3 Matrix Operations

There are several operations one can apply to a matrix. Addition and scalar multiplication behave as you would expect (just like in vectors), but matrix multiplication and its counterpart inverses are more interesting.

3.1 Basic Matrix Operations

We need some notation. For matrix $A$, the notation $a_{ij}$ means the entry in row $i$ and column $j$ of $A$.

Matrix addition requires that the two matrices have the same dimensions. The sum is defined by adding the corresponding entries. For example,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{bmatrix}$$

Similarly, scalar multiplication is defined entry-wise. For example,

$$c \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{bmatrix}$$

Another operation is the transpose:

The transpose of a matrix $A$, denoted $A^T$, exchanges rows and columns. That is, $(A^T)_{ij} = A_{ji}$.

Example.

The transpose of $\begin{bmatrix} 3 & 4 & 7 \\ -2 & 5 & -3 \end{bmatrix}$ is $\begin{bmatrix} 3 & -2 \\ 4 & 5 \\ 7 & -3 \end{bmatrix}$

A square matrix has equal number of rows and columns. The diagonal of a square matrix runs from top-left to bottom-right. A symmetric matrix is a square matrix which is symmetric around its diagonal. In other words, $A = A^T$. 
3.2 Matrix Multiplication

An important operation is matrix multiplication. Matrix multiplication produces a matrix. Only matrices of compatible sizes can be multiplied. One way to define matrix-matrix multiplication is in terms of matrix-vector multiplication:

If matrix $A$ is $m \times n$ and matrix $B$ is $r \times s$, then for the product $AB$ to be valid it must be that $n = r$. If valid, the product $AB$ has size $m \times s$. The columns of the product are the results of multiplying the first matrix by the columns of the second. That is,

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_s \end{bmatrix}$$

where $b_j$ is the $j^{th}$ column of $B$.

**Example.** Here is the product of a $2 \times 3$ and $3 \times 4$ matrix:

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 & 5 \\ -2 & 0 & 3 & -4 \\ 1 & -2 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 3 & -2 \\ -8 & 4 & 5 & -10 \end{bmatrix}$$

An example detail: the 3rd column of the result is given by $-1 \times 1 + 3 \times 2 + (-2) \times 2 = 5$.

Equivalently, one can define matrix multiplication without reference to vectors. There is a formula for each entry in the product; namely the sum

$$(AB)_{ij} = \sum_k a_{ik}b_{kj}$$

That is, to calculate the entry in row $i$ and column $j$ of the product, look at row $i$ of the first matrix and column $j$ of the second matrix; then multiply corresponding entries and add. This is illustrated here:

**Example.** The entry in the 2nd row 3rd column of the previous example is calculated as $0 \times (-1) + 3 \times 3 + (-2) \times 2 = 5$. 

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ 0 & 3 & -2 \end{bmatrix} \begin{bmatrix} \cdots & \cdots & -1 & \cdots \\ \cdots & \cdots & 3 & \cdots \\ \cdots & \cdots & 2 & \cdots \end{bmatrix} = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 5 & \cdots \end{bmatrix}$$
### 3.3 Properties of Matrix Multiplication

It is important to note two facts about multiplication:

1. Matrix multiplication is **associative**. That is, brackets don’t matter. For example \((AB)C = A(BC)\) (and the one product is valid whenever the other one is).

2. However, matrix multiplication is not **commutative**. That is, order matters. There is no guarantee that \(AB = BA\). Indeed, the one product might be valid when the other one is not. Even if \(A\) and \(B\) are square matrices of the same size, so that both products are defined and the results have the same size, there is no guarantee (and indeed it is unlikely that) the two products are the same.

A **diagonal matrix** is a square matrix that has zeros off the diagonal (and might or might not have zeroes on the diagonal). The **identity matrix** \(I_n\) is the \(n \times n\) diagonal matrix with 1’s on the diagonal. (We sometimes write just \(I\).) Its columns are the vectors \(e_i\): these have 0’s in every position except for a 1 in the \(i\)th position.

