# Matrix Operations

If $A$ is an $m \times n$ matrix, then the scalar entry $a_{ij}$ in the $i$th row and $j$th column of $A$ is called the $(i, j)$-entry of $A$. Notice that since the $j$th column of an $m \times n$ matrix $A$ is a list of $m$ real numbers, the $j$th column of $A$ identifies a vector $\mathbf{a}_j \in \mathbb{R}^m$ and, thus, the entry $a_{ij}$ can be thought of as the $i$th entry (from the top) in the $j$th column vector $\mathbf{a}_j$.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \tag{1}$$

The *main diagonal* of an $m \times n$ matrix $A$ is formed the *diagonal* entries $a_{11}, \ldots, a_{ij}$ of $A$, where $i = j$. A *diagonal matrix* is a square $n \times n$ matrix whose *non-diagonal* entries are zero. An $m \times n$ matrix whose entries are *all* zero a *zero matrix*, denoted by 0.

## Sums and Scalar Multiples

The arithmetic operations of vectors has a natural extension to matrices, being as the columns of an $m \times n$ matrix $A$ can be represented by the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. Two matrices $A$ and $B$ are *equal* if $A$ and $B$ are the same size and if their corresponding columns (that is, their corresponding entries) are equal. If $A$ and $B$ are $m \times n$ matrices, then the *sum* of $A$ and $B$ is the $m \times n$ matrix $A + B = \begin{bmatrix} \mathbf{a}_1 + \mathbf{b}_1 & \cdots & \mathbf{a}_n + \mathbf{b}_n \end{bmatrix}$ whose columns are the sums $\mathbf{a}_1 + \mathbf{b}_1, \ldots, \mathbf{a}_n + \mathbf{b}_n$ of the corresponding columns of $A$ and $B$. That is, each entry in the matrix $A + B$ is the sum of the corresponding entries in $A$ and $B$. Thus, the matrix $A + B$ is defined only when $A$ and $B$ are the same size.

**Example.** Let $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$.

The sum of $A$ and $B$ is the matrix $A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$, but the matrix $A + C$ is not defined because $A$ and $C$ are different sizes. ■

If $r$ is a scalar and $A$ is an $m \times n$ matrix, then the *scalar multiple* of $r$ and $A$ is the $m \times n$ matrix $rA$ whose columns are $r$ times the corresponding columns of $A$. For any matrix $A$, the matrix $-A = (-1)A$ and, therefore, the matrix $A - B = A + (-1)B$.

**Example.** Consider the matrices $A$ and $B$ defined in the preceding example.

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$

■

Each equality in the following theorem can be verified by first observing that the matrices $A$, $B$, and $C$ are the same size and then showing, through the analogous algebraic properties of vectors, that the columns of each matrix are equal.

**Theorem** (Algebraic Properties of Matrices).

Let $A$, $B$, and $C$ be matrices of the same size, and let $r$ and $s$ 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. $r(sA)=(rs)A$

## Matrix Multiplication

Consider the $m\times n$ matrix $A=\begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$, the $n\times k$ matrix $B=\begin{bmatrix}\mathbf{b}_1 & \cdots & \mathbf{b}_k\end{bmatrix}$, and the vector $\mathbf{x}\in\mathbb{R}^k$. Observe that since $B$ is $n\times k$ and $\mathbf{x}\in\mathbb{R}^k$, $\mathbf{x}$ can be left-multiplied by $B$ to transform $\mathbf{x}\in\mathbb{R}^k$ into $B\mathbf{x}\in\mathbb{R}^n$, which can be expressed by $B\mathbf{x}=x_1\mathbf{b}_1+\cdots+x_k\mathbf{b}_k$. Also observe that since $A$ is $m\times n$ and $B\mathbf{x}\in\mathbb{R}^n$, it follows that $B\mathbf{x}$ can be left-multiplied by $A$ to transform $B\mathbf{x}\in\mathbb{R}^n$ into $A(B\mathbf{x})\in\mathbb{R}^m$ and, since the matrix transformation $B\mathbf{x}\mapsto A(B\mathbf{x})$ is linear, it further follows that $A(B\mathbf{x})$ can be expressed by the following.

