LINEAR ALGEBRA- LECTURE 2

1. Matrices - ii

Recall that last time we had discussed matrix addition and multiplication. We also noted two properties, namely that matrix multiplication distributes over matrix addition and that matrix multiplication is associative. In this section we will discuss a few more properties and see some examples.

To begin we first note that

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 7 & 7 \end{pmatrix}$$

whereas

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 4 & 6 \end{pmatrix}$$

Thus matrix multiplication is in general not commutative. Thus we need to be careful when dealing with matrix multiplication and equations. For example, if A, B are square matrices of the same size, then

$$(A + B)^2 = A^2 + AB + BA + B^2.$$

If for two matrices A, B the equality

$$AB = BA$$

holds, then we say that the matrices A and B commute. Here are some special kinds of matrices that we often encounter.

Definition 1.1. A matrix all of whose entries are 0 is called the zero matrix and we denote it by the symbol 0. A square matrix $A = (a_{ij})$ is said to be a diagonal matrix if the off diagonal entries are all zero. In other words A is diagonal if $a_{ij} = 0$ whenever $i \neq j$.

Definition 1.2. The $n \times n$ matrix

$$\mathbb{I}_n = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

in which all the off diagonal entries are zero and all the diagonal entries are 1 is called the $n \times n$ identity matrix.

The identity matrix $\mathbb{I}_n$ has the property that for any $n \times n$ matrix A, we have

$$A\mathbb{I}_n = \mathbb{I}_n A = A.$$

Having defined the identity matrix we define an inverse of a matrix.

Definition 1.3. Let A be a $n \times n$ matrix. A $n \times n$ matrix B is said to be an inverse of A if both equalities

$$AB = \mathbb{I}_n, \quad BA = \mathbb{I}_n \tag{1.3.1}$$

hold. If such a B exists, we say that A is invertible. We write $A^{-1} = B$.

Not all matrices are invertible. It is easy to see that if B is an inverse of A, then B is unique. For suppose that B, B' are two $n \times n$ matrices such that

$$B'A = \mathbb{I}_n = AB.$$

Then,

$$B' = B'\mathbb{I}_n = B'(AB) = (B'A)B = \mathbb{I}_n B = B.$$

Thus an inverse if it exists is unique. It turns out that to show $A^{-1} = B$ we need not check that both equalities in (1.3.1) hold. If one holds the other holds too. We shall prove this soon.

It is easy to see that the product of two invertible matrices is invertible. Indeed, if A, B are two invertible matrices, then as

$$ABB^{-1}A^{-1} = \mathbb{I}_n = B^{-1}A^{-1}AB$$

we conclude that (AB) is also invertible and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Thus, the product of finitely many invertible matrices is invertible.

The entries of an invertible matrix are somewhat constrained. For example if a whole row (or column) of an $n \times n$ matrix A is zero, then the matrix cannot be invertible. This verification is left as an exercise.

It is instructive to pause at this point and see the connection between matrices, matrix multiplication and system of linear equations. Suppose we are given a system of m linear equations in n variables

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \tag{1.3.2}$$

The above system gives rise to three matrices : the $m \times n$ matrix A

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \vdots & \vdots & \dots & \dots & \vdots \\ a_{mn} & a_{m2} & \dots & \dots & a_{mn} \end{pmatrix}$$

consisting of the coefficients of the above system; the $n \times 1$ column vectors of variables

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and the $m \times 1$ column vector

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

From our discussion it is clear that the system (1.3.2) may be written as a single equation

$$AX = B.$$

All the m equations in (1.3.2) are captured in the single equation above. This is not just a notational convenience, for if the matrix A were invertible, then the unknowns can be immediately computed from

$$X = A^{-1}B.$$

We now come back to our discussion on some standard types of matrices that we often encounter. We shall denote by e_{ij} the $m \times n$ matrix that has 1 in the ij-th place and 0 elsewhere. The matrix e_{ij} is called a matrix unit. For example,

$$e_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0, \end{pmatrix}$$

is the 2×3 matrix unit e_{12}. It is clear that every $m \times n$ matrix $A = (a_{ij})$ can be written as a linear combination of matrix units in a unique way. Indeed,

$$A = \sum_{i,j} a_{ij} e_{ij}.$$

For example,

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} &= 1\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 4\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= e_{12} + 2e_{12} + 3e_{21} + 4e_{22}. \end{aligned}$$

We use a special notation to denote the matrix unit e_{i1} of size $m \times 1$. Traditionally, we denote it by

$$e_i = e_{i1} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Thus e_i is the column vector with 1 in the i-th place and 0 elsewhere and hence any $m \times 1$ column vector can be uniquely written as

$$X = x_1 e_1 + x_2 e_2 + \cdots + x_m e_m$$

where x_i are the entries of X.

