THE UNIQUENESS OF THE ROW ECHELON FORM[1]

Let M be a matrix. A matrix M' is called a row echelon form of M if the following conditions are satisfied.

(i) M' is obtained from M by a finite number of the following three operations, called *elementary row operations*: 1) interchange of two rows, 2) multiplying a row by a nonzero scalar, and 3) adding a scalar multiple of a row to another row.

(ii) Each row of M' starts with either with a 1, or with a number of zeros followed by a 1, or the row consists entirely of zeros. The first nonzero entry in a row of M' is called the leading entry of that row; according to what we said, this leading entry must be 1.

(iii) If $l > k > 0$, and row l in M' has as a nonzero entry, then row k must also have a nonzero entry, and the leading entry of row k must occur earlier than the leading entry of row l. In particular, this means that all purely zero rows must occur at the bottom of matrix M'.

(iv) If a column of M' contains a leading entry (of a row), then all other entries in this column must be 0.

Theorem. *The row echelon form of a matrix is unique.*

Proof. In the proof, we will need the following notation. If a matrix M has at least n columns, write $M \restriction n$ for the submatrix resulting from M by deleting all columns after the nth column (in particular, if M has exactly n columns then $M \restriction n = M$). $M \restriction n$ is called the restriction of M to n columns.

The assertion says that a matrix M cannot have two different row echelon forms. Assume, on the contrary that both M_1 and M_2 are row echelon forms of M, and $M_1 \neq M_2$. First notice that, in a row echelon form of M, a column consists of all zeros if and only if the corresponding column in M consists only of zeros; this is because the elementary row operations cannot make all zeros from a nonzero column. Further, observe that the first nonzero column in a row echelon form of M starts with a 1, and all other entries of this column are zero. Therefore, the initial all-zero columns (if any) of M_1 and M_2, and the first column containing a leading entry in M_1 and M_2 must be the same (note that M cannot be the zero matrix, since then its row echelon form would also be the zero matrix, so M would not have two different row echelon forms; so the row echelon form of M must have at least one nonzero column).

So, assume that of M_1 and M_2 agree up to the nth column, and the first column that is different in M_1 and M_2 is the $(n+1)$st column.[2] Then $M_1 \restriction (n+1)$ and $M_2 \restriction (n+1)$ are two different row echelon forms of the matrix $M \restriction (n+1)$. Write $A = M \restriction n$, and let $\mathbf{b}$ be the $(n+1)$st column of M. Then $(A, \mathbf{b}) = M \restriction (n+1)$. Similarly, write $D = M_1 \restriction n = M_2 \restriction n$, and let $\mathbf{f}$ be the $(n+1)$st column of M_1 and $\mathbf{g}$, the $(n+1)$st column of M_2. Then $(D, \mathbf{f}) = M_1 \restriction (n+1)$, $(D, \mathbf{g}) = M_2 \restriction (n+1)$, and $\mathbf{f} \neq \mathbf{g}$. As we explained above, we have $n \geq 1$, the initial zero columns of M_1 and M_2 and the first column containing a leading entry in M_1 and M_2 must be the same, so D must have at least one leading entry.

Consider the system of linear equations $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x}$ is an $n \times 1$ matrix (a column vector of length n). This system of equations is equivalent to both of the systems $D\mathbf{x} = \mathbf{f}$ and $D\mathbf{x} = \mathbf{g}$. We will discuss the solvability of the system of equations $D\mathbf{x} = \mathbf{f}$ (a similar discussion applies to the system $D\mathbf{x} = \mathbf{g}$). Label the columns D containing a leading entry as $l(1)$, $l(2)$, ..., and label the columns not containing a leading entry as $z(1)$, $z(2)$, ... Since, as we mentioned above, D contains at least one leading entry, $l(1)$ is always

[1]Notes for Course Mathematics 10.1 at Brooklyn College of CUNY. Attila Máté, September 13, 2009.

[2]That is, $M_1 \restriction n = M_2 \restriction n$ and $M_1 \restriction (n+1) \neq M_2 \restriction (n+1)$.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

defined. As an example, this labeling for a matrix D is shown here:

$$D = \begin{array}{c} \begin{matrix} l(1) & z(1) & l(2) & z(2) & z(3) & l(3) & z(4) & l(4) & z(5) \end{matrix} \\ \begin{pmatrix} 1 & 2 & 0 & -3 & 2 & 0 & 4 & 0 & 2 \\ 0 & 0 & 1 & -2 & 3 & 0 & 3 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array};$$

