Lecture Notes: Matrix Definitions and Operations

Yufei Tao
Department of Computer Science and Engineering
Chinese University of Hong Kong
taoyf@cse.cuhk.edu.hk

1 Matrix Definitions

An $m \times n$ *matrix* is defined as m rows of real numbers, where each row has length n. To represent a matrix, we typically write out all these numbers in a 2d array, enclosed by a pair of square brackets, e.g.:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 8 & 4 & 2 \end{bmatrix} \tag{1}$$

is a 4×3 matrix. We will use capitalized bold symbols to denote arrays, e.g., $\boldsymbol{A}$. The values m and n are called the *dimensions* of $\boldsymbol{A}$.

When the dimensions m, n are clear, we sometimes use the notation $\boldsymbol{A} = [a_{ij}]$ to define a_{ij}, which refers to the number at the i-th row and j-th column of $\boldsymbol{A}$, with $i \in [1, m]$ and $j \in [1, n]$. For example, if A is the array in (1), then $a_{12} = 2$ whereas $a_{21} = 3$.

A *vector* is a matrix that has only one row or one column, namely, either $m = 1$ or $n = 1$. More specifically, a $1 \times n$ matrix is a *row vector*, while an $m \times 1$ matrix is a *column vector*. For example, let $\boldsymbol{A}$ be the matrix in (1). Then, the 3rd row of $\boldsymbol{A}$ is a row vector $[6, 7, 8]$, while the 2nd column is a column vector:

$$\begin{bmatrix} 2 \\ 4 \\ 7 \\ 4 \end{bmatrix}.$$

If $m = n$, $\boldsymbol{A}$ is a *square matrix*, e.g.

$$\begin{bmatrix} 1 & 2 & 3 & 8 \\ 3 & 4 & 5 & 2 \\ 6 & 7 & 8 & 3 \\ 8 & 4 & 2 & 4 \end{bmatrix}. \tag{2}$$

When $\boldsymbol{A} = [a_{ij}]$ is an $n \times n$ square matrix, we refer to the sequence $a_{11}, a_{22}, ..., a_{nn}$ as the *main diagonal* (or just *diagonal* for short). For example, if $\boldsymbol{A}$ is the matrix in (2), then its main diagonal is the sequence $1, 4, 8, 4$.

Again, let $\boldsymbol{A} = [a_{ij}]$ be a square matrix. Then, we say that

- $\boldsymbol{A}$ is *symmetric* if it always holds that $a_{ij} = a_{ji}$;
- $\boldsymbol{A}$ is *skew-symmetric* if it always holds that $a_{ij} = -a_{ji}$.

It is easy to see that $\boldsymbol{A}$ is skew-symmetric, then its main diagonal consists of only 0's. For example,

$$\begin{bmatrix} 1 & 2 & 3 & 8 \\ 2 & 4 & 5 & 2 \\ 3 & 5 & 8 & 3 \\ 8 & 2 & 3 & 4 \end{bmatrix}$$

is symmetric, while

$$\begin{bmatrix} 0 & 2 & -3 & -8 \\ -2 & 0 & 5 & 2 \\ 3 & -5 & 0 & 3 \\ 8 & -2 & -3 & 0 \end{bmatrix}$$

is skew-symmetric.

Still let $\boldsymbol{A}$ be a square matrix. We say that $\boldsymbol{A}$ is a *diagonal matrix* if it has non-zero values only at its main diagonal. If in addition all those non-zero values are 1, then we say that $\boldsymbol{A}$ is an *identity matrix*, e.g.:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Finally, if all the values in an $m \times n$ matrix $\boldsymbol{A}$ are 0, then we say that $\boldsymbol{A}$ as a *zero matrix*. We may denote the matrix as $\mathbf{0}$ if its dimensions are clear from the context.

2 Matrix Operations

Definition 1. *Let $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$ be $m \times n$ matrices. Then, we say that $\boldsymbol{A}$* **equals** *$\boldsymbol{B}$ if $a_{ij} = b_{ij}$ for all $i \in [1, n]$ and $j \in [1, m]$.*

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are equal, then we write $\boldsymbol{A} = \boldsymbol{B}$; otherwise, we write $\boldsymbol{A} \neq \boldsymbol{B}$.

