# Matrix multiplication

**Matrix multiplication** is an operation between two matrices that creates a new matrix such that given two matrices $\mathbf{A}$ and $\mathbf{B}$, each column of the product $\mathbf{AB}$ is formed by multiplying $\mathbf{A}$ by each column of $\mathbf{B}$ (Definition 1).

   Note that this definition requires that if we multiply an $m \times n$ matrix by a $n \times p$ matrix, the result will be an $m \times p$ matrix where the number of rows is determined by the first matrix and the number of columns is determined by the second matrix. More succinctly:

$$\mathbf{A}_{m \times n} \, \mathbf{B}_{n \times p} = (\mathbf{AB})_{m \times p}$$

---

**Definition 1** *Matrix multiplication between an $m \times n$ matrix $\mathbf{A}$ and a $n \times p$ matrix $\mathbf{B}$ is given by*

$$\mathbf{AB} := \begin{bmatrix} \mathbf{Ab}_{*,1} & \mathbf{Ab}_{*,2} & \ldots & \mathbf{Ab}_{*,n} \end{bmatrix}$$

---

## Intuition

There are several perspectives for which one can view matrix multiplication each depending on the perspective taken on each of the matrix factors. Recall we can view a matrix via a number of perspectives including

1. As an ordered list of column vectors

2. As an ordered list of row vectors

3. As a linear transformation

These views of matrices lead to the following views of matrix multiplication:

1. *Matrix multiplication as the linear transformation of a set of vectors*: This perspective follows from viewing $\mathbf{A}$ as a linear transformation and $\mathbf{B}$ as an ordered list of column-vectors.

2. *Matrix multiplication as a composition of linear transformations*: This perspective follows from viewing both $\mathbf{A}$ and $\mathbf{B}$ as linear transformations.

3. *Matrix multiplication as the computation of all pair-wise dot products between two lists of vectors*: This perspective follows from viewing $\mathbf{A}$ as an ordered list of row-vectors and viewing $\mathbf{B}$ as an ordered list of column-vectors.

These perspectives are discussed in more depth in the following sections.

### 1. Matrix multiplication as the linear transformation of a set of vectors

The most obvious way to view matrix multiplication is as simply performing a linear transformation on a set of vectors. That is, we view $\mathbf{A}$ as a linear transformation and we view the matrix $\mathbf{B}$ as an ordered list of column vectors

$$\mathbf{B} := \left[ \mathbf{b}_{*,1}, \mathbf{b}_{*,2}, \ldots, \mathbf{b}_{*,n} \right]$$

Then, we form the matrix $\mathbf{AB}$ by taking the linear transformation characterized by $\mathbf{A}$ of each $\mathbf{b}_{*,i}$.

### 2. Matrix multiplication as a composition of linear transformations

Alternatively if we view *both* $\mathbf{A}$ and $\mathbf{B}$ as linear transformations, then the matrix $\mathbf{AB}$ is the matrix that characterizes the composition of the linear transformations characterized by $\mathbf{A}$ and $\mathbf{B}$ (Theorem 1). That is, given two linear transformations

$$T_{\mathbf{A}}(\mathbf{x}) := \mathbf{A}\mathbf{x}$$
$$T_{\mathbf{B}}(\mathbf{x}) := \mathbf{B}\mathbf{x}$$

the matrix resulting from multiplying $\mathbf{A}$ with $\mathbf{B}$, denoted $\mathbf{AB}$, is the matrix that characterizes the composition $T_{\mathbf{A}} \circ T_{\mathbf{B}}$. That is, $\mathbf{AB}$ characterizes the linear transformation:

$$T_{\mathbf{AB}}(\mathbf{x}) := \mathbf{A}(\mathbf{B}\mathbf{x})$$

This concept is illustrated in Figure 1.

Recall that a matrix's number of rows determines the dimensions of the vectors in its range and the number of columns correspond to the number of dimensions the domain. Since $\mathbf{AB}$ characterizes the composition $T_{\mathbf{A}} \circ T_{\mathbf{B}}$, it follows that the matrix $\mathbf{AB}$ will map from the domain of $\mathbf{B}$ to the range of $\mathbf{A}$.

