Review of Some Matrix Definitions

1. Matrix

A **matrix** is a rectangular array of numbers. An $n \times m$ matrix contains n rows and m columns. Consider these matrices:

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 2 & -2 \\ 1 & 0 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ 0 & 3 \end{pmatrix}$$

$\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices, $\mathbf{C}$ is a 2×3 matrix, and $\mathbf{D}$ is a 3×2 matrix.

2. Addition

To add two matrices, they must have the same number of rows and the same number of columns. These matrices are then added componentwise:

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \dots & a_{nm}+b_{nm} \end{pmatrix}$$

The only two matrices that can be added in Section 1 are $\mathbf{A}$ and $\mathbf{B}$ because they are both 2×2.

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2+2 & 0-2 \\ -2+1 & 1+0 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -1 & 1 \end{pmatrix}$$

2. Scalar Multiplication

Scalar multiplication means multiplying a number times a matrix. For example, if the matrix $\mathbf{C}$ is defined as in Section 1, compute $3\mathbf{C}$:

$$3\mathbf{C} = 3\begin{pmatrix} 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3\cdot 3 & 3\cdot 1 & 3\cdot 0 \\ 3(-2) & 3\cdot 0 & 3\cdot 1 \end{pmatrix} = \begin{pmatrix} 9 & 3 & 0 \\ -6 & 0 & 3 \end{pmatrix}$$

3. Matrix Multiplication

Two matrices **A** and **B** can only multiplied if the number of columns in **A** equals the number of rows in **B**. Thus, of the matrices defined in Section 1, only these matrix products are permissible: **AB**, **BA**, **CD**, and **DC**.

Given an $n \times m$ matrix A and an $m \times p$ matrix B, the product is the $n \times p$ matrix defined as:

$$
\begin{aligned}
\mathbf{AB} &= \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mp} \end{pmatrix} && (1) \\
&= \begin{pmatrix} \sum_{i=1}^{m} a_{1i} b_{i1} & \dots & \sum_{i=1}^{m} a_{1i} b_{ip} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} a_{ni} b_{i1} & \dots & \sum_{i=1}^{m} a_{1i} b_{ip} \end{pmatrix} && (2)
\end{aligned}
$$

An easy way to understand matrix multiplication is to note that the ij-th component of the product matrix is the product of row i from A with column j from B:

$$a_{i1}b_{1j} + a_{21}b_{2j} + \cdots + a_{im}b_{mj}$$

For example:

$$
\begin{aligned}
\mathbf{CD} &= \begin{pmatrix} 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ 3 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 3 \cdot 3 + 1(-2) + 0 \cdot 3 & 3 \cdot 1 + 1 \cdot 0 + 0 \cdot 1 \\ 3(-2) + 0(-2) + 1 \cdot 3 & 1(-2) + 0 \cdot 0 + 1 \cdot 1 \end{pmatrix} \\
&= \begin{pmatrix} 7 & 3 \\ -3 & -1 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{DC} &= \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 3\cdot 3 + 1(-2) & 3\cdot 1 + 1\cdot 0 & 3\cdot 0 + 1\cdot 1 \\ (-2)3 + 0(-2) & (-2)1 + 0\cdot 0 & (-2)0 + 0\cdot 1 \\ 3\cdot 3 + 1(-2) & 3\cdot 1 + 1\cdot 0 & 3\cdot 0 + 1\cdot 1 \end{pmatrix} \\
&= \begin{pmatrix} 7 & 3 & 1 \\ -6 & -2 & 0 \\ 6 & 3 & 1 \end{pmatrix}
\end{aligned}
$$

3. Transpose

To compute the **transpose** $\mathbf{A}^T$ of a matrix $\mathbf{A}$, change the rows to columns and columns to rows:

$$
\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}^T = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \dots & a_{nm} \end{pmatrix}
$$

For example:

$$
\mathbf{C}^T = \begin{pmatrix} 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 3 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

A matrix $\mathbf{A}$ is **symmetric** if $\mathbf{A}^T = \mathbf{A}$. Here is an example of a symmetric matrix:

$$
\begin{pmatrix} 7 & 2 & -3 \\ 2 & 11 & 5 \\ -3 & 5 & 4 \end{pmatrix}
$$

5. Identity Matrices

The $n \times n$ **identity matrix** is an $n \times n$ matrix with ones on the diagonal, but zeros everywhere else.

Here are the identity matrices of sizes 2×2, 3×3, and 4×4:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

An identity matrix $\mathbf{I}$ is used in theoretical calculations because $\mathbf{AI} = \mathbf{A}$ and $\mathbf{IB} = \mathbf{B}$. The size of $\mathbf{I}$ in each identity is chosen to be compatible for multiplication. Verify that, for $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ defined as in Section 1,

$$\mathbf{IA} = \mathbf{A} \quad \mathbf{AI} = \mathbf{A} \quad \mathbf{IB} = \mathbf{B} \quad \mathbf{BI} = \mathbf{B}$$
$$\mathbf{IC} = \mathbf{C} \quad \mathbf{CI} = \mathbf{C} \quad \mathbf{ID} = \mathbf{D} \quad \mathbf{DI} = \mathbf{D}$$

The $n \times m$ **zero matrix** is the analogy of the identity matrix for addition. It is an $n \times m$ matrix filled entirely with zeros. Here is a 3×4 zero matrix:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

5. Inverse Matrices

A matrix $\mathbf{A}^{-1}$ is the **inverse** of the matrix $\mathbf{A}$ if $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{I}$ and $\mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$. The inverse of a 2×2 matrix can be computed with the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ab - dc} \begin{pmatrix} a & -c \\ -b & d \end{pmatrix}$$

For example:

$$\begin{aligned} \mathbf{A}^{-1} &= \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix}^{-1} = \frac{1}{2 \cdot 1 - 0(-2)} \begin{pmatrix} 2 & -(-2) \\ -0 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0.5 \end{pmatrix} \end{aligned}$$

Verify that $\mathbf{A}^{-1}$ is actually the inverse of $\mathbf{A}$:

$$\mathbf{A}\mathbf{A}^{-1} = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0.5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\mathbf{A}^{-1}\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 0.5 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A matrix is called **invertable** if its inverse exists. A 2×2 matrix is invertible if $ad - cd \neq 0$. A matrix must be square to be invertable, but not all square matrices are invertable.

Here is the formula for the inverse of a 3×3 matrix: if

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

then

$$A^{-1} = \frac{1}{aei + bfg + cdh - afh - bdi - ceg} \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}$$

Similar formulas exist for the inverses of larger matrices, although they are not computationally efficient. Computer software, such as R, uses more efficient algorithms for computing matrix inverses.

6. Some Matrix Identities

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $c(\mathbf{A} + \mathbf{B}) = c\,\mathbf{A} + c\,\mathbf{B}$
4. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{A}\,\mathbf{B} + \mathbf{A}\,\mathbf{C}$
5. $(\mathbf{A} + \mathbf{B})\,\mathbf{C} = \mathbf{A}\,\mathbf{C} + \mathbf{B}\,\mathbf{C}$
6. $(\mathbf{A}\,\mathbf{B})\,\mathbf{C} = \mathbf{A}(\mathbf{B}\,\mathbf{C})$

7. $\mathbf{0} + \mathbf{A} = \mathbf{A} = \mathbf{A} + \mathbf{0}$
8. $\mathbf{I}\,\mathbf{A} = \mathbf{A} = \mathbf{A}\,\mathbf{I}$
9. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
10. $(\mathbf{A}^T)^T = \mathbf{A}$
11. $(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$