Definition 2. *Let $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$ be $m \times n$ matrices. We define:*

- **(matrix addition)** *the result of $\boldsymbol{A}+\boldsymbol{B}$ to be the $m \times n$ matrix $\boldsymbol{C} = [c_{ij}]$ where $c_{ij} = a_{ij} + b_{ij}$ for all $i \in [1, n]$ and $j \in [1, m]$;*
- **(matrix subtraction)** *the result of $\boldsymbol{A}-\boldsymbol{B}$ to be the $m \times n$ matrix $\boldsymbol{C} = [c_{ij}]$ where $c_{ij} = a_{ij} - b_{ij}$ for all $i \in [1, n]$ and $j \in [1, m]$.*

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 8 & 4 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 2 & -3 & 1 \\ 0 & -7 & 0 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 3 \\ 5 & 1 & 6 \\ 6 & 0 & 8 \\ 8 & 2 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 8 & 4 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 2 & -3 & 1 \\ 0 & -7 & 0 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 7 & 4 \\ 6 & 14 & 8 \\ 8 & 6 & 0 \end{bmatrix}$$

Definition 3. **(Matrix Scalar Multiplication)** *Let* $\boldsymbol{A} = [a_{ij}]$ *be an* $m \times n$ *matrices, and* c *be a real value. Then, we define* $c\boldsymbol{A}$ *to be the* $m \times n$ *matrix* $\boldsymbol{B} = [b_{ij}]$ *where* $b_{ij} = c \cdot a_{ij}$ *for all* $i \in [1, n]$ *and* $j \in [1, m]$.

For example:

$$2 \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 8 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 6 & 8 & 10 \\ 12 & 14 & 16 \\ 16 & 8 & 4 \end{bmatrix}$$

Definition 4. **(Matrix Multiplication)** *Let* $\boldsymbol{A} = [a_{ij}]$ *be an* $m \times n$ *matrix, and* $\boldsymbol{B} = [b_{ij}]$ *be an* $n \times p$ *matrix. We define* $\boldsymbol{AB}$ *as the* $m \times p$ *matrix* $\boldsymbol{C} = [c_{ij}]$ *where*

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

for all $i \in [1, m]$ *and* $j \in [1, p]$.

Note that matrix multiplication requires that the number of *columns* of the first matrix must equal the number of *rows* of the second matrix. For example:

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 2 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & -2 & 0 \end{bmatrix} = \begin{bmatrix} -5 & 4 & 7 \\ -5 & 8 & 12 \\ -1 & -1 & 2 \end{bmatrix}$$

It is rudimentary to verify:

$$\begin{aligned} \boldsymbol{ABC} &= \boldsymbol{A}(\boldsymbol{BC}) \\ (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} &= \boldsymbol{AC} + \boldsymbol{BC} \\ \boldsymbol{C}(\boldsymbol{A} + \boldsymbol{B}) &= \boldsymbol{CA} + \boldsymbol{CB} \end{aligned}$$

Note that, in general, matrix multiplication does *not* necessarily obey commutativity. In fact, $\boldsymbol{AB}$ does not always guarantee that $\boldsymbol{BA}$ is well defined (recall the dimension requirement in Definition 4).

Definition 5. **(Matrix Transposition)** *Let* $\boldsymbol{A} = [a_{ij}]$ *be an* $m \times n$ *matrix. Then, the* **transpose** *of* $\boldsymbol{A}$*, denoted as* $\boldsymbol{A}^T$*, is the* $n \times m$ *matrix* $B = [b_{ij}]$ *where* $a_{ij} = b_{ji}$ *for all* $i \in [1, n]$ *and* $j \in [1, m]$.

For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \\ 8 & 4 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 6 & 8 \\ 2 & 4 & 7 & 4 \\ 3 & 5 & 8 & 2 \end{bmatrix}$$

It is rudimentary to verify:

$$\begin{aligned} (\boldsymbol{A}^T)^T &= \boldsymbol{A} \\ (\boldsymbol{A} + \boldsymbol{B})^T &= \boldsymbol{A}^T + \boldsymbol{B}^T \\ (c\boldsymbol{A})^T &= c\boldsymbol{A}^T \\ (\boldsymbol{AB})^T &= \boldsymbol{B}^T \boldsymbol{A}^T \end{aligned}$$