Chapter 2: Matrices

Cheng Chang, Ph.D.
Department of Math & CS
Mercy College

Contents

1. Operations with Matrices

Matrices can be represented in one of the following three ways:

(1) An uppercase letter such as A, B, or C
(2) A representative element enclosed in brackets, such as $[a_{ij}], [b_{ij}]$, or $[c_{ij}]$
(3) A rectangular array of numbers

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Two matrices are *equal* when their corresponding entries are equal.

Definition 1.1. *Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are **equal** when they have the same size $m \times n$ and $a_{ij} = b_{ij}$ for $1 \le i \le m$ and $1 \le j \le n$.*

A matrix that has only one column is a **column matrix** or **column vector**. Similarly, a matrix that has only one row is a **row matrix** or **row vector**. We usually use boldface lowercase letters (e.g. $\mathbf{a}, \mathbf{b}, \mathbf{c}$) or lowercase letters accented with right arrows (e.g. $\vec{a}, \vec{b}, \vec{c}$) to represent column/row vectors. Sometimes we would think of a matrix as a sequence of column vectors, for instance, $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ can be partitioned into the two column vectors $\vec{a}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ as shown below

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = [\vec{a}_1 \quad \vec{a}_2]$$

To **add** two matrices (of the same size), add their corresponding entries.

Definition 1.2. *If $A = [a_{ij}]$ and $B = [b_{ij}]$ are matrices of size $m \times n$, then their **sum** is the $m \times n$ matrix $A + B = [a_{ij} + b_{ij}]$.*

The sum of two matrices of different sizes is undefined.

Example 1.1. *Add these matrices:*

$$\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & -2 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

When working with matrices, real numbers are referred to as **scalars**. To multiply a matrix A by a scalar c, multiply each entry in A by c.

Definition 1.3. *If $A = [a_{ij}]$ is a $m \times n$ matrix and c is a scalar, then the **scalar multiple** of A by c is the $m \times n$ matrix $cA = [ca_{ij}]$.*

Example 1.2. *For the matrices A and B, find (a) $3A$ (b) $-B$, and (c) $3A - B$.*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix}$$

Another basic matrix operation is **matrix multiplication**.

Definition 1.4. *If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the* ***product*** *AB is an $m \times p$ matrix $AB = [c_{ij}]$ where*

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

This definition means that to find the entry in the ith row and the jth column of the product AB, multiply the entries in the ith row of A by the corresponding entries in the jth column of B and then add the results.

Example 1.3. *Find the product AB, where*

$$A = \begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix}.$$

Remark 1.1. *Keep in mind that for the product of two matrices to be defined, the number of columns of the first matrix must equal the number of rows of the second matrix. In general, the multiplication of two matrices is not commutative. For instance, in the previous example, if we switch the order of the product, then BA is not even defined.*

Example 1.4. *Multiply:*

a. $\begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & -2 \end{bmatrix} \begin{bmatrix} -2 & 4 & 2 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{bmatrix}$

b. $\begin{bmatrix} 3 & 4 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

d. $\begin{bmatrix} 1 & -2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

e. $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & -3 \end{bmatrix}$

One application of matrix multiplication is representing a system of linear equations. For example, a system

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

can be written as the matrix equation $A\vec{x} = \vec{b}$, where A is the coefficient matrix of the system, and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ are column vectors.

There is yet another way to represent a linear system. Given a linear system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ and } \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

We can multiply the left-hand-side of $A\vec{x} = \vec{b}$ together and have the following equation

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \vec{b}$$

Next we separate the left-hand-side as a sum of n column vectors:

$$\begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = \vec{b}$$

Finally, pull out the common factor in each column, we obtain

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \vec{b}$$

If we use $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ to represent the column vectors of the matrix A, then the above equation can be written as

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b} \tag{1}$$

The left-hand-side of the equation (1) is called a **linear combination** of the column vectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ with **coefficients** $x_1, x_2, \ldots, x_n$. And the equation (1) is called the vector equation representation of the original linear system $A\vec{x} = \vec{b}$.

Definition 1.5 (**Linear Combinations of Column Vectors**). *The matrix product $A\vec{x}$ is a linear combination of the column vectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ that form the coefficient matrix A.*

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Furthermore, the system

$$A\vec{x} = \vec{b}$$

is consistent if and only if $\vec{b}$ can be expressed as such a linear combination, where the coefficients of the linear combination are a solution of the system.

Example 1.5. *Solve the linear system*

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 0 \\ 4x_1 + 5x_2 + 6x_3 = 3 \\ 7x_1 + 8x_2 + 9x_3 = 6 \end{cases}$$

Express the solution set as a linear combination of column vectors.