$$\begin{aligned}A(B\mathbf{x})&=A\left(B\mathbf{x}=x_1\mathbf{b}_1+\cdots+x_k\mathbf{b}_k\right)=A\left(x_1\mathbf{b}_1\right)+\cdots+A\left(x_k\mathbf{b}_k\right)\\&=x_1A\mathbf{b}_1+\cdots+x_kA\mathbf{b}_k=\begin{bmatrix}A\mathbf{b}_1 & \cdots & A\mathbf{b}_k\end{bmatrix}\mathbf{x}\end{aligned}$$

Thus, since $A(B\mathbf{x})=(AB)\mathbf{x}$ by the linearity of $B\mathbf{x}\mapsto A(B\mathbf{x})$, it follows that the matrix-product $AB$ is defined as follows.

**Definition** (The Matrix-Product $AB$).

If $A$ is an $m\times n$ matrix and $B$ is an $n\times k$ matrix with columns $\mathbf{b}_1,\dots,\mathbf{b}_k$, then the matrix-product $AB$ is the $m\times k$ matrix with columns $A\mathbf{b}_1,\dots,A\mathbf{b}_k$.

$$AB=A\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p\end{bmatrix}=\begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p\end{bmatrix}$$

*Note.* A vector $\mathbf{x}$ can be transformed into the vector $B\mathbf{x}$ through multiplication by $B$ which, in turn, can be transformed into the vector $A(B\mathbf{x})$ through multiplication by $A$. So as a consequence of the definition of $AB$, the equation $A(B\mathbf{x})=(AB)\mathbf{x}$ is true for all $\mathbf{x}\in\mathbb{R}^p$ and, thus, the multiplication of matrices corresponds to a *composition of linear transformations.*

**Example.** Compute $AB$, where $A=\begin{bmatrix}2 & 3\\1 & -5\end{bmatrix}$ and $B=\begin{bmatrix}4 & 3 & 6\\1 & -2 & 3\end{bmatrix}$.

Express $B$ as $B=\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix}$ and compute $A\mathbf{b}_1$, $A\mathbf{b}_2$, and $A\mathbf{b}_3$ as follows.

$$\begin{aligned}A\mathbf{b}_1&=\begin{bmatrix}2 & 3\\1 & -5\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}\\&=\begin{bmatrix}11\\-1\end{bmatrix}\end{aligned}\qquad\begin{aligned}A\mathbf{b}_2&=\begin{bmatrix}2 & 3\\1 & -5\end{bmatrix}\begin{bmatrix}3\\-2\end{bmatrix}\\&=\begin{bmatrix}0\\13\end{bmatrix}\end{aligned}\qquad\begin{aligned}A\mathbf{b}_3&=\begin{bmatrix}2 & 3\\1 & -5\end{bmatrix}\begin{bmatrix}6\\3\end{bmatrix}\\&=\begin{bmatrix}21\\-9\end{bmatrix}\end{aligned}$$

Thus $AB=A\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix}=\begin{bmatrix}11 & 0 & 21\\-1 & 13 & -9\end{bmatrix}$.

■

Notice that since the $j$th column $A\mathbf{b}_j$ of the matrix-product $AB$ is computed by multiplying the columns of $A$ by the corresponding entries in $\mathbf{b}_j$, the $j$th column $A\mathbf{b}_j$ is a linear combination of the columns of $A$ using the entries in $\mathbf{b}_j$ as weights.

**Fact** (Columns of the Matrix-Product $AB$).

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

By definition, for a linear combination of the form $A\mathbf{b}_j$ to exist, the number of columns in $A$ must match the number of rows in $\mathbf{b}_j$. Notice that by definition, $AB$ has the same number of *rows* as $A$ and the same number of *columns* as $B$.

**Example.** If $A$ is a $3 \times 5$ matrix and $B$ is a $5 \times 2$ matrix, what are the sizes of $AB$ and $BA$, if they are defined?

