Lecture `linalg1`: Basic Gaussian Elimination

September 23, 2022

Summary: Triangular systems and Gaussian elimination, with 3-by-3 example.

Triangular systems

Diagonal systems are trivial to solve, so let us examine the easiest of nontrivial systems. Namely, triangular systems. An *upper triangular* system has form

$$U\mathbf{x} = \mathbf{b} \quad \Longleftrightarrow \quad \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{1}$$

Here $u_{jk} = 0$ whenever $j > k$ (the row index exceeds the column index). Solve system by **backward substitution**. Note last equation is $u_{nn}x_n = b_n \implies x_n = b_n/u_{nn}$ (provided $u_{nn} \neq 0$). Note second to last equation (implied but not shown) is $u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n = b_{n-1} \implies x_{n-1} = (b_{n-1} - u_{n-1,n}x_n)/u_{n-1,n-1} = (b_{n-1} - u_{n-1,n}b_n/u_{nn})/u_{n-1,n-1}$ (provided both $u_{n-1,n-1}$ and $u_{nn} \neq 0$). We have the algorithm

$$x_k = \left(b_k - \sum_{j=k+1}^{n} u_{kj}x_j \right) \Big/ u_{kk}, \quad \text{for } k = n, n-1, \cdots, 1. \qquad \textbf{(backward substitution).} \tag{2}$$

Note the only possible obstruction to solving the system this way is if one of the diagonal elements $u_{kk} = 0$. Fact: for an upper triangular matrix U, we have $\det U = u_{11}u_{22}\cdots u_{nn}$. You may compute the determinant of a triangular matrix as the product of its diagonal elements.

Similar formulas hold for lower triangular systems,

$$L\mathbf{x} = \mathbf{b} \quad \Longleftrightarrow \quad \begin{pmatrix} \ell_{11} & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \cdots & \ell_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{3}$$

Now $\ell_{jk} = 0$ whenever $k > j$ (column index exceeds row index), and $\det L = \ell_{11}\ell_{22}\cdots\ell_{nn}$. We may solve a lower triangular system via

$$x_k = \left(b_k - \sum_{j=1}^{k-1} \ell_{kj}x_j \right) \Big/ \ell_{kk}, \quad \text{for } k = 1, 2, \cdots, n. \qquad \textbf{(forward substitution),} \tag{4}$$

provided all $\ell_{kk} \neq 0$. Often the strategy is to put a general system into an upper triangular form (by convention, lower triangular would be equally good) using row operations. Let's see how to do that

Gaussian elimination

We will use *row operations* to simplify a general system. Consider the matrix

$$A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 2 \\ -4 & -4 & 1 \end{pmatrix}, \tag{5}$$

here with entries chosen for easy algebra. Say we are interested in solving $Ax = b$ and/or finding A^{-1}. Not unrelated problems, and both may be attacked using *augmented* matrices. We form

$$[A|b] = \left(\begin{array}{rrr|c} 1 & -1 & 3 & b_1 \\ 2 & -2 & 2 & b_2 \\ -4 & -4 & 1 & b_3 \end{array}\right), \qquad [A|I] = \left(\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 2 & -2 & 2 & 0 & 1 & 0 \\ -4 & -4 & 1 & 0 & 0 & 1 \end{array}\right), \tag{6}$$

with the first augmented matrix appropriate for solving $Ax = b$, and the second for finding finding A^{-1}. Now consider the *replacement* matrices

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}. \tag{7}$$

Note that $\det R_1 = 1 = \det R_2$ (each is triangular with a 1 for each diagonal entry). Upon multiplication from the left, R_1 replaces row② with row② − 2 row①, and R_2 replaces row③ with row③ + 4 row①. These matrices have been tailored to zero out the last two entries of the first column of A as follows:

$$\begin{aligned} [R_2R_1A|R_2R_1b] &= \left(\begin{array}{rrr|r} 1 & -1 & 3 & b_1 \\ 0 & 0 & -4 & b_2 - 2b_1 \\ 0 & -8 & 13 & b_3 + 4b_1 \end{array}\right), \\ [R_2R_1A|R_2R_1I] &= \left(\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 0 & -4 & -2 & 1 & 0 \\ 0 & -8 & 13 & 4 & 0 & 1 \end{array}\right), \end{aligned} \tag{8}$$

where $\det R_2R_1A = \det A$ (fact: $\det(AB) = \det A \det B$). Next, we introduce a *permutation* matrix,

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{9}$$

which exchanges row② and row③. Therefore,

$$\begin{aligned} [PR_2R_1A|PR_2R_1b] &= \left(\begin{array}{rrr|r} 1 & -1 & 3 & b_1 \\ 0 & -8 & 13 & b_3 + 4b_1 \\ 0 & 0 & -4 & b_2 - 2b_1 \end{array}\right), \\ [PR_2R_1A|PR_2R_1I] &= \left(\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & -8 & 13 & 4 & 0 & 1 \\ 0 & 0 & -4 & -2 & 1 & 0 \end{array}\right). \end{aligned} \tag{10}$$

