Matrices

Preliminaries. A $(m \times n)$-**matrix** is a rectangular array of numbers comprised of m rows and n columns. We refer to $(m \times n)$ as the **dimension** of the matrix. When a matrix has the same number of rows as columns we refer to it as a square matrix. We often denote matrices by capital letters i.e. A,B,C etc. In addition, denote the set of all $(n \times n)$-matrices with real numbers as entries by $M_n(\mathbb{R})$. The entry in the i^{th} row and j^{th} column of a matrix A is denoted A_{ij} and sometimes a_{ij}.

Example. *Below we give examples of a* (2×2)*-matrix,* (4×1)*-matrix,* (2×4)*-matrix, and* (3×3)*-matrix, respectively.*

$$\begin{bmatrix} 1 & 0 \\ -2 & 17 \end{bmatrix} \qquad \begin{bmatrix} 9 \\ -8 \\ 0 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & -1 & \pi \\ 0 & 0 & 0 & 42 \end{bmatrix} \qquad \begin{bmatrix} 2 & 2 & 11 \\ 4 & 3 & 13 \\ 6 & 5 & 19 \end{bmatrix}$$

A **scalar** is a quantity which can be described by a single real number. In contrast, a **vector** is a quantity which has both magnitude and direction, hence, it is described by several real numbers. An example of a scalar is the speed of a particle travelling through 3-dimensional space, whilst an example of a vector would be the particle's velocity (which is its speed and direction of motion). The first quantity can be described by a non-negative real number, whilst the second quantity can be described by an arrow in 3-dimensional space, whose length is the speed of the particle, and whose direction is given by the direction of motion of the particle. Such an arrow may be specified by three real numbers which represent the x, y, and z coordinates of the arrowhead. Hence, a vector may be thought of as an arrow, or as a matrix consisting entirely of one row, or entirely of one column. We often refer to $(1 \times n)$-matrices as row vectors, and $(n \times 1)$-matrices as column vectors.

Matrix Addition. Two matrices A, B are equal, written $A = B$ iff their dimensions agree, and their corresponding entries are equal.

We can add (or subtract) matrices if their dimensions agree. We do so by adding (or subtracting) their corresponding entries.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

Scalar Multiplication. Given $k \in \mathbb{R}$, we can multiply a matrix A by k. This is defined by multiplying each entry in A by k. We denote the product by kA, and refer to this operation as scalar multiplication.

$$k \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}$$

Matrix Multiplication. We have seen how to multiply a matrix by a real number, can we multiply matrices themselves? Matrix multiplication has a somewhat involved definition - the rough idea is that, given a $(m \times n)$-matrix A and a $(n \times p)$-matrix B their product is a $(m \times p)$-matrix AB. This product is obtained by multiplying along the rows of A and down the columns of B. Formally, the definition can be written

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

Consider the case where both A and B are (2×2)-matrices:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}$$

Note that fundamentally we are multiplying along the rows of the first matrix, and down the columns of the second matrix. The number of columns of the first matrix must agree with the number of rows of the second matrix in order for this product to be defined. That is, the matrix product AB is only defined if the number of columns of A is equal to the number of rows of B.

Note that even if AB is defined, this does not necessarily imply that BA is defined. However, both are defined if A and B are square matrices of the same dimension.

One dramatic difference between matrix multiplication, and the multiplication of ordinary numbers is that even when AB and BA are both defined, they are not necessarily equal! That is, in general, $AB \neq BA$. We say that matrix multiplication is **not commutative** because it matters in what order the multiplication is performed.

Example. *Consider* $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}$.

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix} =$$

$$BA = \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} =$$

Example. *Consider* $A = \begin{bmatrix} 2 & 1 & 5 \\ 0 & -1 & 0 \\ 6 & 6 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 5 & 2 & -3 \\ -1 & 1 & 3 \\ 5 & 0 & 2 \end{bmatrix}$.

$$AB = \begin{bmatrix} 2 & 1 & 5 \\ 0 & -1 & 0 \\ 6 & 6 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & -3 \\ -1 & 1 & 3 \\ 5 & 0 & 2 \end{bmatrix} =$$

$$BA = \begin{bmatrix} 5 & 2 & -3 \\ -1 & 1 & 3 \\ 5 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 5 \\ 0 & -1 & 0 \\ 6 & 6 & 1 \end{bmatrix} =$$

Transpose. Given a matrix A we define its **transpose** matrix A^T to be that matrix obtained from A by switching the first row and first column, switching the second row and second column, switching the third row and third column, etc. The transpose operation takes a $(m \times n)$-matrix, and maps it to a $(n \times m)$-matrix. Formally, $(A^T)_{ij} = A_{ji}$.

