LINEAR ALGEBRA- LECTURE 5

1. Matrices - Elementary Row operations...continued...

Our discussion in the previous sections shows that matrices are very well suited to understanding a method of finding solutions to a system of linear equations. For we noticed that if we start with a system of linear equations

$$AX = B,$$

and use (a sequence of) elementary row operations on the augmented matrix $(A|B)$ to get a matrix $(A'|B')$, then the system of equations

$$A'X = B'$$

has the same set of solutions as the system $AX = B$. The system of equations $A'X = B'$ can be easier to understand than the original equations if the augmented matrix $(A'|B')$ has a nice form. We now focus on this.

So we will now try to understand if there is a particularly nice form to which every matrix can be row reduced. Recall our convention : for a matrix A, A_i denotes the i-th row vector of A. We make the following definition.

Definition 1.1. A matrix A is said to be a row echelon matrix (or to be in row echelon form) if the following conditions are satisfied.

(1) The first non-zero entry in each row is 1. This is called a pivot.
(2) The first non-zero entry of the $(i+1)$-th row is to the right of the first non-zero entry of the i-th row. That is, the pivot in the $(i+1)$-th row is to the right of the pivot in the i-th row.
(3) The entries above a pivot are zero.

For example, the 4×5 matrix

$$A = (a_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

is a row echelon matrix. The entries a_{12}, a_{23}, a_{35} are pivots. There is no pivot in the 4-th row.

Observe that in a row echelon matrix A if the i-th row A_i consists of zeros, then by property (2) each row A_j, $j > i$ also consists of zeros. Also note that all entries (other than the pivot) in a column containing a pivot are zero. This follows from property (2).

Just so that we understand the definition here is an example of a matrix not in the row echelon form. Consider the matrix

$$A = \begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \end{pmatrix}$$

This violates all the conditions.

It turns out that every matrix can be row reduced to a row echelon matrix. In other words, every matrix can be converted to a row echelon matrix by a sequence of elementary row operations.

Proposition 1.2. Every matrix can be row reduced to a row echelon matrix.

Proof. This can be done by induction using the following algorithmic method. Clearly, the proposition is true for row matrices.

(1) Identify the first column, say the j-th column, that contains a non-zero entry. If this entry is in the i-th row, we interchange the first row and the i-th row (a type (ii) row operation). Then multiply the first row by a suitable scalar (a type (iii) row operation) so that the first non-zero entry in the j-th column becomes 1. The pivot in the first row is in the $1j$-th position.

(2) Now clear out the entries in the ij-th position $i > 1$ by row operations of type(i).

After carrying out operations (1) and (2) on a matrix A, the resultant matrix will have the form

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & \cdots & 0 & \vdots & \vdots & \cdots & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & * & \cdots & * \end{pmatrix}$$

which may be now written in the block form

$$\left(\begin{array}{ccc|c} 0 & \cdots & 1 & B \\ \hline 0 & \cdots & 0 & C \end{array}\right)$$

where the matrix C now has lesser number of rows. By induction hypothesis we may now convert C into a row echelon matrix C' by row operations to get a matrix of the form

$$\left(\begin{array}{ccc|c} 0 & \cdots & 1 & B \\ \hline 0 & \cdots & 0 & C' \end{array}\right)$$

where C' is in row echelon form. We now clear out the entries in the row vector B that are above pivots in C' to get a row vector B'. The matrix

$$\left(\begin{array}{ccc|c} 0 & \cdots & 1 & B' \\ \hline 0 & \cdots & 0 & C' \end{array}\right)$$

is now a row echelon matrix. This completes the proof. □

Let us try to understand why and how the existence of an echelon form can be of help. Suppose we are looking for solutions to the system of m linear equations in n variables given by

$$AX = B.$$

We consider the augmented matrix $(A|B)$ and row reduce it to a row echelon matrix $(A' \,|\, B')$ which represents the system of equations

$$A'X = B'.$$

We know that both these systems have the same set of solutions. Suppose that there is a pivot in the last column B'. Then we claim that the system $A'X = B'$ has no solutions. For, a pivot in the last column B' gives us an equation of the form

$$0 = 1.$$

This implies that the original system $AX = B$ does not have any solutions either. Thus *not* having a pivot in the last column B' is a necessary condition for $A'X = B'$ to have a solution. This is evidently sufficient too. For suppose that the row echelon matrix $(A'|B')$ does not have a pivot in the last column B'. Assume that the pivots appear in columns $j_1, \ldots, j_r$ with

