

# Matrix Algebra

A **matrix** is a rectangular array of numbers (real or complex), usually enclosed by large square brackets. The **dimension** of a matrix is  $m \times n$ , where  $m$  is the number of rows (horizontal lines in the array) and  $n$  is the number of columns (vertical lines in the array).

**EXAMPLE 1:** The matrix

$$A = \begin{bmatrix} 2 & 3 & -9 \\ -1 & 0 & \frac{1}{5} \end{bmatrix}$$

has two rows and three columns. Therefore the dimension of  $A$  is  $2 \times 3$ .  $\square$

If a matrix has only one row or only one column, we call it a **vector**. If we wish to specify what the situation is, we can call it a **row vector** (so that it is  $1 \times n$ ) or a **column vector** (so that it is  $n \times 1$ ).

The values in each row and column of a matrix are called the **entries** of the matrix. We can specify an entry by stating the row and column in which it appears. This is usually done using two subscripts attached to the symbol representing the matrix; the first subscript indicates the row, and the second subscript indicates the column. For example, the entry in the  $i^{th}$  row and  $j^{th}$  column of a matrix  $A$  is denoted by  $A_{ij}$ . Two matrices of the same dimension are equal if and only if all their entries are equal.

**EXAMPLE 2:** Let

$$B = \begin{bmatrix} 2 & 4 \\ -1 & 0 \end{bmatrix}.$$

The entry  $B_{11}$  is the entry in the first row and first column, so  $B_{11} = 2$ . Also,  $B_{12}$  is the entry in the first row and second column, so  $B_{12} = 4$ . Similarly,  $B_{21} = -1$  and  $B_{22} = 0$ .  $\square$

If two matrices are of the same dimension, we can add and subtract one from another by adding or subtracting each corresponding entry. Formally, we state this by saying that if  $C = A + B$ , then  $C_{ij} = A_{ij} + B_{ij}$ . Also, if  $D = A - B$ , then  $D_{ij} = A_{ij} - B_{ij}$ . We can also multiply a matrix by a scalar by distributing the multiplication to each entry in the matrix: if  $E = cA$ , then  $E_{ij} = cA_{ij}$ . Note that for every scalar  $c$ ,  $cA = Ac$ .

**EXAMPLE 3:** Let

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

Then

$$A + B = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, \quad A - B = \begin{bmatrix} 0 & 7 \\ -1 & -5 \end{bmatrix} \quad \text{and} \quad 3A = \begin{bmatrix} 3 & 9 \\ 0 & -6 \end{bmatrix}$$

□

We would also like to multiply matrices together, and it turns out that the useful way to define this is not the obvious way. We will not simply multiply corresponding entries. For application purposes, it is far more useful to define matrix multiplication as follows: if  $A$  is an  $m \times k$  matrix and  $B$  is an  $k \times n$  matrix, then the entries of the product matrix  $AB$  are

$$(AB)_{ij} = \sum_{l=1}^k A_{il}B_{lm}.$$

This operation is only defined when the number of columns of the first matrix matches the number of rows of the second matrix, and the result is a matrix that has the same number of rows as  $A$  and the same number of columns as  $B$ . That is to say, if  $A$  is  $m \times k$  and  $B$  is  $k \times n$ , then  $AB$  is  $m \times n$ .

**EXAMPLE 4:** Let

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

Then

$$AB = \begin{bmatrix} (1 \cdot 1 + 3 \cdot 1) & (1 \cdot 4 + 3 \cdot 3) & (1 \cdot 0 + 3 \cdot 1) \\ (0 \cdot 1 + 2 \cdot 1) & (0 \cdot 4 + 2 \cdot 3) & (0 \cdot 0 + 2 \cdot 1) \end{bmatrix} = \begin{bmatrix} 4 & 13 & 3 \\ 2 & 6 & 2 \end{bmatrix},$$

but  $BA$  is not defined. □

**EXERCISE 1:** Let  $A = \begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix}$ . Calculate  $AB$  and  $BA$ .

The last exercise shows that matrix multiplication is not commutative: even if both multiplications  $AB$  and  $BA$  are defined,

$$AB \neq BA \text{ in general.}$$

(Both products could be equal, but usually that is not the case.)  
On the other hand, matrix multiplication is distributive:

$$A(B + C) = AB + AC,$$

provided the matrices have dimensions for which the additions and multiplications are all defined.