Example.

$$\begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix}^T = \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ -1 & -2 \\ 5 & 7 \\ 0 & 112 \end{bmatrix}^T = \begin{bmatrix} \quad & \quad & \quad & \quad \\ \quad & \quad & \quad & \quad \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 6 \\ 0.2 & -1 & 0 \\ 3 & -4 & 5 \end{bmatrix}^T = \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

If a matrix A satisfies $A^T = A$, then we say that the matrix **symmetric**.

Note that for any matrix A we have that $(A^T)^T = A$.

Theorem. *Suppose we have matrices A and B whose product AB is defined. Then,*

$$(AB)^T = B^T A^T.$$

Special Matrices. For each dimension $(m \times n)$ we have a unique matrix whose every entry is 0. This is referred to (by abuse of grammar) as the **zero matrix**. See the examples below.

$$\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

If a square matrix has 1 down the lead diagonal, and 0 for every other entry, then we refer to it as the identity matrix, and denote it by I. There is a unique identity matrix for each positive integer n. See the examples below.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The zero matrix and identity matrix play roles in matrix arithmetic, analogous to the role of the numbers 0 and 1 in ordinary arithmetic.

Inverses. Suppose we have a square matrix A, and their exists another square matrix B such that

$$AB = BA = I.$$

Then B is said to be the **inverse** of A, and is written $B = A^{-1}$. Matrices which are not square do not have inverses. Moreover, even if a matrix is square it may fail to have an inverse. We call a square matrix **invertible** if it has an inverse, and **singular** if it does not have an inverse.

Theorem. *If a matrix is invertible, then its inverse is unique.*

Theorem. *Suppose that A and B are invertible matrices of the same dimension. Then the following properties hold:*

(i) $$(A^{-1})^{-1} = A$$

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

(iii) $$(A^T)^{-1} = (A^{-1})^T$$

In general, finding the inverse of a matrix can be computationally difficult, however, for (2×2)-matrices there is a simple formula.

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $A^{-1} = \dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Example. *Multiply out AA^{-1} and $A^{-1}A$ to check that they do indeed yield the (2×2) identity matrix.*

$$AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} =$$

$$A^{-1}A = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} =$$

Determinants. There is a very important function associated with matrices, called the **determinant**

$$\det : \Big\{ \text{all square matrices} \Big\} \to \mathbb{R}$$

which determines whether a square matrix is invertible, or not.

Theorem. *Suppose A is a square matrix.*

$$\det(A) \neq 0 \iff A \textit{ is invertible.}$$

The definition of the determinant is somewhat complicated in general, however, for (2×2)-matrices it is straightforward.

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = ad - bc$.

From this definition, and the formula for the inverse of a (2×2)-matrix above, it should be clear that $\det(A) \neq 0$ is a necessary condition for the inverse to exist. We will not prove it here, but it is also a sufficient condition.

Use the determinant to ascertain whether the following matrices are invertible. If they are, then write down their inverse.

(i) $\begin{bmatrix} 1 & 7 \\ -2 & 5 \end{bmatrix}$

(ii) $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$

(iii) $\begin{bmatrix} 2 & -1 \\ 10 & \pi \end{bmatrix}$

(iv) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

Theorem. *Suppose A and B are square matrices. Then,*

(i) $$\det(AB) = \det(A)\det(B).$$

(ii) $$\det(A^T) = \det(A).$$

From the first statement in the theorem above, it follows that if A is also invertible, then

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Matrices as Transformations. Let A be a $(m\times n)$-matrix. We can multiply A by a column vector of length n (that is, an $(n\times 1)$ matrix) on the right, and the product will be a column vector of length m (that is, an $(m\times 1)$ matrix). This allows us to think of $(m\times n)$-matrices as functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. In particular, (2×2)-matrices map $\mathbb{R}^2$ to itself, so they transform one vector in the plane to another. This perspective, where $(m\times n)$-matrices are viewed as functions which transform vectors of length n to vectors of length m, is incredibly powerful, and forms the foundation of the branch of mathematics known as Linear Algebra.