Let \( m \) and \( n \) be two positive integers. A rectangular array (of numbers) consisting of \( m \) rows and \( n \) columns as

\[
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

is called an \( m \) by \( n \) matrix, or an \( m \times n \) matrix. We call \( a_{ij} \) the \( ij \)-entry or the \( ij \)-element of the matrix; The element* \( a_{ij} \) are indexed so that the first subscript \( i \) indicates the row, and the second subscript \( k \) indicates the column in which \( a_{ij} \) occurs. The more compact notation

\[(a_{ij})_{i,j=1}^{m,n}, \quad \text{or} \quad (a_{ij})\]

is also used. The collection of \( m \) by \( n \) matrices is denoted by \( M_{m \times n} \), and as such we write \((a_{ij}) \in M_{m \times n}\).

If \( m = n \), the matrix is said to be a square matrix. The row vectors are the \( 1 \) by \( n \) matrices, and the column vectors are the \( m \) by \( 1 \) matrices. Thus, the \( i \)-th row of \((a_{ij})_{i,j=1}^{m,n}\) is the row vector \((a_{i1}, a_{i2}, \ldots, a_{in})\), and the \( j \)-th column of \((a_{ij})_{i,j=1}^{m,n}\) is the column vector

\[
\begin{bmatrix}
a_{i1} \\
\vdots \\
a_{in}
\end{bmatrix}
\]

Two matrices \( A = (a_{ij}) \) and \( B = (b_{ij}) \) are equal if and only if they have the same number of rows, the same number of columns, and equal entries \( a_{ij} = b_{ij} \) for each pair \( i \) and \( j \).

Matrices arise naturally as representation of linear transformations, but they can also considered as objects existing in their own right, without necessarily being connected to linear transformations. As such, they form another class of mathematical objects on which algebraic operations can be defined. The connection with linear transformations serves as motivation for these definitions, but this connection, which is discussed in [Apo, 16.1–16.11], will be ignored in our discussion.

**Linear spaces of matrices.** We now define addition of matrices and multiplication by scalars. If \( A = (a_{ij}) \) and \( B = (b_{ij}) \) are in \( M_{m \times n} \) and if \( c \) is any scalar then we define \( m \) by \( n \) matrices \( A + B \) and \( cA \) as

\[
A + B = (a_{ij} + b_{ij}), \quad cA = (ca_{ij}).
\]

*The element \( a_{ij} \) may be arbitrary objects of any kind. Usually, they are real (or complex) numbers, but sometimes it is convenient to consider matrices whose elements are other objects, for example, functions.*
In other words, \( A + B \) is obtained by adding the corresponding entries of \( A \) and \( B \) and \( cA \) is obtained from \( A \) by multiplying each entry of \( A \) by \( c \). The sum is defined only when \( A \) and \( B \) have the same size. For example, if 

\[
A = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix},
\]

then

\[
A + B = \begin{bmatrix} 6 & 2 & -2 \\ 0 & -2 & 7 \end{bmatrix}, \quad 2A = \begin{bmatrix} 2 & 4 & -6 \\ -2 & 0 & 8 \end{bmatrix}.
\]

The zero matrix \( O \) is the \( m \times n \) matrix all of whose entries are 0. It is straightforward to verify that \( M_{m \times n} \) is a linear space with these definitions.

**Exercise.** The linear space \( M_{m \times n} \) has dimension \( mn \). Find a basis for \( M_{2 \times 3} \).

**Matrix multiplication.** It is not surprising that \( M_{m \times n} \) is a linear space since an \( m \) by \( n \) matrix is very much like an \( mn \)-vector; the only difference is that the components are written in a rectangular array instead of a linear array. Unlike tuples, however, matrices have a further operation − product.

If \( A = (a_{ij}) \in M_{m \times p} \) and \( B = (b_{ij}) \in M_{p \times n} \) then the product \( AB \) is defined to be \( C = (c_{ij}) \in M_{m \times n} \) whose \( ij \)-entry is given by

\[
c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}.
\]

That is, the \( ij \)-entry of \( AB \) is to take the dot product of the \( i \)-th row vector of \( A \) and the \( j \)-th column vector of \( B \). Thus, matrix multiplication can be regarded as a generalization of the dot product.

**Nota Bene:** The product \( AB \) is defined only when the number of columns of \( A \) is equal to the number of rows of \( B \).

For example, if

\[
A = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 6 \\ 5 & -1 \\ 0 & 2 \end{bmatrix},
\]

then the product \( AB \in M_{2 \times 2} \) and \( BA \in M_{3 \times 3} \) are given by

\[
AB = \begin{bmatrix} 17 & 21 \\ 1 & -7 \end{bmatrix}, \quad BA = \begin{bmatrix} 6 & 10 & 8 \\ 16 & 4 & 10 \\ -2 & 2 & 0 \end{bmatrix}.
\]

The next theorem states some algebraic properties of matrix multiplication.

**Theorem 1.** Given matrices \( A, B \) and \( C \), and any scalar \( c \),

(a) (associative law) \( A(BC) = (AB)C \), provided that \( A(BC) \) and \( (AB)C \) are meaningful.

(b) (distributive law) \( (A + B)C = AC + BC \), provided that \( AC \) and \( BC \) are meaningful.

(c) (homogeneity) \( (cA)B = c(AB) \), provided that \( AB \) is meaningful.
These properties can be deduced directly from the definition of multiplication, and I will leave the proof of the theorem to readers. An alternative proof using the ideas of linear transformation is given in [Apo, pp. 603].

The square matrix
\[
\begin{bmatrix}
1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 1
\end{bmatrix}
\]
denoted by \( I_n \) or simply by \( I \), is called the identity matrix. In other words, in an identity matrix, all diagonal entries are 1 and other entries are all zeros. It can be represented using the Kronecker's delta\(^\dagger\) as \( I_n = (\delta_{ij}) \), where
\[
\delta_{ij} = \begin{cases} 
1 & \text{if } i = j \\
0 & \text{if } i \neq j.
\end{cases}
\]
Note that \( AI = IA = A \) for any square matrix \( A \).

If \( A \) is a square matrix, we define integral powers of \( A \) inductively as \( A^0 = I \) and \( A^n = AA^{n-1} \) for \( n \geq 1 \).

Nota Bene: In general \( AB \neq BA \), that is, matrix multiplication is not commutative, even if \( AB \) and \( BA \) are both defined and of the same size. For example, if
\[
A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix},
\]
then
\[
AB = \begin{bmatrix} 13 & 8 \\ 2 & -2 \end{bmatrix}, \quad BA = \begin{bmatrix} -1 & 10 \\ 3 & 12 \end{bmatrix},
\]
and therefore \( AB \neq BA \). There are, though, a few important exceptions. For example, \( AI = IA \) for any square matrix \( A \).

At this point, a natural question to ask would be: a matrix has a multiplicative inverse (such that \( AB = I \) and \( BA = I \))? We will answer for this question in a later chapter.

REFERENCES


©2008 BY VERA MIKYOUNG HUR
E-mail address: verahur@math.mit.edu

\(^\dagger\)It is named after Leopold Kronecker (1823–1891), useful in signal processing and linear algebra.