Example 1.6. *Can we express the column vector* $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ *as a linear combination of the column vectors* $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, *and* $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$*?*

2. Properties of Matrix Operations

In this section, we will look at more operations of matrices. The next theorem lists several properties of matrix addition and scalar multiplication.

Theorem 2.1. *If A, B, and C are $m \times n$ matrices, and c and d are scalars, then the properties below are true.*

(1) $A + B = B + A$
(2) $A + (B + C) = (A + B) + C$
(3) $(cd)A = c(dA)$
(4) $1A = A$
(5) $c(A + B) = cA + cB$
(6) $(c + d)A = cA + dA$

A $m \times n$ **zero matrix** is the matrix with all 0 entries, denoted by O_{mn}. It is the **additive identity** for the set of all $m \times n$ matrices, i.e. $A + O_{mn} = A$ for all $m \times n$ matrix A. When the size of the matrix is understood, we may denote a zero matrix simply by $\vec{0}$ or $\mathbf{0}$.

Theorem 2.2. *If A is an $m \times n$ matrix and c is a scalar, then the properties below are true.*

(1) $A + O_{mn} = A$
(2) $A + (-A) = O_{mn}$
(3) *If* $cA = O_{mn}$, *then* $c = 0$ *or* $A = O_{mn}$.

Example 2.1. *Solve for X in the equation $3X + A = B$, where*

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix}$$

The next theorem states some useful properties of matrix multiplication.

Theorem 2.3. *If A, B, and C are matrices (with sizes such that the matrix products are defined), and c is a scalar, then the properties below are true.*

(1) $A(BC) = (AB)C$
(2) $A(B + C) = AB + AC$
(3) $(A + B)C = AC + BC$
(4) $c(AB) = (cA)B = A(cB)$

Example 2.2. *Find the matrix product ABC by grouping the factors first as $(AB)C$ and then as $A(BC)$. Show that you obtain the same result from both processes.*

$$A = \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 2 \\ 3 & -2 & 1 \end{bmatrix} \quad C = \begin{bmatrix} -1 & 0 \\ 3 & 1 \\ 2 & 4 \end{bmatrix}$$

The next example shows that even when both products AB and BA are defined, they may not be equal.

Example 2.3. *Show that AB and BA are not equal for the matrices*

$$A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$$

Another important quality of matrix multiplication is that it does not have a general cancellation property. That is, when $AC = BC$, it is not necessarily true that $A = B$. The next example demonstrates this.

Example 2.4. *Show that $AC = BC$.*

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$$

Now we define a special type of square matrix that has 1's on the main diagonal and 0's elsewhere.

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

When the size of the matrix is understood to be $n \times n$, we may denote I_n simply as I. I_n is called the **identity matrix of order** n, the name is an obviously choice as is shown in the next theorem.

Theorem 2.4. *If A is a matrix of size $m \times n$, then the properties below are true.*

(1) $AI_n = A$

(2) $I_m A = A$

For repeated multiplication of square matrices, use the same exponential notation used with real numbers, i.e. $A^k = AA \cdots A$ where A occurs k times on the right-hand side. It is convenient also to define $A^0 = I_n$.

Example 2.5. *For the matrix* $A = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix}$, *compute* A^3.

Example 2.6 (Optional). *Use the properties of the matrix operations to show that if a linear system $A\vec{x} = \vec{b}$ has two distinct solutions $\vec{x}_1$ and $\vec{x}_2$, i.e. $\vec{x}_1 \neq \vec{x}_2$, then there are infinitely many solutions to the system.*

The **transpose** of a matrix is formed by writing its rows as columns. For example, if A is the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

then the transpose, denoted by A^T, is the $n \times m$ matrix

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Example 2.7. *Find the transpose of each matrix.*

(1) $A = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$ *(2)* $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ *(3)* $C = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *(4)* $D = \begin{bmatrix} 0 & 1 \\ 2 & 4 \\ 1 & -1 \end{bmatrix}$

Theorem 2.5. *If A and B are matrices (with sizes such that the matrix operations are defined) and c is a scalar, then the properties below are true.*

(1) $(A^T)^T = A$
(2) $(A + B)^T = A^T + B^T$
(3) $(cA)^T = c(A^T)$
(4) $(AB)^T = B^T A^T$

Example 2.8. *Show that $(AB)^T$ and $B^T A^T$ are equal.*

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -1 & 0 & 3 \\ 0 & -2 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 1 \\ 2 & -1 \\ 3 & 0 \end{bmatrix}$$

In general, $A \neq A^T$. If a matrix A has the property that $A^T = A$, then we say A is symmetric.

