Linear Algebra

Hamid Dehghani
School of Computer Science
Birmingham
September 2020

UNIVERSITY OF BIRMINGHAM

A **matrix** is a rectangular array of numbers or functions which we will enclose in brackets. For example,

$$(1)\qquad \begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \quad \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}$$

The numbers (or functions) are called **entries** or, less commonly, *elements* of the matrix.

The first matrix in (1) has two **rows**, which are the horizontal lines of entries.

Furthermore, it has three **columns**, which are the vertical lines of entries.

$$
(1) \qquad \begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},
$$

$$
\begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \quad \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}
$$

The second and third matrices are **square matrices**, which means that each has as many rows as columns—3 and 2, respectively.

The entries of the second matrix have two indices, signifying their location within the matrix. The first index is the number of the row and the second is the number of the column, so that together the entry's position is uniquely identified. For example, a_{23} (read *a two three*) is in Row 2 and Column 3, etc. The notation is standard and applies to all matrices, including those that are not square.

$$
\begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},
$$

$$
(1) \qquad \begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \quad \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}
$$

Matrices having just a single row or column are called **vectors**.
The fourth matrix in (1) has just one row and is called a **row vector**.
The last matrix in (1) has just one column and is called a **column vector**.

General Concepts and Notations

We shall denote matrices by capital boldface letters **A**, **B**, **C**, … , or by writing the general entry in brackets; thus **A** = $[a_{jk}]$, and so on.

By an $\boldsymbol{m \times n}$ **matrix** (read *m by n matrix*) we mean a matrix with m rows and n columns—rows always come first!

$m \times n$ is called the **size** of the matrix. Thus an $m \times n$ matrix is of the form

(2)
$$\mathbf{A} = \left[a_{jk}\right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Vectors

A **vector** is a matrix with only one row or column. Its entries are called the **components** of the vector.

We shall denote vectors by ***lowercase*** boldface letters **a**, **b**, … or by its general component in brackets, $\mathbf{a} = [a_j]$, and so on.

A (general) **row vector** is of the form

$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$. For instance, $\mathbf{a} = \begin{bmatrix} -2 & 5 & 0.8 & 0 & 1 \end{bmatrix}$.

Vectors (continued)

A **column vector** is of the form

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \quad \text{For instance,} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ -7 \end{bmatrix}.$$

Definition

Equality of Matrices

Two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ are **equal**, written $\mathbf{A} = \mathbf{B}$, if and only if (1) they have the same size and (2) the corresponding entries are equal, that is, $a_{11} = b_{11}$, $a_{12} = b_{12}$, and so on.

Matrices that are not equal are called **different**. Thus, matrices of different sizes are always different.

Definition

Addition of Matrices

The **sum** of two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ ***of the same size*** is written $\mathbf{A} + \mathbf{B}$ and has the entries $a_{jk} + b_{jk}$ obtained by adding the corresponding entries of $\mathbf{A}$ and $\mathbf{B}$.
Matrices of different sizes cannot be added.

Definition

Scalar Multiplication (Multiplication by a Number)

The **product** of any $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ and any **scalar** c (number c) is written $c\mathbf{A}$ and is the $m \times n$ matrix $c\mathbf{A} = [ca_{jk}]$ obtained by multiplying each entry of $\mathbf{A}$ by c.

Rules for Matrix Addition and Scalar Multiplication.

From the familiar laws for the addition of numbers we obtain similar laws for the addition of matrices of the same size $m \times n$, namely,

$$
\text{(3)}\quad
\begin{array}{ll}
\text{(a)} & \mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A} \\
\text{(b)} & (\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}) \qquad (\text{written } \mathbf{A}+\mathbf{B}+\mathbf{C}) \\
\text{(c)} & \mathbf{A}+\mathbf{0}=\mathbf{A} \\
\text{(d)} & \mathbf{A}+(-\mathbf{A})=\mathbf{0}.
\end{array}
$$

Here $\mathbf{0}$ denotes the **zero matrix** (of size $m \times n$), that is, the $m \times n$ matrix with all entries zero.