that is, $l(1) = 1$, $l(2) = 3$, $l(3) = 6$, $l(4) = 8$, and $z(1) = 2$, $z(2) = 4$, $z(3) = 5$, $z(4) = 7$, $z(5) = 9$. Using this labeling, the solutions of the equation $D\mathbf{x} = \mathbf{f}$ can be easily described; however, here we need to know only somewhat less. Namely, we need to know the following: 1) If column f of the matrix (D, f) contains a leading entry then the equation is unsolvable. Continuing the previous example, in this case the situation we are facing is as follows:

$$(D, \mathbf{f}) = \begin{pmatrix} 1 & 2 & 0 & -3 & 2 & 0 & 4 & 0 & 2 & & 0 \\ 0 & 0 & 1 & -2 & 3 & 0 & 3 & 0 & 5 & & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 & 3 & & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & 1 \end{pmatrix}.$$

Here we somewhat offset the last column of the matrix to indicate that this column corresponds to the right-hand sides of the equation. This system of equation is unsolvable, since the fifth equation requires $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 + 0x_7 + 0x_8 + 0x_9 = 1$, that is, $0 = 1$. On the other hand, if the last column of the matrix the matrix does not contain a leading entry, then the equation is solvable. This is easy to see, but the best way to visualize it is to look at a continuation of the above example:

$$(D, \mathbf{f}) = \begin{pmatrix} 1 & 2 & 0 & -3 & 2 & 0 & 4 & 0 & 2 & & f_1 \\ 0 & 0 & 1 & -2 & 3 & 0 & 3 & 0 & 5 & & f_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 & 3 & & f_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & & f_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & 0 \end{pmatrix}.$$

In this example, the corresponding system of equations can be written as

$$\begin{aligned} x_1 + 2x_2 \qquad - 3x_4 + 2x_5 \qquad + 4x_7 \qquad + 2x_9 &= f_1 \\ x_3 - 2x_4 + 3x_5 \qquad + 3x_7 \qquad + 2x_9 &= f_2 \\ x_6 + 3x_7 \qquad + 3x_9 &= f_3 \\ x_8 + 4x_9 &= f_4, \end{aligned}$$

or else as

$$\begin{aligned} x_{l(1)} + 2x_{z(1)} \qquad - 3x_{z(2)} + 2x_{z(3)} \qquad + 4x_{z(4)} \qquad + 2x_{z(5)} &= f_1 \\ x_{l(2)} - 2x_{z(2)} + 3x_{z(3)} \qquad + 3x_{z(4)} \qquad + 2x_{z(5)} &= f_2 \\ x_{l(3)} + 3x_{z(4)} \qquad + 3x_{z(5)} &= f_3 \\ x_{l(4)} + 4x_{z(5)} &= f_4. \end{aligned}$$

In the example, $x_{l(1)} = x_1 = f_1$, $x_{l(2)} = x_3 = f_2$, $x_{l(3)} = x_6 = f_3$, $x_{l(4)} = x_8 = f_4$, and $x_{z(1)} = x_2 = x_{z(2)} = x_4 = x_{z(3)} = x_5 = x_{z(4)} = x_7 = x_{z(5)} = x_9 = 0$ is a solution of the system equations (there are other solutions, but this is of no interest to us here). In general, if $f^T = [f_1, f_2, \ldots]$,[3] then a solution of the equation $D\mathbf{x} = \mathbf{f}$ is $x_{l(1)} = f_1$, $x_{l(2)} = f_2$, ..., and $x_{z(1)} = x_{z(2)} = \ldots = 0$.

[3] To save space, we describe a column vector here as the transpose of a row vector, since a row vector is easier to print.

Using this, we can complete the proof as follows. If $\mathbf{x}$ is a solution of the equation $A\mathbf{x} = \mathbf{b}$, then $\mathbf{x}$ is also a solution of the equations $D\mathbf{x} = \mathbf{f}$ and $D\mathbf{x} = \mathbf{g}$, and then $\mathbf{f} = D\mathbf{x} = \mathbf{g}$, showing that $\mathbf{f} = \mathbf{g}$, contradicting our assumption that $\mathbf{f} \neq \mathbf{g}$. If the equation $A\mathbf{x} = \mathbf{b}$ is unsolvable, then the equation $D\mathbf{x} = \mathbf{f}$ is also unsolvable. In this case the column $\mathbf{f}$ of $(D, \mathbf{f})$ contains a leading entry in the first row in which the matrix A contains all zeros. The same argument shows that the column $\mathbf{g}$ of $(D, \mathbf{g})$ contains a leading entry at the same place. This shows that, $\mathbf{f} = \mathbf{g}$ again (because the leading entry is 1, and all other entries are 0 in both $\mathbf{f}$ and $\mathbf{g}$). This contradiction completes the proof. $\square$