Example 2.9. *For the matrix $A = \begin{bmatrix} 1 & 3 \\ 0 & -2 \\ -2 & -1 \end{bmatrix}$, find the product AA^T and show that is is symmetric.*

3. The Inverse of a Matrix

When try to solve a simple linear equation, for example

$$2x = 5$$

we can isolate the variable x by multiplying the multiplicative inverse of 2, i.e. $\frac{1}{2}$ or 2^{-1}, on both sides of the equation:

$$2^{-1} \cdot 2x = 2^{-1} \cdot 5$$

so $x = 5/2$. Likewise, given a linear system $A\vec{x} = \vec{b}$, we want to explore the possibility of solving this system by finding the multiplicative inverse of the matrix A.

Definition 3.1. *An $n \times n$ matrix A is **invertible** (or **nonsingular**) when there exists an $n \times n$ matrix B such that*

$$AB = BA = I_n$$

*where I_n is the identity matrix of order n. The matrix B is the (multiplicative) **inverse** of A. A matrix that does not have an inverse is **non-invertible** (or **singular**).*

Nonsquare matrices do not have inverses. Not all square matrices have inverses. The next theorem, however, states that if a matrix does have an inverse, then that inverse is unique.

Theorem 3.1. *If A is an invertible matrix, then its inverse is unique. The inverse of A is denoted by A^{-1}.*

Example 3.1. *Show that B is the inverse of A, where*

$$A = \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}$$

Example 3.2. *Find the inverse of the matrix*

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}.$$

In general, to find the inverse matrix, we follow the procedure below.

Procedure 3.1 (Finding the Inverse of a Matrix by Gauss-Jordan Elimination). *Let A be a square matrix of order n.*

*(1) Write the $n \times 2n$ matrix that consists of A on the left and the $n \times n$ identity matrix I on the right to obtain $[A\ I]$. This process is called **adjoining** matrix I to matrix A.*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

(2) If possible, row reduce A to I using elementary row operations on the entire matrix $[A\ I]$. The result will be the matrix $[I\ A^{-1}]$. It this is not possible, then A is non-invertible (or singular).

(3) Check your work by multiplying to see that $AA^{-1} = I = A^{-1}A$.

Example 3.3. *Find the inverse of the matrix.*

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -6 & 2 & 3 \end{bmatrix}$$

Example 3.4. *Show that the matrix has no inverse.*

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ -2 & 3 & -2 \end{bmatrix}$$

For 2×2 matrices, we can use a formula for the inverse rather than Gauss-Jordan elimination. For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Example 3.5. *If possible, find the inverse of each matrix.*

$$A = \begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -1 \\ -6 & 2 \end{bmatrix}$$

3.1. **Properties of Inverses.**

Theorem 3.2. *If A is an invertible matrix, k is positive integer, and c is a nonzero scalar, then A^{-1}, A^k, cA, and A^T are invertible and the statements below are true.*

(1) $(A^{-1})^{-1} = A$
(2) $(A^k)^{-1} = (A^{-1})^k$
(3) $(cA)^{-1} = \frac{1}{c}A^{-1}$
(4) $(A^T)^{-1} = (A^{-1})^T$

The next theorem gives a formula for computing the inverse of a product of two matrices.

Theorem 3.3. *If A and B are invertible matrices of order n, then AB is invertible and*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

With the previous theorem, for nonsingular matrices, we may define an exponential term A^{-k} as $A^{-1}A^{-1} \cdots A^{-1}$ where A^{-1} occurs k times.

Example 3.6. *Compute A^{-2} two different ways (one way as $(A^{-1})^2$, the other as $A^{-1}A^{-1}$) and show that the results are equal.*

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$$

Theorem 3.4. *If C is an invertible matrix, then the properties below are true.*

(1) If $AC = BC$, then $A = B$.
(2) If $CA = CB$, then $A = B$.

Remark 3.1. *In order for the cancellation properties to be true, it's very important that C is invertible.*

Now we're able to solve a linear system using the multiplicative inverse of the coefficient matrix, assuming the coefficient matrix is a square matrix and invertible.

Theorem 3.5. *If A is an invertible matrix, then the system of linear equations $A\vec{x} = \vec{b}$ has a unique solution $\vec{x} = A^{-1}\vec{b}$.*

Example 3.7. *Use an inverse matrix to solve the system.*

$$2x + 3y + z = 4$$
$$3x + 3y + z = 8$$
$$2x + 4y + z = 5$$