Notice that $\det P = -1$, $\det(PR_2R_1A) = -\det A$. Since PR_2R_1A is upper triangular, we see that $\det A = -\det(PR_2R_1A) = -32$, whence A is an invertible matrix, and we can find a unique inverse A^{-1}. However, if our goal is only to solve $Ax = b$, it's less work for us to stop here, finishing off the job with backwards substitution,

$$\begin{aligned} x_3 &= -\tfrac{1}{4}(b_2 - 2b_1) \\ x_2 &= -\tfrac{1}{8}(b_3 + 4b_1 - 13x_3) = -\tfrac{1}{32}(4b_3 + 13b_2 - 10b_1) \\ x_1 &= b_1 + x_2 - 3x_3 = -\tfrac{1}{32}(6b_1 - 11b_2 + 4b_3). \end{aligned} \tag{11}$$

Now, we could in fact "read-off" A^{-1} from these results. Indeed, since we left $\mathbf{b}$ general, the equation for x_1, for example, tells us that the first row of A^{-1} is $[-\frac{3}{16}, \frac{11}{32}, -\frac{1}{8}]$. Nevertheless to explicitly get A^{-1}, we carry on and introduce the following diagonal *pivot* matrices

$$D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{8} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{4} \end{pmatrix}, \tag{12}$$

where $\det D_1 = -\frac{1}{8}$ and $\det D_2 = -\frac{1}{4}$. Upon multiplication from the left, we then have

$$[D_2D_1PR_2R_1A|D_2D_1PR_2R_1I] = \left(\begin{array}{rrr|rrr} 1 & -1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -\frac{13}{8} & -\frac{1}{2} & 0 & -\frac{1}{8} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{4} & 0 \end{array}\right). \tag{13}$$

Next, we use the matrix

$$R_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{14}$$

to replace row① with row① + row②, with result

$$[R_3D_2D_1PR_2R_1A|R_3D_2D_1PR_2R_1I] = \left(\begin{array}{rrr|rrr} 1 & 0 & \frac{11}{8} & \frac{1}{2} & 0 & -\frac{1}{8} \\ 0 & 1 & -\frac{13}{8} & -\frac{1}{2} & 0 & -\frac{1}{8} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{4} & 0 \end{array}\right). \tag{15}$$

Finally, we use

$$R_4 = \begin{pmatrix} 1 & 0 & -\frac{11}{8} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad R_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{13}{8} \\ 0 & 0 & 1 \end{pmatrix} \tag{16}$$

to replace row① with row① $-\frac{11}{8}$ row③, and to replace row② with row② $+\frac{13}{8}$ row③. The result of all row operations is then the following:

$$[R_5R_4R_3D_2D_1PR_2R_1A|R_5R_4R_3D_2D_1PR_2R_1I] = \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{3}{16} & \frac{11}{32} & -\frac{1}{8} \\ 0 & 1 & 0 & \frac{5}{16} & -\frac{13}{32} & -\frac{1}{8} \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{4} & 0 \end{array}\right). \tag{17}$$

Argument has shown that $A^{-1} = R_5R_4R_3D_2D_1PR_2R_1$ and, moreover, that

$$A^{-1} = \begin{pmatrix} -\frac{3}{16} & \frac{11}{32} & -\frac{1}{8} \\ \frac{5}{16} & -\frac{13}{32} & -\frac{1}{8} \\ \frac{1}{2} & -\frac{1}{4} & 0 \end{pmatrix}. \tag{18}$$

We also find

$$\det(A^{-1}) = \det(R_5)\det(R_4)\det(R_3)\det(D_2)\det(D_1)\det(P)\det(R_2)\det(R_1) = -\frac{1}{32}, \tag{19}$$

as expected since $\det(AA^{-1}) = \det I = 1$.

Factorization $A = P^T LU$

Here we show how the matrix (5) can be factorized such that $PA = LU \iff A = P^T LU$, where P is a permutation matrix, L is lower triangular, and U is upper triangular. From the last equation in (10) we have $PMA = U$, where P is the permutation matrix in (9) and

$$U = \begin{pmatrix} 1 & -1 & 3 \\ 0 & -8 & 13 \\ 0 & 0 & -4 \end{pmatrix}, \qquad M = R_2R_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}. \tag{20}$$

Since $P = P^T$ and $PP^T = I = P^2$, we have $MA = PU \implies A = M^{-1}PU \implies PA = PM^{-1}PU$. We now define $L \equiv PM^{-1}P$ and verify that it is indeed lower triangular. First, check that

$$M^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}, \tag{21}$$

so the inverse M^{-1} of M (lower triangular) is also lower triangular. In general, the inverse of a lower triangular matrix is itself lower triangular. Since P swaps the second and third rows,

$$PM^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \implies PM^{-1}P = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}. \tag{22}$$

The last matrix is L, and in all $PA = LU$ reads

$$\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}^{P} \overbrace{\begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 2 \\ -4 & -4 & 1 \end{pmatrix}}^{A} = \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}}^{L} \overbrace{\begin{pmatrix} 1 & -1 & 3 \\ 0 & -8 & 13 \\ 0 & 0 & -4 \end{pmatrix}}^{U}. \tag{23}$$