Rules for Matrix Addition and Scalar Multiplication.
(continued)

Hence matrix addition is *commutative* and *associative* [by (3a) and (3b)]. Similarly, for scalar multiplication we obtain the rules

$$
\textbf{(4)}\qquad
\begin{aligned}
&\text{(a)}\ c(\mathbf{A}+\mathbf{B}) = c\mathbf{A}+c\mathbf{B}\\
&\text{(b)}\ (c+k)\mathbf{A} = c\mathbf{A}+k\mathbf{A}\\
&\text{(c)}\quad c(k\mathbf{A}) = (ck)\mathbf{A} \qquad\qquad \text{(written } ck\mathbf{A}\text{)}\\
&\text{(d)}\qquad 1\mathbf{A} = \mathbf{A}.
\end{aligned}
$$

Definition

Multiplication of a Matrix by a Matrix

The **product C** = **AB** (in this order) of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ times an $r \times p$ matrix $\mathbf{B} = [b_{jk}]$ is defined if and only if $r = n$ and is then the $m \times p$ matrix $\mathbf{C} = [c_{jk}]$ with entries

(1) $$c_{jk} = \sum_{l=1}^{n} a_{jl} b_{lk} = a_{j1} b_{1k} + a_{j2} b_{2k} + \cdots + a_{jn} b_{nk} \qquad \begin{aligned} j &= 1, \cdots, m \\ k &= 1, \cdots, p. \end{aligned}$$

The condition $r = n$ means that the second factor, **B**, must have as many rows as the first factor has columns, namely n. A diagram of sizes that shows when matrix multiplication is possible is as follows:

$$\begin{array}{cccc} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ [m \times n] & [n \times p] & = & [m \times p]. \end{array}$$

The entry c_{jk} in (1) is obtained by multiplying each entry in the jth row of **A** by the corresponding entry in the kth column of **B** and then adding these n products. For instance, $c_{21} = a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2n}b_{n1}$, and so on. One calls this briefly a ***multiplication of rows into columns***. For $n = 3$, this is illustrated by

$$m = 4\left\{\overbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}}^{n=3} \right. \overbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}}^{p=2} = \left. \overbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix}}^{p=2} \right\} m = 4$$

where we shaded the entries that contribute to the calculation of entry c_{21} just discussed.

EXAMPLE 1
Matrix Multiplication

$$\mathbf{AB} = \begin{bmatrix} 3 & 5 & -1 \\ 4 & 0 & 2 \\ -6 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 & 1 \\ 5 & 0 & 7 & 8 \\ 9 & -4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 22 & -2 & 43 & 42 \\ 26 & -16 & \boxed{14} & 6 \\ -9 & 4 & -37 & -28 \end{bmatrix}$$

Here $c_{11} = 3 \cdot 2 + 5 \cdot 5 + (-1) \cdot 9 = 22$, and so on.
The entry in the box is $c_{23} = 4 \cdot 3 + 0 \cdot 7 + 2 \cdot 1 = 14$.

EXAMPLE 2 CAUTION!

Matrix Multiplication Is *Not Commutative*, AB ≠ BA in General

This is illustrated by Example 1, where one of the two products is not even defined. But it also holds for square matrices. For instance,

$$\begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\text{but} \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix} = \begin{bmatrix} 99 & 99 \\ -99 & -99 \end{bmatrix}.$$

It is interesting that this also shows that $\mathbf{AB} = \mathbf{0}$ does ***not*** necessarily imply $\mathbf{BA} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$.

Our examples show that in matrix products ***the order of factors must always be observed very carefully***. Otherwise matrix multiplication satisfies rules similar to those for numbers, namely.

$$
(2)\quad
\begin{aligned}
&\text{(a)}\quad (k\mathbf{A})\mathbf{B} = k(\mathbf{AB}) = \mathbf{A}(k\mathbf{B}) \qquad \textit{written } k\mathbf{AB} \textit{ or } \mathbf{A}k\mathbf{B}\\
&\text{(b)}\quad \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \qquad \textit{written } \mathbf{ABC}\\
&\text{(c)}\quad (\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}\\
&\text{(d)}\quad \mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA}+\mathbf{CB}
\end{aligned}
$$

provided **A**, **B**, and **C** are such that the expressions on the left are defined; here, k is any scalar. (2b) is called the **associative law**. (2c) and (2d) are called the **distributive laws**.

Transposition

We obtain the **transpose of a matrix** by writing its rows as columns (or equivalently its columns as rows). This also applies to the transpose of vectors. Thus, a row vector becomes a column vector and vice versa.
In addition, for square matrices, we can also "reflect" the elements along the main diagonal, that is, interchange entries that are symmetrically positioned with respect to the main diagonal to obtain the transpose. Hence a_{12} becomes a_{21}, a_{31} becomes a_{13}, and so forth.
Also note that, if $\mathbf{A}$ is the given matrix, then we denote its transpose by $\mathbf{A}^\mathsf{T}$.