$$1 \leq j_1 < i_2 < \cdots < j_r \leq n$$

where A' is a $m \times n$ matrix. The first row of A' gives us the equation

$$a_{1j_1}x_{j_1} + \sum_{j \neq j_1, \ldots, j_r} a_{1j}x_j = b_1.$$

We may now solve this equation for x_{j_1} by assigning arbitrary but fixed values to x_j, $j \neq j_1, \ldots, j_r$ say $x_j = c_j$ to get $x_{j_1} = c_{j_1}$, say. Next, the second row of $(A'|B')$ gives us the equation

$$a_{2j_2}x_{j_2} + \sum_{j \neq j_1, \ldots, j_r} a_{2j}x_j = b_2.$$

We may now solve this equation for x_{j_2} using the above fixed values x_j, $j \neq j_1, i_2, \ldots, j_r$ to get $x_{j_2} = c_{j_2}$, say. Continuing this way we can solve for the unknows $x_{j_1}, \ldots, x_{j_r}$. It is clear that $(c_1, \ldots, c_n)$ is a solution for the system $A'X = B'$. We record these observations below.

Proposition 1.3. Suppose $(A'|B')$ is a row echelon matrix. Then the system $A'X = B'$ has a solution if and only if there is no pivot in the last last column B'. In this case a solution can be obtained by assigning arbitrary value to the variable x_j if there is no pivot in the j-th column. $\square$

Here are two examples.

Example 1.4. Consider the system of equations

$$\begin{aligned} x_2 + 5x_3 &= -4 \\ x_1 + 4x_2 + 3x_3 &= -2 \\ 2x_1 + 7x_2 + x_3 &= 2. \end{aligned} \tag{1.4.1}$$

The associated augmented matrix is

$$(A \mid B) = \begin{pmatrix} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & 2 \end{pmatrix}$$

We now perform elementary row operations to row reduce the above matrix to a row echelon matrix as follows.

$$\begin{pmatrix} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & 2 \end{pmatrix} \xrightarrow{X_2:X_1} \begin{pmatrix} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & 2 \end{pmatrix} \xrightarrow{X_3:X_3-2X_1} \begin{pmatrix} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 0 & -1 & -5 & 6 \end{pmatrix}$$

$$\xrightarrow{X_1:X_1-4X_2} \begin{pmatrix} 1 & 0 & -17 & 14 \\ 0 & 1 & 5 & -4 \\ 0 & -1 & -5 & 6 \end{pmatrix} \xrightarrow{X_3:X_3+X_2} \begin{pmatrix} 1 & 0 & -17 & 14 \\ 0 & 1 & 5 & -4 \\ 0 & 0 & 0 & 2 \end{pmatrix} \cdots \longrightarrow \begin{pmatrix} 1 & 0 & -17 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (A' \mid B')$$

Thus the augmented matrix $(A' \mid B')$ represents the system of equations

$$\begin{aligned} x_1 - 17x_3 &= 0 \\ x_2 + 5x_3 &= 0 \\ 0 &= 1 \end{aligned} \tag{1.4.2}$$

and hence the system (1.4.2) cannot have any solutions. This implies that the original system (1.4.1) cannot have any solutions either. We note that the (augmented) row echelon matrix $(A' \mid B')$ has a pivot in the last column and therefore the system of equations that it represents cannot have a solution (Proposition 1.3).

Example 1.5. Consider the system of equations

$$\begin{aligned} 2x_1 - 6x_3 &= -8 \\ x_2 + 2x_3 &= 3 \\ 3x_1 + 6x_2 - 2x_3 &= -4 \end{aligned} \tag{1.5.1}$$

The associated augmented matrix is

$$\begin{pmatrix} 2 & 0 & -6 & -8 \\ 0 & 1 & 2 & 3 \\ 3 & 6 & -2 & -4 \end{pmatrix}$$

It is now easy to check that the above augmented matrix can be row reduced by elementary row operations to the row echelon matrix

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

This represents the system of equations

$$\begin{aligned} x_1 &= 2 \\ x_2 &= -1 \\ x_3 &= 2 \end{aligned} \tag{1.5.2}$$

which now gives us the solution to the original system of equations.

Having looked at two examples, we now again analyse Proposition 1.3 and derive some consequences. First observe that every homogeneous system of m linear equations in n variables

$$AX = 0$$

has a solution, namely, $X = 0$. In trying to obtain a solution to the above system we need not look at the augemented matrix since the last column of the augmented matrix is zero. We just try to row reduce the matrix A of coefficients by elementary row operations to a row echelon matrix A' to get the system of equations

$$A'X = 0.$$

Now observe that if the number of pivots in the matrix A' is r, then

$$r \leq \min\{m, n\}.$$

Now suppose that $m < n$, that is, the number of equations is leass than the number of unknowns, the proof of Proposition 1.3 shows that we may assign arbitrary values to $n - r$ many variables to get a solution of the system. This tells us the following.

