

Calculus II - Basic Matrix Operations

Ryan C. Daileda

1 Terminology

A *matrix* is a rectangular array of numbers, for example

$$\begin{pmatrix} 1 & 5 & 7 \\ -2 & -7 & -5 \end{pmatrix}, \begin{pmatrix} \sqrt{3} & -1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -3 \\ -7 & 9 \end{pmatrix}, \text{ or } \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \end{pmatrix}.$$

The numbers in any matrix are called its *entries*. The entries of a matrix are organized into *rows* and *columns*, which are simply the horizontal and vertical (resp.) lists of entries appearing in the matrix. For example, if

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

then the rows of M are $(1 \ 0 \ 2 \ 0 \ -3)$, $(0 \ 4 \ 0 \ 5 \ 0)$ and $(-6 \ 0 \ 7 \ 0 \ 8)$ whereas the columns of M are

$$\begin{pmatrix} 1 \\ 0 \\ -6 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}, \text{ and } \begin{pmatrix} -3 \\ 0 \\ 8 \end{pmatrix}.$$

The *dimensions* of a matrix are the numbers of rows and columns it has. If a matrix has m rows and n columns we say that it is an $m \times n$ matrix (note that we always list the number of rows first). So, the first four matrices above have dimensions 2×3 , 2×2 , 4×2 and 3×4 , respectively. The dimensions of the matrix M are 3×5 . An $m \times n$ matrix is called *square* if $m = n$. Thus, the only example of a square matrix above is the second.

It is worth noting that an $m \times n$ matrix will have m rows with n entries each, and n columns with m entries each. That is, the number of entries in any row of a matrix is the number of columns of that matrix, and vice versa. This is readily apparent in each of the examples above.

So that we can more easily refer to various entries in matrices, we index the columns of a matrix from left to right and the rows from top to bottom. For example, the *first column* of M (above) is

$$\begin{pmatrix} 1 \\ 0 \\ -6 \end{pmatrix},$$

the *second column* is

$$\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix},$$

the *third column* is

$$\begin{pmatrix} 2 \\ 0 \\ 7 \end{pmatrix},$$

etc. The *first row* of M is $(1 \ 0 \ 2 \ 0 \ -3)$, the *second row* is $(0 \ 4 \ 0 \ 5 \ 0)$ and the *third row* is $(-6 \ 0 \ 7 \ 0 \ 8)$. We can use this numbering scheme to easily refer to entries in a matrix: we call the entry located in row i and column j the (i, j) -*entry*. For the matrix

$$B = \begin{pmatrix} 1 & 3 & 4 \\ -1/6 & 0 & -5 \\ 2 & -1 & 7 \\ 1/4 & 2/3 & 9 \end{pmatrix}$$

the $(1, 1)$ -entry is 1, the $(3, 2)$ -entry is -1 , the $(4, 3)$ -entry is 9 and the $(2, 1)$ -entry is $-1/6$.

To write down a matrix with unspecified entries we use variables with subscripts that indicate their position in the matrix, using the convention described above. A generic $m \times n$ matrix can therefore be denoted

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \quad (1)$$

or just $A = (a_{ij})$ for short. Using this notation, if $A = (a_{ij})$ and $B = (b_{ij})$ are two $m \times n$ matrices then $A = B$ (i.e. A and B are the same matrix) if and only if $a_{ij} = b_{ij}$ for every pair (i, j) .

An $m \times 1$ matrix has the form

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}$$

and is called, appropriately, a *column vector*. Notice that since a column vector has only a single column we have used only single subscripts to index its entries. Likewise, a $1 \times n$ matrix looks like

$$(r_1 \ r_2 \ r_3 \ \cdots \ r_n)$$

and is called a *row vector*. When we use the word *vector* with no qualification we will usually mean a column vector. Column vectors give us another shorthand for writing down generic matrices. Notice that if we use the matrix A in (1) and set

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{pmatrix}, \cdots, \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

(i.e. we use the entries in the j -th column of A as the entries in \mathbf{a}_j) then we can write

$$A = (\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \cdots \ \mathbf{a}_n).$$

In a similar way one can also use the rows of A to express A in terms of row vectors, but since we won't be using this idea later there's no need to write it out explicitly.

