Department of Mathematics, University of Utah

Analysis of Numerical Methods I

MTH6610 – Section 001 – Fall 2017

Lecture notes – LU factorizations

Monday October 14, 2019

These notes are not a substitute for class attendance. Their main purpose is to provide a lecture overview summarizing the topics covered.

Reading: Trefethen & Bau III, Lecture 20

Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Given any $b \in \mathbb{C}^n$, there is a unique $x \in \mathbb{C}^n$ solving

$$Ax = b.$$

The most elementary way of computing this solution is to solve the above system of n equations and n unknowns via Gaussian elimination. Practically, this entails forming the augmented matrix

$$\left(\begin{array}{c:c} A & b \end{array} \right),$$

and subsequently performing row operations to bring the left $n \times n$ block into upper triangular form, followed by back-substitution to bring the left $n \times n$ block to the identity. The resulting vector in the last column is the solution x.

This procedure of performing row operations can be described without the need for the right-hand side vector b. In this way, we need not form an augmented matrix and instead just use A, but the price we pay is that we need to record the row operations that we use. This recording process can be codified by means of the LU factorization of the matrix A.

The matrix A has columns

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}, \qquad\qquad a_j = \begin{pmatrix} a_{j,1} & a_{j,2)} & \ldots & a_{j,n} \end{pmatrix}^T$$

For notational convenience, we define $a_j^{(1)} \equiv a_j$ for $j = 1, \ldots, n$ and $A_1 \equiv A$. The first step in Gaussian elimination would eliminate (zero-out) all entries in the column a_1 except for the first entry. If $r_j(A)$ represents the jth row of A, this first elimination step performs

$$\begin{aligned}
r_1(A) &\leftarrow r_1(A) \\
r_2(A) &\leftarrow r_2(A) - \frac{a_{1,2}}{a_{1,1}} r_1(A) \\
r_3(A) &\leftarrow r_3(A) - \frac{a_{1,3}}{a_{1,1}} r_1(A) \\
&\;\;\vdots \\
r_n(A) &\leftarrow r_n(A) - \frac{a_{1,n}}{a_{1,1}} r_1(A).
\end{aligned}$$

We can write this in matrix operations:

$$A = \underbrace{\begin{pmatrix} \ell_1 & e_2 & \cdots & e_n \end{pmatrix}}_{L_1} \underbrace{\begin{pmatrix} a_1^{(2)} & a_2^{(2)} & \ldots & a_n^{(2)} \end{pmatrix}}_{A_2},$$

$$\ell_1 = \begin{pmatrix} 1 & \frac{a_{1,2}}{a_{1,1}} & \frac{a_{1,3}}{a_{1,1}} & \cdots & \frac{a_{1,n}}{a_{1,1}} \end{pmatrix}^T,$$

where $a_1^{(2)}$ has zeros in entries $2, \ldots, n$.
We proceed by eliminating entries $3, \ldots, n$ from the vector $a_2^{(2)}$. This results in a similar formula:

$$A_2 = \underbrace{\begin{pmatrix} e_1 & \ell_2 & e_3 & \cdots & e_n \end{pmatrix}}_{L_2} \underbrace{\begin{pmatrix} a_1^{(3)} & a_2^{(3)} & a_3^{(3)} & \ldots & a_n^{(3)} \end{pmatrix}}_{A_3},$$

$$\ell_2 = \begin{pmatrix} 0 & 1 & \frac{a_{2,3}^{(2)}}{a_{2,2}^{(2)}} & \frac{a_{2,4}^{(2)}}{a_{2,2}^{(2)}} & \cdots & \frac{a_{2,n}^{(2)}}{a_{2,2}^{(2)}} \end{pmatrix}^T,$$

We can iterate this procedure: at step k, the first $k-1$ columns of A_k are in upper triangular form. We then focus on the k column of A_k, which is $a_k^{(k)}$. We form a matrix L_k that eliminates entries $k+1, \ldots, n$ from $a_k^{(k)}$ by performing row operations. Putting this all together, we have

$$A = L_1 L_2 \cdots L_{n-1} A_n$$

Note that each L_k is lower-triangular, with each column being a column of the identity, except for column k which is the vector ℓ_k. A computation with this structure of L_k shows that

$$L_k L_{k+1} = \begin{pmatrix} e_1 & e_2 & \cdots & e_{k-1} & \ell_k & \ell_{k+1} & e_{k+2} & \cdots & e_n \end{pmatrix},$$

which is also lower-triangular. Then making the definitions

$$L \coloneqq \prod_{j=1}^{n-1} L_j, \qquad U \coloneqq A_n,$$

we have the equality

$$A = LU,$$

with L lower triangular and U upper triangular. This is the LU factorization of A. Note that, given L and U, computing $A^{-1}b$ requires only forward- and back-substitution operations.
Note that not all invertible matrices have an LU factorization: the steps above require us to divide by $a_{k,k}^{(k)}$ at step k. If any of these values vanish, then Gaussian elimination cannot be completed.