---

**Theorem 1** *Matrix multiplication between an $m \times n$ matrix $\mathbf{A}$ and a $n \times p$ matrix $\mathbf{B}$ results in a matrix $\mathbf{AB}$ such that given a vector $\mathbf{x} \in \mathbb{R}^m$, the following holds:*
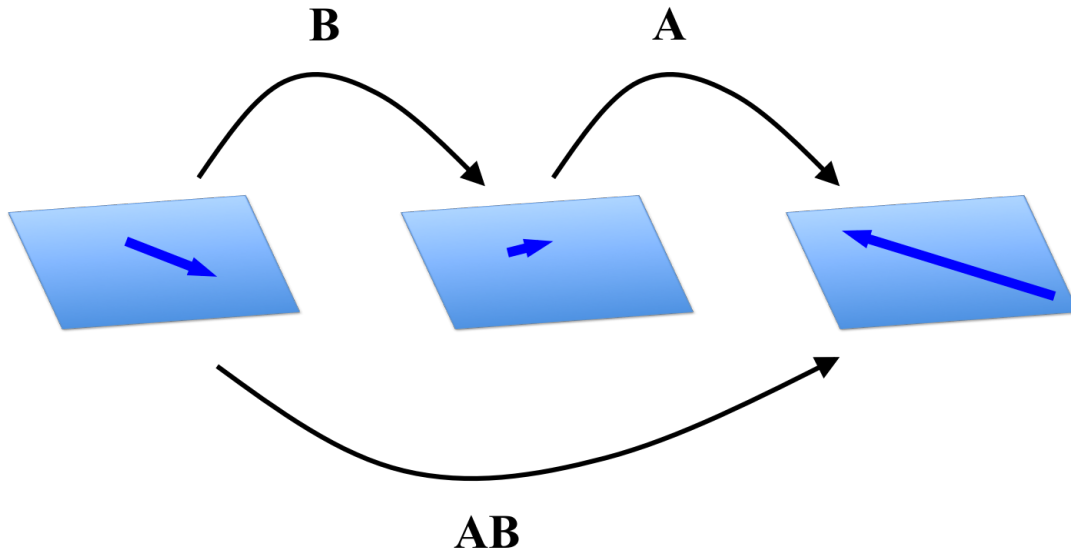
$$(\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{x})$$

**Proof:**

---

Figure 1: The matrix **AB** performs the mapping by **B** followed by **A**.

First, let's expand **Bx**.

$$\mathbf{Bx} = \mathbf{b}_{*,1}x_1 + \mathbf{b}_{*,2}x_2 + \cdots + \mathbf{b}_{*,n}x_n$$

$$= \begin{bmatrix} b_{1,1}x_1 + b_{1,2}x_2 + \cdots + b_{1,n}x_n \\ b_{2,1}x_1 + b_{2,2}x_2 + \cdots + b_{2,n}x_n \\ \vdots \\ b_{n,1}x_1 + b_{n,2}x_2 + \cdots + b_{n,n}x_n \end{bmatrix}$$

Now,

$$(\mathbf{AB})\mathbf{x} = \begin{bmatrix} \mathbf{Ab}_{*,1}, \mathbf{Ab}_{*,2}, \dots, \mathbf{Ab}_{*,n} \end{bmatrix} \mathbf{x}$$

$$= \mathbf{Ab}_{*,1}x_1 + \mathbf{Ab}_{*,2}x_2 + \cdots + \mathbf{Ab}_{*,n}x_n$$

$$= (\mathbf{a}_{*,1}b_{1,1} + \dots \mathbf{a}_{*,n}b_{n,1})\,x_1 + (\mathbf{a}_{*,1}b_{1,2} + \dots \mathbf{a}_{*,n}b_{n,2})\,x_2 + \cdots + (\mathbf{a}_{*,1}b_{1,n} + \dots \mathbf{a}_{*,n}b_{n,n})\,x_n$$

$$= (\mathbf{a}_{*,1}b_{1,1}x_1 + \dots \mathbf{a}_{*,n}b_{n,1}x_1) + (\mathbf{a}_{*,1}b_{1,2}x_2 + \dots \mathbf{a}_{*,n}b_{n,2}x_2) + \cdots + (\mathbf{a}_{*,1}b_{1,n}x_n + \dots \mathbf{a}_{*,n}b_{n,n}x_n)$$

$$= \mathbf{a}_{*,1}(b_{1,1}x_1 + \cdots + b_{1,n}x_n) + \mathbf{a}_{*,2}(b_{2,1}x_1 + \cdots + b_{2,n}x_n) + \cdots + \mathbf{a}_{*,n}(b_{n,1}x_1 + \cdots + b_{n,n}x_n)$$

$$= \mathbf{a}_{*,1}(\mathbf{Bx})_1 + \mathbf{a}_{*,2}(\mathbf{Bx})_2 + \dots \mathbf{a}_{*,n}(\mathbf{Bx})_n$$