2 Scalar multiplication and addition of matrices

Having dispensed with the basic terminology and notation of matrices, we now turn to how they are manipulated algebraically. We will see that it is possible to add, subtract and multiply matrices together, but only if certain restrictions on their dimensions are met. We begin with the notion of scalar multiplication. Given an $m \times n$ matrix $A = (a_{ij})$ and a number (scalar) c we define

$$cA = (ca_{ij}).$$

That is, cA is the matrix obtained by multiplying *every* entry of A by c . As examples, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 0 & 3 & -6 \\ 2 & 5 & -1 \end{pmatrix}$$

then

$$2A = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}, 0A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \frac{1}{2}B = \begin{pmatrix} 0 & 3/2 & -3 \\ 1 & 5/2 & -1/2 \end{pmatrix}, -B = (-1)B = \begin{pmatrix} 0 & -3 & 6 \\ -2 & -5 & 1 \end{pmatrix}.$$

Adding two matrices is also done entry-by-entry. If $A = (a_{ij})$ and $B = (b_{ij})$ are two $m \times n$ matrices, then their *sum* is $A + B = (a_{ij} + b_{ij})$. That is, the (i, j) -entry of $A + B$ is the sum of the (i, j) -entries of A and B . It is important to note that it is only possible to add two matrices if they have *exactly the same dimensions*. Here's an example: if

$$A = \begin{pmatrix} 6 & 5 \\ 3 & 4 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & -1 \\ -2 & 3 \\ 4 & 1 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} 6 & 4 \\ 1 & 7 \\ 6 & 2 \end{pmatrix}$$

and

$$2A - 5B = \begin{pmatrix} 12 & 15 \\ 16 & -7 \\ -16 & -3 \end{pmatrix}.$$

The following theorem summarizes the main properties of matrix addition. The proofs of these properties follow directly from the definitions made so far and are left to the reader. We will find it useful to be able to refer to the $m \times n$ *zero matrix*, which is the matrix all of whose entries are zero.

Theorem 1. *Let A , B and C be $m \times n$ matrices, let c be a real number and let $\mathbf{0}$ denote the $m \times n$ zero matrix. Then*

1. $A + B = B + A$;
2. $A + (B + C) = (A + B) + C$;
3. $\mathbf{0} + A = A + \mathbf{0} = A$;
4. $c(A + B) = cA + cB$;
5. $0A = \mathbf{0}$;

3 Matrix multiplication

Defining the matrix product is a two step process. First we will define what it means to multiply a matrix by a column vector and then we'll use that to tell us how to multiply two matrices in general.

3.1 Matrix-vector multiplication

Let A be an $m \times n$ matrix and let \mathbf{v} be an $n \times 1$ column vector (notice that the vector \mathbf{v} has as many entries as A has columns). Write A in terms of its columns as above,

$$A = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \cdots \quad \mathbf{a}_n)$$

and write out the entries of \mathbf{v} as

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix}.$$

The *product* of A with \mathbf{v} is defined to be

$$A\mathbf{v} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3 + \cdots + v_n\mathbf{a}_n.$$

In words, we multiply the columns of A by the respective entries of \mathbf{v} and then add the results together. According to this definition, the product of an $m \times n$ matrix and an $n \times 1$ column vector is an $m \times 1$ column vector, i.e. the product is a column with as many entries as A has rows.

