = LU,
\]

where \( L \) is unit lower triangular, \( U \) is upper triangular, and \( P \) and \( Q \) are permutation matrices that reorder the rows and columns, respectively, of \( A \). To solve the linear system \( A x = b \), we first solve the lower triangular system \( L y = Pb \) by forward-substitution, then the upper triangular system \( U z = y \) by back-substitution, and finally we permute the solution components to obtain \( x = Qz \). Although the numerical stability of complete pivoting is theoretically superior, it requires a much more expensive pivot search than partial pivoting. Because the numerical stability of partial pivoting is more than adequate in practice, it is almost universally used in solving general linear systems by Gaussian elimination.

Pivot selection depends on the magnitudes of individual matrix entries, so the particular choice obviously depends on the scaling of the matrix. A diagonal scaling of the matrix (recall Example 2.10) may result in a different sequence of pivots.
For example, any nonzero entry in a given column can be made the largest in magnitude simply by giving that row a sufficiently heavy weighting. This does not mean that an arbitrary pivot sequence is acceptable, however: a badly skewed scaling can result in an ill-conditioned system and a correspondingly inaccurate solution (see Example 2.20). A well-formulated problem should have appropriately commensurate units for measuring the unknown variables (column scaling), and a weighting of the individual equations that properly reflects their relative importance (row scaling). It should also account for the relative accuracy of the input data. Under these circumstances, the pivoting procedure will usually produce a solution that is as accurate as the problem warrants (see Section 2.3.4).

We saw in Section 2.3.5 that the relative residual for a computed solution satisfies the inequality

\[
\frac{\|r\|}{\|A\| \cdot \|\hat{x}\|} \leq \frac{\|E\|}{\|A\|},
\]

where \( E \) is the backward error in the matrix \( A \). But how large is \( \|E\| \) likely to be in practice? Wilkinson [417] showed that for LU factorization by Gaussian elimination, a bound of the form

\[
\frac{\|E\|}{\|A\|} \leq \rho \cdot n \cdot \epsilon_{\text{mach}}
\]

holds, where \( \rho \), called the *growth factor*, is the ratio of the largest entry of \( U \) to the largest entry of \( A \) in magnitude (technically, the growth factor depends on the largest entry produced at any stage of the factorization process, but this is typically the last, or \( U \)). Without pivoting, \( \rho \) can be arbitrarily large, and hence Gaussian elimination without pivoting is unstable, as we have already seen. With partial pivoting, the growth factor can still be as large as \( 2^{n-1} \) (since in the worst case the size of the entries can double at each stage of elimination), but such behavior is extremely rare. In practice, there is little or no growth in the size of the entries, so that

\[
\frac{\|E\|}{\|A\|} \approx n \cdot \epsilon_{\text{mach}}.
\]

This relation means that solving a linear system by Gaussian elimination with partial pivoting followed by back-substitution almost always yields a very small relative residual, regardless of how ill-conditioned the system may be. Thus, a small relative residual does not necessarily indicate that a computed solution is accurate unless the system is well-conditioned. Complete pivoting yields an even smaller growth factor, both in theory and in practice, but the additional margin of stability it provides is usually not worth the extra expense.

**Example 2.17 Small Residual.** Consider the linear system

\[
Ax = \begin{bmatrix} 0.913 & 0.659 \\ 0.457 & 0.330 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.254 \\ 0.127 \end{bmatrix} = b
\]

from Example 2.8. Using four-digit decimal arithmetic, Gaussian elimination yields