$$= \mathbf{A}(\mathbf{Bx})$$

□

## 3. Matrix multiplication as the computation of all pair-wise dot products between two lists of vectors

If we view the matrix $\mathbf{A}$ as an ordered list of row-vectors and the matrix $\mathbf{B}$ as an ordered list of column vectors, then the product $\mathbf{AB}$ is the matrix that stores all of the pairwise dot products of the vectors in $\mathbf{A}$ and $\mathbf{B}$. That is, the $i,j$th element of $\mathbf{AB}$ is the the dot product of the $i$th row of $\mathbf{A}$ and the $j$th column of $\mathbf{B}$ (Theorem 2). This fact, known as **row-column rule**, can be used for computing each element of $\mathbf{AB}$. Figure 2 demonstrates this view of matrix multiplication.
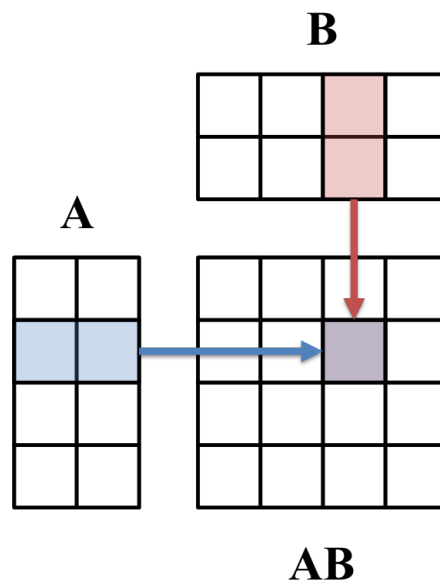


Figure 2: The $i,j$th element of $\mathbf{AB}$ is the dot product of the $i$th row of $\mathbf{A}$ and $j$th column of $\mathbf{B}$.

---

**Theorem 2** *The $i$, $j$th element of $\mathbf{AB}$ is computed by*

$$\mathbf{a}_{i,*} \cdot \mathbf{b}_{*,j}$$

---

4

**Proof:**

$$(\mathbf{AB}) = \left[\mathbf{Ab}_{*,1}, \mathbf{Ab}_{*,2}, \ldots, \mathbf{Ab}_{*,n}\right]$$

$$= \left[\mathbf{a}_{*,1}b_{1,1} + \cdots + \mathbf{a}_{*,n}b_{n,1}, \quad \ldots \quad \mathbf{a}_{*,1}b_{1,n} + \cdots + \mathbf{a}_{*,n}b_{n,n}\right]$$

$$= \begin{bmatrix} a_{1,1}b_{1,1} + \cdots + a_{1,n}b_{n,1} & \ldots & a_{1,1}b_{1,n} + \cdots + a_{1,n}b_{n,n} \\ a_{2,1}b_{1,1} + \cdots + a_{2,n}b_{n,1} & \ldots & a_{2,1}b_{1,n} + \cdots + a_{2,n}b_{n,n} \\ \vdots & \ddots & \vdots \\ a_{n,1}b_{1,1} + \cdots + a_{n,n}b_{n,1} & \ldots & a_{n,1}b_{1,n} + \cdots + a_{n,n}b_{n,n} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{a}_{1,*} \cdot \mathbf{b}_{*,1} & \ldots & \mathbf{a}_{1,*} \cdot \mathbf{b}_{*,n} \\ \mathbf{a}_{2,*} \cdot \mathbf{b}_{*,1} & \ldots & \mathbf{a}_{2,*} \cdot \mathbf{b}_{*,n} \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{m,*} \cdot \mathbf{b}_{*,1} & \ldots & \mathbf{a}_{m,*} \cdot \mathbf{b}_{*,n} \end{bmatrix}$$

□

## Properties of matrix multiplication

The following properties hold for matrix multiplication:

1. **Associative law for matrices** (Theorem 3)

$$\mathbf{A(BC) = (AB)C}$$

2. **Commutative property of scalars** (Theorem 4)

$$r(\mathbf{AB}) = (r\mathbf{A})\mathbf{B} = \mathbf{A}r\mathbf{B}$$

   where $r$ is a scalar.

3. **Left distributive law** (Theorem 5)

$$\mathbf{A(B + C) = AB + AC}$$

4. **Right distributive law** (Theorem 6)

$$\mathbf{(B + C)A = BA + CA}$$

5. **Identity** (Theorem 7)

$$\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{AI}_n$$

**Theorem 3** *Given matrices* $\mathbf{A} \in \mathbb{R}^{m \times l}$, $\mathbf{B} \in \mathbb{R}^{l*p}$, *and* $\mathbf{C} \in \mathbb{R}^{p \times n}$, *the following holds:*

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

**Proof:**

Since matrix-multiplication can be understood as a composition of functions, and since compositions of functions are associative, it follows that matrix-multiplication is associative