The process of multiplying a matrix by a vector is straightforward enough once one is used to the definition. Let's look at some examples. Suppose that we take

$$A = \begin{pmatrix} 6 & 5 \\ 3 & 4 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 3 & -6 & 5 \\ 2 & 5 & -1 & 0 \end{pmatrix}.$$

The matrix A can only be multiplied by column vectors with 2 entries while B can only be multiplied by column vectors with 4 entries. So, if we take

$$\mathbf{v} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} 2 \\ 0 \\ -5 \\ 1 \end{pmatrix}$$

then

$$A\mathbf{v} = \begin{pmatrix} 6 & 5 \\ 3 & 4 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = 3 \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 13 \\ 5 \\ 5 \end{pmatrix}$$

and

$$B\mathbf{w} = \begin{pmatrix} 0 & 3 & -6 & 5 \\ 2 & 5 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ -5 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 0 \begin{pmatrix} 3 \\ 5 \end{pmatrix} - 5 \begin{pmatrix} -6 \\ -1 \end{pmatrix} + \begin{pmatrix} 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 35 \\ 9 \end{pmatrix}$$

There is a close connection between matrix vector multiplication and systems of linear equations. A *linear equation* in the variables x_1, x_2, \dots, x_n has the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, where the a_i and b are constants. So, for example, $3x - 2y + 7z = 8$ is a linear equation in the variables x, y and z . Suppose we are faced with the following system of m linear equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

This system of inequalities is encompassed by the single vector equality

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \quad (2)$$

According to the definitions of matrix addition and scalar multiplication, on the left hand side we have

$$\begin{aligned} \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} &= \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}. \end{aligned}$$

If we let $A = (a_{ij})$ (the so-called *coefficient matrix* of the system) this is precisely $A\mathbf{x}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

If we now set

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

then it follows that (2) is the same as the matrix equation $A\mathbf{x} = \mathbf{b}$. That is, *a system of linear equations is the same thing as a matrix equation of the form $A\mathbf{x} = \mathbf{b}$* , where A , \mathbf{x} and \mathbf{b} are related to the system as described above.

As an example, consider the linear system

$$\begin{aligned} x - 2y + 3z + w &= 1, \\ y + 2z - 7w &= 5, \\ 6x + y - 5z &= 0. \end{aligned}$$

This collection of three equations is equivalent to the single matrix equation

$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 1 & 2 & -7 \\ 6 & 1 & -5 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 0 \end{pmatrix}.$$

Here we have included zeros in the second and third row of the coefficient matrix to account for the fact that the variables x and w do not appear in the second and third equations, respectively. Likewise, the simpler system

$$\begin{aligned} 2x - y &= 5, \\ 3x + 4y &= -7 \end{aligned}$$

is the same as the matrix equation

$$\begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ -7 \end{pmatrix}.$$

3.2 Matrix-matrix multiplication

Since we can now multiply matrices by (suitably sized) column vectors, we can develop a way to multiply matrices by other (suitably sized) matrices. Let A be an $m \times n$ matrix and let B be a $n \times p$ matrix. Notice that B has as many rows as A has columns. In particular, the columns of B are $n \times 1$ column vectors and can therefore individually be multiplied by A . To be more specific, write B in terms of its columns:

$$B = (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \cdots \quad \mathbf{b}_p)$$

where each \mathbf{b}_j is an $n \times 1$ column vector. We define the *product* of A and B to be

$$AB = A (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \cdots \quad \mathbf{b}_p) = (A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad A\mathbf{b}_3 \quad \cdots \quad A\mathbf{b}_p).$$

That is, to multiply two matrices simply multiply the first matrix by the columns of the second and use the results as the columns in a new matrix. Since each $A\mathbf{j}$ is an $m \times 1$ column vector, and there are exactly p of them, we find that AB is an $m \times p$ matrix.