Corollary 1.6. Let $AX = 0$ be a homogeneous system of m linear equations in n variables. If $m < n$, then the system has a non-zero solution. □

Here is a simple fact about square row echelon matrices whose proof is left as an exercise.

Lemma 1.7. If A is a square row echelon matrix, the either A is the identity matrix or the last row of A is zero.

Proof. Exercise. □

We may now characterize invertible matrices.

Proposition 1.8. Let A be a square matrix. Then the following statements are equivalent.

(1) A is row equivalent to the identity matrix.
(2) A is a product of elementary matrices.
(3) A is invertible.
(4) The system of homogeneous equations $AX = 0$ has only the trivial solution.

Proof. Assume (1) holds. Thus after a sequence of, say s many, elementary row operations on A we get the identity matrix. Since every elementary row operations corresponds to a left multiplication by an elementary matrix, we may write

$$E_s E_{s-1} \cdots E_1 A = \mathbb{I}$$

and therefore

$$A = E_1^{-1} \cdots E_s^{-1}.$$

Thus (1) implies (2) as the inverse of an elementary matrix is again an elementary matrix.

Next assume that $A = E_1 \cdots E_s$ is a product of elementary matrices. Then as each E_i is invertibe, so is their product and hence A is invertible. Thus (2) implies (3).

Next assume that (3) holds so that A is invertible. Given a homogeneous system$AX = 0$, it is clear that $X = A^{-1}0 = 0$ is the only solution.

Finally assume that the homogeneous system $AX = 0$ has only the trivial solution. We now row reduce A to a row echelon matrix A'. Then it is clear that either A' is the identity matrix or the last row is zero. If the last row is not zero we are done. So assume that the last row of A' is zero. Then, by Corollary 1.6, the system $A'X = 0$ has a non-trivial solution and hence so does $AX = 0$. Thus A' must be the identity matrix and therefore (4) implies (1). □

The above proposition presents us with several ways to decide if a square matrix is invertible. We can also compute the inverse of a matrix. For example, suppose A is an invertible matrix. Then we know that A can be row reduced to the identity matrix. Thus we can write

$$E_s E_{s-1} \cdots E_1 A = \mathbb{I}.$$

where E_i's are elementary matrices. Hence,

$$E_s \cdots E_1 \mathbb{I} = A^{-1}.$$

This means that we can apply the same set of elementary row operations to the identity matrix (that we have applied on A) to get A^{-1}.

Example 1.9. Suppose we wish to find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

So we find out what are the elementary row operations that will reduce A to the identity matrix. The same operations then are applied to the identity matrix. To do this it is convenient to work on both A and $\mathbb{I}$ simultaneously by considering the 2×4 matrix below.

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix} \xrightarrow{X_2 : X_2 - X_1} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix} \xrightarrow{X_1 : X_1 - 2X_2} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$

Thus

$$A^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$$

Here are some exercises.

Exercise 1.10. Complete the proof of Lemma 1.7.

Exercise 1.11. Find solutions to the system of equations

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9. \end{aligned}$$

Exercise 1.12. Find solution to the system of equations

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 + -9x_4 + 6x_5 &= 15 \\ 2x_2 - 4x_3 + 4x_4 + 2x_5 &= -6 \end{aligned}$$

Exercise 1.13. Find inverses of the following matrices using the method discussed

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$$

Exercise 1.14. Let A, B be $n \times n$ matrices such that AB is invertible. Show that A, B are both invertible.

Exercise 1.15. Let $AX = B$ be a system of linear equations where A, B are real matrices. Prove that if the system $AX = B$ has more than one solution, then it has infinitely many solutions. Prove that if there is a solution in the complex numbers, then there is a also a real solution.

Exercise 1.16. Let A be a square matrix. Show that if the system $AX = B$ has a unique solution for some particular column vector B, then it has a unique solution for all B.

Exercise 1.17. Use elementary row operations to determine whether the matrix

$$\begin{pmatrix} 2 & 5 & -1 \\ 4 & -1 & 2 \\ 6 & 4 & 1 \end{pmatrix}$$

is invertible, and to find the inverse if it is.

Exercise 1.18. Let A be a $n \times n$ matrix. If A is invertibe and $AB = 0$ for some $n \times n$ matrix B, then show that $B = 0$. If A is not invertible show that there exists a $n \times n$ matrix B with $AB = 0$ and $B \neq 0$.