---

**Theorem 4** *Given matrices* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\mathbf{B} \in \mathbb{R}^{n*p}$, *the following holds:*

$$r(\mathbf{AB}) = (r\mathbf{A})\mathbf{B} = \mathbf{A}(r\mathbf{B})$$

**Proof:**

First we prove $r(\mathbf{AB}) = (r\mathbf{A})\mathbf{B}$:

$$
\begin{aligned}
r(\mathbf{AB}) &= r\begin{bmatrix} \mathbf{Ab}_{*,1} & \dots & \mathbf{Ab}_{*,p} \end{bmatrix} \\
&= \begin{bmatrix} r\mathbf{Ab}_{*,1} & \dots & r\mathbf{Ab}_{*,p} \end{bmatrix} \\
&= (r\mathbf{A})\mathbf{B}
\end{aligned}
$$

Next, we prove $r(\mathbf{AB}) = \mathbf{A}(r\mathbf{B})$:

$$
\begin{aligned}
r(\mathbf{AB}) &= r\begin{bmatrix} \mathbf{Ab}_{*,1} & \dots & \mathbf{Ab}_{*,p} \end{bmatrix} \\
&= \begin{bmatrix} r\mathbf{Ab}_{*,1} & \dots & r\mathbf{Ab}_{*,p} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}(r\mathbf{b}_{*,1}) & \dots & \mathbf{A}(r\mathbf{b}_{*,p}) \end{bmatrix} \quad \text{linearity of matrix-vector multiplication} \\
&= \mathbf{A}(r\mathbf{B})
\end{aligned}
$$

$\square$

---

**Theorem 5** *Given matrices* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n*p}$, *and* $\mathbf{C} \in \mathbb{R}^{n \times p}$, *the following holds:*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

**Proof:**

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \begin{bmatrix} \mathbf{A}(\mathbf{b}_{*,1}+\mathbf{c}_{*,1}) & \ldots & \mathbf{A}(\mathbf{b}_{*,p}+\mathbf{c}_{*,p}) \end{bmatrix} \qquad \text{definition of matrix multiplication}$$

$$= \begin{bmatrix} (\mathbf{A}\mathbf{b}_{*,1}+\mathbf{A}\mathbf{c}_{*,1}) & \ldots & (\mathbf{A}\mathbf{b}_{*,p}+\mathbf{A}\mathbf{c}_{*,p}) \end{bmatrix} \qquad \text{linearity of matrix-vector multiplication}$$

$$= \begin{bmatrix} \mathbf{A}\mathbf{b}_{*,1} & \ldots & \mathbf{A}\mathbf{b}_{*,p} \end{bmatrix} + \begin{bmatrix} \mathbf{A}\mathbf{c}_{*,1} & \ldots & \mathbf{A}\mathbf{c}_{*,p} \end{bmatrix} \qquad \text{definition of matrix-addition}$$

$$= \mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{C} \qquad \text{definition of matrix multiplication}$$

□

---

**Theorem 6** *Given matrices* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n*p}$, *and* $\mathbf{C} \in \mathbb{R}^{n \times p}$, *the following holds:*

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{B}\mathbf{A} + \mathbf{C}\mathbf{A}$$

**Proof:**

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \begin{bmatrix} (\mathbf{B}+\mathbf{C})\mathbf{a}_{*,1} & \ldots & (\mathbf{B}+\mathbf{C})\mathbf{a}_{*,p} \end{bmatrix} \qquad \text{definition of matrix-matrix multiplication}$$

$$= \begin{bmatrix} (\mathbf{B}\mathbf{a}_{*,1}+\mathbf{C}\mathbf{a}_{*,1}) & \ldots & (\mathbf{B}\mathbf{a}_{*,p}+\mathbf{C}\mathbf{a}_{*,p}) \end{bmatrix} \qquad \text{definition of matrix addition}$$

$$= \begin{bmatrix} \mathbf{B}\mathbf{a}_{*,1} & \ldots & \mathbf{B}\mathbf{a}_{*,p} \end{bmatrix} + \begin{bmatrix} \mathbf{C}\mathbf{a}_{*,1} & \ldots & \mathbf{C}\mathbf{a}_{*,p} \end{bmatrix}$$

$$= \mathbf{B}\mathbf{A} + \mathbf{C}\mathbf{A}$$

□

---

**Theorem 7** *Given an* $m \times n$ *matrix* $\mathbf{A}$, *the following holds:*

$$\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$$

**Proof:**

By the fact that an identity function simply maps each element in its domain back to itself. It follows that the composition of a function $f$ and identity function is simply the function $f$. Thus, it follows that any matrix multiplied on the left or right by the identity matrix returns the original matrix. Thus, $\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_n$.

□