Let's look at a quick example. Take

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -3 & 2 \\ 1 & 2 & 4 \end{pmatrix}, B = \begin{pmatrix} 3 & 5 \\ 2 & -1 \\ 1 & 3 \end{pmatrix}.$$

The product AB makes sense since A has as many columns as B has rows. The definition of matrix multiplication says that

$$AB = \left(A \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \quad A \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix} \right).$$

We find that

$$A \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 11 \end{pmatrix}$$

and

$$A \begin{pmatrix} 5 \\ -1 \\ 3 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 9 \\ 10 \end{pmatrix}$$

so that

$$AB = \begin{pmatrix} 2 & 2 \\ -4 & 9 \\ 11 & 10 \end{pmatrix}.$$

The $n \times n$ *identity matrix* I is the (square) matrix all of whose entries are zero except for those along the “main diagonal” which are all equal to 1. Symbolically

$$I = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The 2×2 , 3×3 and 4×4 identity matrices are then

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

respectively.

The following theorem gives the main properties of matrix multiplication. These all follow directly from the definitions, but some are harder to prove than others, most notably that matrix multiplication is associative.

Theorem 2. *Let A be $m \times n$, B and C be $n \times p$, D be $p \times q$, and let c be a real number. Then*

1. $A(B + C) = AB + AC$;
2. $(B + C)D = BD + CD$;
3. $(AB)D = A(BD)$;
4. if I is the $m \times m$ identity matrix then $IA = A$;
5. if I is the $n \times n$ identity matrix then $AI = A$;
6. $c(AB) = (cA)B = A(cB)$.

4 Exercises

In exercises 1 and 2, let

$$A = \begin{pmatrix} -6 & 3 & -4 \\ 0 & 0 & -6 \end{pmatrix}, B = \begin{pmatrix} 5 & 0 \\ -7 & -3 \\ -9 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & -5 \\ -3 & 9 \end{pmatrix},$$

$$D = \begin{pmatrix} 5 & 0 & -2 \\ 3 & -8 & 7 \\ 6 & -1 & -6 \end{pmatrix}, E = \begin{pmatrix} -8 & 1 & 6 \\ 4 & -3 & 0 \end{pmatrix}$$

and compute each matrix sum or product if it is defined. If it is not defined, explain why.

Exercise 1.

- a. $A - B$
- b. $A - 3E$
- c. $2A + DB$
- d. AC

Exercise 2.

- a. $A + CB$
- b. $3BC - A$
- c. CAD
- d. $CA - E$

Exercise 3. If $A = \begin{pmatrix} 2 & 1 \\ 3 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 3 \\ -4 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 9 & 1 \\ 3 & 5 \end{pmatrix}$ show that $AB \neq BA$ but that $AC = CA$.

Exercise 4. If $A = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}$, construct a nonzero 2×2 matrix B (with two *distinct* columns) so that AB is the zero matrix.

If A is an $n \times n$ matrix, we say the $n \times n$ matrix B is the *inverse* of A if $AB = BA = I$, where I is the $n \times n$ identity matrix. Use this definition in exercises 5 - 7.

Exercise 5. Show that if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc \neq 0$ then the inverse of A is $B = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Exercise 6. If $A = \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 8 \\ 16 \end{pmatrix}$, use the inverse of A (see the previous exercise) to solve the matrix equation $A\mathbf{x} = \mathbf{b}$ for \mathbf{x} .

Exercise 7. Consider the system

$$\begin{aligned} 2x - y - z &= 3, \\ x - y - 5z &= 0, \\ -x - 3z &= 2. \end{aligned} \tag{3}$$

a. If A is the coefficient matrix for this system show that

$$B = \begin{pmatrix} -3 & 3 & -4 \\ -8 & 7 & -9 \\ 1 & -1 & 1 \end{pmatrix}$$

is the inverse of A .

b. Use part (a) to solve the system (3).

c. Suppose the constants 3, 0 and 2 on the right hand sides of the system (3) are replaced with -1 , 2 and 1, respectively. Use part (a) to help you solve this new system.

Exercise 8. If $A = \begin{pmatrix} -7 & 6 \\ -12 & 10 \end{pmatrix}$ find a nonzero vector \mathbf{v} so that $A\mathbf{v} = \mathbf{v}$.

Exercise 9. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Rewrite the matrix equation $AB = BA$ as a system of four linear equations, and write down the coefficient matrix for this system.