Definition

Transposition of Matrices and Vectors

The transpose of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^\mathsf{T}$ (read *A transpose*) that has the first *row* of $\mathbf{A}$ as its first *column*, the second *row* of $\mathbf{A}$ as its second *column*, and so on. Thus the transpose of $\mathbf{A}$ in (2) is $\mathbf{A}^\mathsf{T} = [a_{kj}]$, written out

$$(9) \qquad \mathbf{A}^\mathsf{T} = \left[a_{kj}\right] = \begin{bmatrix} a_{11} & \mathrm{a21} & \cdots & a_{m1} \\ \mathrm{a12} & a_{22} & \cdots & a_{m2} \\ \cdot & \cdot & \cdots & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

As a special case, transposition converts row vectors to column vectors and conversely.

Rules for transposition are

$$\text{(10)} \qquad \begin{aligned} &\text{(a)} \quad (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{A} \\ &\text{(b)} \quad (\mathbf{A}+\mathbf{B})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} + \mathbf{B}^{\mathsf{T}} \\ &\text{(c)} \quad (c\mathbf{A})^{\mathsf{T}} = c\mathbf{A}^{\mathsf{T}} \\ &\text{(d)} \quad (\mathbf{A}\mathbf{B})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}. \end{aligned}$$

CAUTION! Note that in (10d) the transposed matrices are ***in reversed order.***

Special Matrices

Symmetric and Skew-Symmetric Matrices.
Transposition gives rise to two useful classes of matrices. **Symmetric** matrices are square matrices whose transpose equals the matrix itself. **Skew-symmetric** matrices are square matrices whose transpose equals ***minus*** the matrix. Both cases are defined in (11) and illustrated by Example 8.

(11) $\mathbf{A}^\mathsf{T} = \mathbf{A}$ (thus $a_{kj} = a_{jk}$), Symmetric Matrix

$\mathbf{A}^\mathsf{T} = -\mathbf{A}$ (thus $a_{kj} = -a_{jk}$), hence $a_{jj} = 0$). Skew-Symmetric Matrix

EXAMPLE 8

Symmetric and Skew-Symmetric Matrices

$$\mathbf{A} = \begin{bmatrix} 20 & 120 & 200 \\ 120 & 10 & 150 \\ 200 & 150 & 30 \end{bmatrix} \quad \text{is symmetric, and}$$

$$\mathbf{B} = \begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix} \quad \text{is skew-symmetric.}$$

Diagonal Matrices.
These are square matrices that can have nonzero entries only on the main diagonal. Any entry above or below the main diagonal must be zero.

Identity Matrices.
These are square matrices that does not change any vector when we multiply that vector by that matrix. An example is a diagonal matrix, whose values on the main diagonal are 1s.

If all the diagonal entries of a diagonal matrix **S** are equal, say, c, we call **S** a **scalar matrix** because multiplication of any square matrix **A** of the same size by **S** has the same effect as the multiplication by a scalar, that is,

$$\mathbf{AS} = \mathbf{SA} = c\mathbf{A}. \tag{12}$$

In particular, a scalar matrix, whose entries on the main diagonal are all 1, is called a **unit matrix** (or **identity matrix**) and is denoted by $\mathbf{I}_n$ or simply by **I**. For **I**, formula (12) becomes

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}. \tag{13}$$

***We consider* square *matrices exclusively*.**

The **inverse** of an $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is denoted by $\mathbf{A}^{-1}$ and is an $n \times n$ matrix such that

(1) $$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

where $\mathbf{I}$ is the $n \times n$ unit matrix

If $\mathbf{A}$ has an inverse, then $\mathbf{A}$ is called a **nonsingular matrix**. If $\mathbf{A}$ has no inverse, then $\mathbf{A}$ is called a **singular matrix**.

If $\mathbf{A}$ *has an inverse, the inverse is unique.*

Indeed, if both $\mathbf{B}$ and $\mathbf{C}$ are inverses of $\mathbf{A}$, then $\mathbf{AB} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{I}$ so that we obtain the uniqueness from

$$\mathbf{B} = \mathbf{IB} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}.$$

Given a system of linear equations

$$Ax = b$$

For A^{-1} to exists, $\boldsymbol{x}$ must have only ONE solution for every value of **b**

Also possible to have no solutions, or infinite solutions!

- By definition, each A must have at least as many independent columns, as there are elements of $\boldsymbol{x}$

A set of vectors (matrix A) is linearly independent if no vectors in that set is a linear combination of other vectors in the set

5 Norms

The size of a vector is measured using a function called the ***norm, L^p***

$$||\boldsymbol{x}||_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$$

for p being real and >=1

Simply, it measures the distance from origin to point x

L^2 norm (p=2) is known as Euclidean norm, often denoted by $||\boldsymbol{x}||$

When differences between zero and nonzero elements are important, we us L^1 norm

$$||\boldsymbol{x}||_1 = \sum_i |x_i|.$$