Let us look at the definition of multiplication of matrices a bit more closely. Let $X = (x_{ij})$ be a $n \times r$ matrix. Let the row vector X_i denote the i-th row of X. Thus

$$X_i = \begin{pmatrix} x_{i1}, & x_{i2}, & \ldots, & x_{ir} \end{pmatrix}$$

We denote the matrix X in the following fashion

$$X = \begin{pmatrix} - & X_1 & - \\ - & X_2 & - \\ \vdots & \vdots & \vdots \\ - & X_n & - \end{pmatrix}$$

Now suppose we left multiply the $n \times r$ matrix X by the $n \times n$ matrix A to get the $n \times r$ matrix $Y = (y_{ij})$, that is

$$AX = Y$$

then we wish to understand how the matrix

$$Y = \begin{pmatrix} - & Y_1 & - \\ - & Y_2 & - \\ \vdots & \vdots & \vdots \\ - & Y_n & - \end{pmatrix}$$

looks like. It is therefore enough to understand what the row vector Y_i is. By the definition of the matrix product we obtain

$$\begin{aligned} Y_i &= \left(\Sigma_k a_{ik}x_{k1}, \quad \Sigma a_{ik}x_{k2}, \quad \cdots, \quad \Sigma_k a_{ik}x_{kr}\right) \\ &= \left(a_{i1}x_{11} + a_{i2}x_{21} + \cdots + a_{in}x_{n1}, \quad \cdots, \quad a_{i1}x_{1r} + a_{i2}x_{2r} + \cdots + a_{in}x_{nr}\right) \\ &= a_{i1}\left(x_{11}, \quad x_{12}, \cdots, \quad x_{1n}\right) + \cdots + a_{in}\left(x_{n1}, \quad x_{12}, \cdots, \quad x_{nr}\right) \\ &= a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{in}X_n \end{aligned}$$

Thus we see that when we multiply two matrices AX as above we get a matrix Y whose i-th row is a linear combination of the rows of X with coefficients the i-th row of A. This is an easy but important observation that we should keep in mind. In particular, each row of Y is a linear combination of the rows of X with coefficients coming from an appropriate row of X.

We shall study products AX when A are some specific type of matrices. This will lead us to the definition of row operations. As of now let us end this section with some examples.

Example 1.4. Let us look at a product AX where A and X are clear from the expression below.

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 + a \cdot 3 & 2 + 3 \cdot 4 \\ 3 & 4 \end{pmatrix}$$

In this example, left multiplication by A modifies the matrix X in this fashion : it modifies the row vector X_1 of X to

$$Y_1 = X_1 + aX_2.$$

and does not change X_2 so that $Y_2 = X_2$. Note that A is invertible.

Example 1.5. Let us look at a product AX where A and X are clear from the expression below.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

In this example, left multiplication by A interchanges the rows of the matrix X. Again, the matrix A is invertible.

Example 1.6. Let us look at a product AX where A and X are clear from the expression below.

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} a & 2a \\ 3 & 4 \end{pmatrix}$$

In this example, left multiplication by A modifies the matrix X in this fashion : it modifies the row vector X_1 of X to

$$Y_1 = aX_1.$$

and does not change X_2 so that $Y_2 = X_2$. Note that A is invertible.

We shall try to understand such operations next time. Note that in all the three examples above, the matrix A can be obtained as a combination of the identity matrix and suitable unit matrices.

Here are some exercises.

Exercise 1.7. Let $A = e_{ij}$ be an $n \times n$ matrix unit and $X = (x_{ij})$ a $n \times r$ matrix. If $AX = Y$, show that

$$Y_k = \begin{matrix} 0 & k \neq i \\ X_j & k = i \end{matrix}$$

Thus left multiplication by a matrix unit e_{ij} transforms the matrix X to a matrix Y all of whose rows are zero except the i-th row which now equals the j-th row X_j of X.

Exercise 1.8. Let e_{ij} be an $n \times n$ matrix unit. Show that

$$e_{ij}e_{ij} = \begin{matrix} 0 & i \neq j \\ e_{ij} & i = j. \end{matrix}$$