**Example.** If \(A\) is a square matrix, then \(IA = AI = A\), where \(I\) is the identity matrix of the same size. (This is why \(I\) is called the identity matrix.)

We use \(A^p\) to mean the product of \(p\) copies of \(A\). (This needs \(A\) to be square.)

**Example.** If \(D\) is a diagonal matrix, then \(D^p\) is a diagonal matrix with the entries taken to power \(p\). For example:

\[
\begin{bmatrix}
3 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{bmatrix}^{100} = \begin{bmatrix}
3^{100} & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

Using the above formula for the entries of the product, the following fact about transposes can be shown. Note that the order is swapped!

**Fact 3.1** \((AB)^T = B^T A^T\)

We conclude this section with the following fact.

**Fact 3.2** Each elementary row operation is equivalent to multiplying on the left by a matrix called an **elementary matrix**.
Instead of providing a proof, we just give an example of each type.

**Example.** If $A$ is a $2 \times 2$ matrix, then:

- **left-multiplication by** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **interchanges the first and second row;**
- **left-multiplication by the diagonal matrix** $\begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$ **divides the second row by 3; and**
- **left-multiplication by** $\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$ **subtracts twice the first row from the second.**

### 3.4 Matrix Transforms

A function/transform/mapping has a **domain** and **codomain**: the domain specifies the possible inputs and the codomain specifies the possible outputs. The function maps each element of the domain to an **image** in the codomain. The **range** is the set of all images: it is a subset of the codomain.

A transform is called **onto** if the range is the whole codomain. A transform is called **one-to-one** if every vector in the range is the image of exactly one vector. (An onto transform is sometimes called a **surjection** and a one-to-one transform an **injection**.)

If the matrix $A$ is $m \times n$, then the matrix transform $x \mapsto Ax$ has domain $\mathbb{R}^n$, codomain $\mathbb{R}^m$, and range some subset of $\mathbb{R}^m$.

**Example.** For example, if $A$ is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then the transform maps $(5, 3)$ to $(3, 5)$.

**Example.** Examples of transforms include:

- **projections**, such as $P = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
- **shears**, such as $S = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$
- **contractions/dilations**, such as $C = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}$
- **rotations**, such as $R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$

The reasons for the names can be seen by observing the effect of these transforms on sets of points in the plane.
One useful property when applying matrix transforms is that the composition of transforms (meaning applying one transform after another) is equivalent to matrix multiplication. For example, in \( \mathbb{R}^2 \) to rotate by \( \theta \) and then contract by a factor of 3, transform by the matrix product \( CR \), where the matrices \( C \) and \( R \) are as in the above example.

**Practice**

3.1. Compute \( XY \), \( X + Y \), \( YZ \), \( Y + Z \), and \( ZX \), when they exist, for the following the matrices:

\[
X = \begin{bmatrix} -3 & 1 & 4 \\ 1 & 2 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 5 & -4 \\ 3 & -1 \end{bmatrix} \quad Z = \begin{bmatrix} -2 & 7 \\ 5 & -1 \end{bmatrix}
\]

3.2. Suppose that \( A \) is an \( n \times n \) matrix and \( D \) is an \( n \times n \) diagonal matrix. Explain what the two products \( AD \) and \( DA \) look like.

3.3. Calculate \( R^{50} \) where \( R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \)

---

Solutions to Practice Exercises

3.1. \( XY \) and \( X + Y \) do not exist.

\[
YZ = \begin{bmatrix} -30 & 39 \\ -11 & 22 \end{bmatrix} \quad Y + Z = \begin{bmatrix} 3 & 3 \\ 8 & -2 \end{bmatrix} \quad ZX = \begin{bmatrix} 13 & 12 & -8 \\ -16 & 3 & 20 \end{bmatrix}
\]

3.2. The product \( AD \) has each column of \( A \) scaled by the corresponding entry in \( D \).

The product \( DA \) has each row of \( A \) scaled by the corresponding entry in \( D \).

3.3. \( R^3 = I \). Therefore \( R^{50} = R^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \).