Mathematics 307—October 25, 1995

Gauss elimination and row reduction II

I have been a bit careless about the best way to perform Gauss elimination on a matrix M so as to factor it.

I recall the setup. At the beginning we are given and $m \times n$ matrix M, and at the end we will have an expression

$$WM = LU$$

where W is an $m \times m$ permutation matrix, and L and U look like this, say, in the 4×4 case:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}, \quad U = \begin{bmatrix} \# & * & * & * \\ 0 & \# & * & * \\ 0 & 0 & \# & * \\ 0 & 0 & 0 & \# \end{bmatrix}$$

although more generally all we can say of U is that it is in row-echelon form. The process works this way: at step i we will have partially calculated versions W_i, L_i, U_i of these factors with the property that

$$W_i M = L_i U_i, \quad \text{or } L_i^{-1} W_i M = U_i \ .$$

To start with $W_0 = I$, $L_0 = I$, $U_0 = M$. At every step W_i will be a permutation matrix, and L_i and U_i will be on the way to their final form: at step i the matrix L_i will look like the final L except that in columns $i+1$, $i+2$, etc. will still be the same as the columns of the identity matrix, and the columns of U_i will be in echelon form only up through the i-th column.

At step i we swap rows of U_i if necessary by choosing the pivot row, according to one of several criteria (according to magnitude of the first non-zero entry, if dealing with matrices in other than exact arithmetic). This gives us a matrix $U_{*,i}$. We apply the same swap $\sigma = \sigma_{i+1}$ to the rows of W_i to get the new W_{i+1}, and to the non-diagonal entries of L_i to get a matrix I'll call $L_{*,i} = \sigma L_i \sigma^{-1}$. Then as explained last time, we now have the equation

$$\sigma L_i^{-1} W_i M = \sigma U_i = U_{*,i}$$

or

$$W_{i+1} M = \sigma W_i M = \sigma L_i \sigma^{-1} \sigma U_i = L_{*,i} U_{*,i} \ .$$

We now have to apply some more elementary row operations to U_i^* to get certain column entries to be 0. This means we multiply $U_{*,i}$ on the left by a certain matrix Λ_{i+1}^{-1}. Thus Λ_{i+1}^{-1} is what we get by applying the same row operations to the identity matrix, or in other words it has entry $-u_{*,i,c}/u_{*,r,c}$ as entry (i,c), with $i > r$. Equivalently, we get Λ_{i+1} by applying the inverse row operation, so its (i,c)-th entry for $i > r$ is $u_{*,i,c}/u_{*,r,c}$. Then

$$\Lambda_{i+1}^{-1} L_{*,i}^{-1} W_{i+1} M = \Lambda_{i+1}^{-1} U_{*,i} = U_{i+1} \ .$$

We then set $L_{i+1} = L_{*,i} \Lambda_{i+1}$, so we have

$$W_{i+1} M = L_{i+1} U_{i+1}$$

and continue on, as long as there remain columns or rows to consider.

In the 4×4 case for example we have

$$L = \Lambda_1 \Lambda_2 \Lambda_1$$

where the Λ_i have the forms

$$\Lambda_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & 0 & 1 & 0 \\ * & 0 & 0 & 1 \end{bmatrix}, \quad \Lambda_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & * & 1 & 0 \\ 0 & * & 0 & 1 \end{bmatrix}, \quad \Lambda_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & * & 1 \end{bmatrix}$$

Example. Take

$$M = \begin{bmatrix} -1 & 2 & 1 & 0 \\ 2 & 4 & -1 & 2 \\ 1 & 2 & -2 & 3 \\ 2 & 3 & 4 & -1 \end{bmatrix}$$

Then in succession

$$U = \begin{bmatrix} -1 & 2 & 1 & 0 \\ 2 & 4 & -1 & 2 \\ 1 & 2 & -2 & 3 \\ 2 & 3 & 4 & -1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 2 & 4 & -1 & 2 \\ 0.0 & 4.0 & 0.5 & 1.0 \\ 0.0 & 0.0 & -1.5 & 2.0 \\ 0.0 & -1.0 & 5.0 & -3.0 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.5 & 1 & 0 & 0 \\ 0.5 & 0 & 1 & 0 \\ 1.0 & 0 & 0 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 2 & 4 & -1 & 2 \\ 0.0 & 4.0 & 0.5 & 1.0 \\ 0.0 & 0.0 & -1.5 & 2.0 \\ 0.0 & 0.0 & 5.125 & -2.75 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.5 & 1 & 0 & 0 \\ 0.5 & 0.0 & 1 & 0 \\ 1.0 & -0.25 & 0 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 2 & 4 & -1 & 2 \\ 0.0 & 4.0 & 0.5 & 1.0 \\ 0.0 & 0.0 & 5.125 & -2.75 \\ 0.0 & 0.0 & 0.0 & 1.19512 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.5 & 1 & 0 & 0 \\ 1.0 & -0.25 & 1 & 0 \\ 0.5 & 0.0 & -0.292683 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Why we choose the pivot carefully

Let's solve the system of equations

$$\begin{bmatrix} 0.001 & 2.000 & 3.000 \\ -1.000 & 3.712 & 4.623 \\ -2.000 & 1.072 & 5.643 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

where we assume we are using a calculator which only stores 4 significant figures. If we are not fussy about pivots, we get in succession

$$U = \begin{bmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0.000 & 2004. & 4.623 + 3000. = 3005. & 3001. \\ 0.000 & 4001. & 5.643 + 6000. = 6006. & 6001. \end{bmatrix}, \quad \begin{bmatrix} 0.001 & 2.000 & 3.000 & 1.000 \\ 0.000 & 2004. & 3005. & 3001. \\ 0.000 & 0.000 & 5.000 & 8.000 \end{bmatrix}$$

which tells us that $z = 0.6250$, whereas if we pivot carefuly we get $z = 0.3670$, which is correct to 4 figures.