2 Matrix Algebra

2.1 Matrix Operations

If A is an $m \times n$ matrix, that is, a matrix with m rows and n colmumns, then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j))-entry of A. Each column of A is a list of m numbers, which identifies a vector in $\mathbb{R}^m$. If we denote these columns by $\mathbf{a}_1, \dots, \mathbf{a}_n$, we may write A as

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

The number a_{ij} is the ith entry of the jth column vector $\mathbf{a}_j$.

The diagonal entries in an $m \times n$ matrix $A = [a_{ij}]$ are $a_{11}, a_{22}, \dots$ and they form the main diagonal of A. A diagonal matrix is a square $n \times n$ matrix whose nondiagonal entries are zero. An example is the $n \times n$ identity matrix I_n. An $m \times n$ matrix whose entries are all zero is a zero matrix and is written as 0.

We say that two matrices are equal if they have the same size and their corresponding entries are equal. If A and B are $m \times n$ matrices, then the sum $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns of A and B. Since vector addition of the columns is done entrywise, each entry in $A + B$ is the sum of the corresponding entries in A and B. The sum $A + B$ is defined only when A and B are the same size.

If r is a scalar and A is a matrix, then the scalar multiple rA is the matrix whose columns are r times the corresponding columns in A. As with vectors, $-A$ stands for $(-1)A$.

Theorem 1. *Let A, B, C be matrices of the same size and let r, s be scalars.*

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

When a matrix B is multiplied by vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$. If this vector is then multiplied by a matrix A, the resulting vector is $A(B\mathbf{x})$. Thus $A(B\mathbf{x})$ is produced from $\mathbf{x}$ by a composition of mappings, the linear transformations. Our goal is to represent this composite mapping as a single mapping: $A(B\mathbf{x}) = (AB)\mathbf{x}$.

If A is $m \times n$, B is $n \times p$, and $\mathbf{x} \in \mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1, \dots, \mathbf{b}_p$ and the entries in $\mathbf{x}$ by $x_1, \dots, x_p$. Then

$$B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p$$

By the linearity of multiplication by A,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1) + \cdots + A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1 + \cdots + x_pA\mathbf{b}_p \end{aligned}$$

The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \dots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights. In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = [A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p]\mathbf{x}$$

Definition. *If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, {}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$. That is,*

$$AB = A[\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_p] = [A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p]$$

Multiplication of matrices corresponds to composition of linear transformations.

Each column of AB is a linear combination of the columns of A using weights from the corresponding columns of B.

If the product AB is defined, then the entry in row i and column j of AB is the sum of products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i,j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Example 1. *If* $C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}$ *and* $E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$, *compute:*

1. $2C - 3D$
2. $2C - 3E$
3. EC
4. CE

Theorem 2. *Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.*

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(B + C)A = BA + CA$
4. $r(AB) = (rA)B = A(rB)$
5. $I_m A = A = AI_n$

The left-to-right order in products is critical because the columns of AB are linear combinations of columns of A, whereas the columns of BA are linear combinations of columns of B. If $AB = BA$, we say that A and B commute with one another. In general,

1. $AB \neq BA$
2. $AB = AC \not\Rightarrow B = C$
3. $AB = 0 \not\Rightarrow A = 0$ or $B = 0$

If A is an $n \times n$ matrix and if k is a positive integer, then we define

$$A^k = \underbrace{AA \cdots A}_{k}$$

Since $A^0\mathbf{x} = \mathbf{x}$, we define A^0 as the identity matrix.

Given an $m \times n$ matrix A, the transpose of A, denoted by A^T, is the $n \times m$ matrix whose columns are formed from the corresponding rows of A.

Theorem 3. *Let A and B denote matrices whose sizes are appropriate for the following sums and products.*

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(rA)^T = rA^T$ *for any scalar* r
4. $(AB)^T = B^T A^T$

The transpose of a product of matrices equals the product of their transposes in reverse order.