Since $A$ has 5 columns and $B$ has 5 rows, the the column $A\mathbf{b}_j$, and hence the matrix-product $AB$ is defined as a $3 \times 2$ matrix.

$$A \times B = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} = \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} = AB$$

The matrix-product $BA$ is not defined because the 2 columns of $A$ do not match the 3 rows of $B$. ■

If $A$ is an $m \times n$ matrix and $B$ is a $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then $A\mathbf{b}_j$ is the $j$th column of the matrix-product $AB$ and the $i$th entry of $A\mathbf{b}_j$ is the sum $a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$ of each entry $a_{i1}, \ldots, a_{in}$ in the $i$th row of $A$ multiplied by the corresponding entry $b_{1j}, \ldots, b_{nj}$ in $\mathbf{b}_j$, the $j$th column of $B$.

**Theorem** (Row-Column Rule for Computing $AB$).

If $A$ is a $m \times n$ matrix and $B$ is a $n \times p$ matrix, then the entry $(AB)_{ij}$ in the $i$th row and $j$th column of the $m \times p$ matrix $AB$ is the sum of the products of the corresponding entries from row $i$ of $A$ and column $j$ of $B$.

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$$

**Example.** Consider the matrices $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$, where the matrix-product $AB$ is defined. Use the row-column rule to compute $(AB)_{13}$ and $(AB)_{22}$. That is, the entry in the 2nd row and 3rd column of $AB$, and the entry in the 2nd row and 2nd column of $AB$.

The entry $(AB)_{13}$ is computed by multiplying each entry $a_{11}, a_{12}$ in row 1 of $A$ by the corresponding entry $b_{13}, b_{23}$ in column 3 of $B$ and then adding each product.

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

The entry $(AB)_{22}$ is computed by multiplying each entry $a_{21}, a_{22}$ in row 2 of $A$ by the corresponding entry $b_{12}, b_{22}$ in column 2 of $B$ and then adding each product.

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

■

Observe that as a consequence of the row-column rule, since the entries in $\text{row}_i(AB)$ come from $\text{row}_i(A)$ and the columns of $B$, and since $\text{row}_i(A)$ is $1 \times n$ and $B$ is $n \times p$, $\text{row}_i(AB)$ can be computed by right-multiplying the $\text{row}_i(A)$ by $B$.

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B \tag{2}$$

## Properties of Matrix Multiplication

Recall that $I_m$ represents the $m \times m$ identity matrix and $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^m$.

**Theorem** (Properties of Matrix Multiplication).

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

a. $A(BC) = (AB)C$ (associative law of multiplication)

b. $A(B + C) = AB + AC$ (left distributive law)

c. $(B + C)A = BA + CA$ (right distributive law)

d. $r(AB) = (rA)B = A(rB)$ for any scalar $r$

e. $I_m A = A = AI_n$ (identity for matrix multiplication)

The associative and distributive properties of matrix multiplication say that the parentheses in matrix expressions can be inserted and removed in the same way that parentheses can be inserted and removed in the algebra of real numbers as long as the left-to-right order of the matrices is preserved. The left-to-right order must be preserved because the columns of $AB$ are linear combinations of the columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of $A$ using the corresponding entries $b_1, \ldots, b_p$ in $B$ as weights, whereas the columns of $BA$ are constructed in the opposite manner. In the case that $AB = BA$, the matrices $A$ and $B$ are said to **commute** each other.

## Powers of a Matrix

If $A$ is an $n \times n$ matrix and $k \in \mathbb{Z}^+$, then $A^k$ denotes the product (3) which, by the associative law of matrix multiplication, may be computed as $A^k = A \cdot A^{k-1}$ or $A^k = A^{k-1} \cdot A$.

$$A^k = \underbrace{A \cdots A}_{k} \tag{3}$$

## The Transpose of a Matrix

For any $m \times n$ matrix $A$, the **transpose** of $A$ is the $n \times m$ matrix $A^T$ which is formed by interchanging the rows and columns of $A$.

**Example.** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. ■

**Theorem** (Properties of the Transpose).

Let $A$ and $B$ be matrices with sizes for which the indicated sums are products are defined.

a. $\left(A^T\right)^T = A$

b. $(A + B)^T = A^T + B^T$

c. $(rA)^T = rA^T$ for any scalar $r$

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