$PA = LU$ Factorization

Recall that "naive" or Classical Gaussian Elimination encounters two significant issues: one, that it cannot find a solution at all if it encounters a zero pivot, and two, that it can compute wildly inaccurate solutions if there are elements of very different magnitude, called swamping. For a nonsingular matrix A, both can be avoided by reintroducing row swaps to the algorithm.

The following strategy on choosing a pivot is called *partial pivoting*. There are also other pivoting strategies, primarily *complete* pivoting and *rook* pivoting, both of which require more work than partial pivoting.

Partial pivoting is: select the largest element (in magnitude) in the column to use as the pivot.

Question 1. Why does partial pivoting solve the issue of a zero pivot?

Question 2. Why does partial pivoting prevent swamping?

Example 1. Use Partial Pivoting to solve the system $A\mathbf{x} = \mathbf{b}$ for $A = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ in double precision.

Example 2. Solve the system $A\mathbf{x} = \mathbf{b}$ using partial pivoting for $A = \begin{bmatrix} 1 & 9 & 1 \\ -2 & 2 & 1 \\ 4 & 4 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 8 \\ 1 \\ 3 \end{bmatrix}$.

How do we store the row swaps to re-use the work, when solving multiple right-hand-sides, while preserving the triangular structure of L and U?

We use a *permutation matrix* to store the row exchanges.

Definition 3. A permutation matrix is an $n \times n$ matrix consisting of all zeros except for a single 1 in every row and column.

Question 4. Are all permutation matrices invertible?

The permutation matrix P goes on the left of A so that $PA = LU$.

Theorem 5. *Let P be the $n \times n$ permutation matrix formed by a particular set of row exchanges applied to the identity matrix. Then for any $n \times n$ matrix A, PA is the matrix obtained by applying exactly the same set of row exchanges to A.*

Question 6. What is the permutation matrix P when A has an $PA = LU$ factorization without row swaps?

As we go through the process of forming LU, we swap the rows in P and also in L. Start with $P = I$. First, we row swap 1 and 3.

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 4 & 4 & 1 \\ -2 & 2 & 1 \\ 1 & 9 & 1 \end{bmatrix}$$

Now $R_2 + \frac{1}{2}R_1$ and $R_3 - \frac{1}{4}R_1$ gives two entries in L, though they may now change position later.

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{4} & & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 4 & 4 & 1 \\ 0 & 4 & \frac{3}{2} \\ 0 & 8 & \frac{3}{4} \end{bmatrix}.$$

We now also need to swap rows 2 and 3. Swap the rows in P and all earlier relevant coefficients in L.

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ -\frac{1}{2} & & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 4 & 4 & 1 \\ 0 & 8 & \frac{3}{4} \\ 0 & 4 & \frac{3}{2} \end{bmatrix}.$$

Next, $R_3 - \frac{1}{2}R_2$ gives

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 4 & 4 & 1 \\ 0 & 8 & \frac{3}{4} \\ 0 & 0 & \frac{9}{8} \end{bmatrix}$$

In practice, L and U are stored as one matrix, with the subdiagonal part being L and the upper triangular part being U, as $\begin{bmatrix} 4 & 4 & 1 \\ \frac{1}{4} & 8 & \frac{3}{4} \\ -\frac{1}{2} & \frac{1}{2} & \frac{9}{8} \end{bmatrix}$.

In order to use the $PA = LU$ factorization, we do also need to use P before the back substitution to permute $\mathbf{b}$ to match. That is,

1. Solve $L\mathbf{y} = P\mathbf{b}$ for $\mathbf{y}$.
2. Solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$.

Example 3. Use the above $PA = LU$ factorization to solve $A\mathbf{x} = \begin{bmatrix} -32 \\ -24 \\ 20 \end{bmatrix}$.

Question 7. Classical Gaussian Elimination runs in approximately $\frac{2n^3}{3}$. How much more work is Gaussian Elimination with Partial Pivoting (GEPP)?

The largest number of elements we search to choose the pivot is $n-1$, which is $n-1$ comparisons. A row swap costs $2n$ assignments. So the total (roughly $3n$) is less than order n^3, and the overall approximate operation count remains $\frac{2n^3}{3}$. Nice!

Question 8. Is GEPP now stable?

Well, in practical terms it is, but in a rigorous sense, no. Think of this roughly in terms of the infinity norms of A, L, and U - a good way to measure if the elements of the three matrices are similar in magnitude. Partial pivoting guarantees that $||L||_\infty \leq 1$, so it won't have significantly larger elements.

So we want to know if $||U||_\infty = \mathcal{O}(\rho||A||_\infty)$, that is, if the order of growth in the elements of U is some constant ρ, called the growth factor. If ρ is about one, no growth is occurring. But if ρ can be a lot more than 1, then we have to expect instability.

There's a famous example in which ρ is 2^{n-1}, displayed here with $n = 5$.

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

Since all elements of A are 1 or 0, $||A||_\infty = 1$. No row swaps will occur, since all sub-diagonal elements are the same magnitude as the original diagonal element. Each time you do a row elimination, the elements of the last row double. At the end, the last element of U is 2^{n-1}.

So in the strictest sense, no, partial pivoting is not stable. But such matrices are incredibly rare (although research continues on why...). If you were working on such a matrix, then the other pivoting strategies (rook, complete) would be worth considering (but they involve a second permutation matrix and more than $3n$ search time).