The notation  $I_n$  is used to denote an  $n \times n$  **identity matrix**. It has entries of 1 along the main diagonal (from top left to bottom right) and the entries are zero everywhere else. For example,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If the appropriate dimension is clear from context, we often omit the subscript and just write  $I$ . For example, if  $I$  is used to multiply a matrix  $A$  on the left, as in the product  $IA$ , then the multiplication is only defined if the number of columns of  $I$  matches the number of rows of  $A$ . When this is defined, it turns out that  $IA = A$  for every matrix  $A$ .

**EXERCISE 2:** Let

$$A = \begin{bmatrix} 2 & 4 \\ -1 & 0 \end{bmatrix}.$$

Verify that  $IA = A$  and  $AI = A$  by actually carrying out the matrix multiplications.

One of the main reasons this definition of multiplication is useful is that it allows us to write a system of linear equations as a matrix equation.

**EXAMPLE 5:** Consider the following system of two linear equations in two unknowns:

$$\begin{aligned} 3x - 2y &= 4 \\ 2x + y &= 0 \end{aligned}$$

These two scalar equalities can be written as a matrix equality:

$$\begin{bmatrix} 3x - 2y \\ 2x + y \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

But notice that the left side can be written as a product of matrices:

$$\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

Thus if  $A = \begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}$ ,  $X = \begin{bmatrix} x \\ y \end{bmatrix}$  and  $B = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ , then the system of equations above can be represented by

$$AX = B.$$

□

When writing a system of linear equations as a matrix equation as in the last example, the matrix  $A$  is called the **coefficient matrix**. In a course on linear algebra, students learn how to solve this equation for  $X$  in a variety of ways. If the coefficient matrix  $A$  is a **square matrix**, meaning that it has the same number of rows as columns, then it is sometimes possible to find an inverse matrix  $A^{-1}$  such that  $A^{-1}A = I$ . In such an instance, we would multiply both sides of the matrix equation to obtain

$$A^{-1}AX = A^{-1}B,$$

or

$$IX = A^{-1}B,$$

which is the same as

$$X = A^{-1}B.$$

If it is possible to solve the equation this way, then there must be exactly one solution for each matrix  $B$ . (We will pursue this approach to solving equations a little in the problem set.) This motivates the following question which we will pursue immediately: under what conditions on  $A$  does the matrix equation  $AX = B$  have exactly one solution for the unknown column vector  $X$ ?

Let us write the general situation for a  $2 \times 2$  matrix  $A$ :

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} k \\ l \end{bmatrix}.$$

The matrix equation  $AX = B$  is equivalent to the system of scalar equations:

$$\begin{aligned} ax + by &= k \\ cx + dy &= l \end{aligned}$$

We can think of this system as a collection of lines in the  $xy$ -plane and the solution of this system as the point of intersection of those lines. There will be infinitely many solutions if the lines are actually the same line, and there will be no solution if the lines are parallel. There will be exactly one solution if the lines are not the same and not parallel. Furthermore, we can assume that no row

of  $A$  can contain only zeros, otherwise there will be no solution when  $B$  has a non-zero entry in the same row.

In the latter case, there are two situations to consider: either one line is vertical, or neither line is vertical. (Both lines cannot be simultaneously vertical because we have assumed they are not parallel). If  $ax + by = k$  is vertical, then  $b = 0$  and  $a \neq 0$ ; also,  $d \neq 0$  since  $cx + dy = l$  cannot be vertical; consequently

$$ad - bc = ad - 0c = ad \neq 0.$$

On the other hand, if  $cx + dy = l$  is the vertical line, we have

$$ad - bc = a0 - bc = -bc \neq 0$$

because  $b \neq 0$  and  $c \neq 0$ . Finally, if neither line is vertical, then their slopes are given by the ratios  $-\frac{a}{b}$  and  $-\frac{c}{d}$ , and if the lines are not parallel, then these slopes are not equal:

$$\frac{-a}{b} \neq \frac{-c}{d} \implies ad \neq bc \implies ad - bc \neq 0.$$

In every possible case we came to the conclusion  $ad - bc \neq 0$ . If this condition fails, then the lines have the same slope, and therefore matrix equation  $AX = B$  will either have infinitely many solutions or no solutions, depending on the choice of  $B$ .

**EXERCISE 3:** Prove that if  $ad - bc = 0$ , then the matrix equation  $AX = B$  with unknown column vector  $X$  will have either no solution or infinitely many solutions, depending on the column vector  $B$ . Here,  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . (*Hint: Start by writing the matrix equation as a system of linear equations. Multiply the first equation by  $d$  and the second equation by  $b$ . Solve the system by the elimination (or addition) method. From this, deduce conditions on  $B$ .)*

The point is that the quantity  $ad - bc$  can be used to completely determine whether the matrix equation  $AX = B$  has exactly one solution or not. This quantity is called the **determinant** of the matrix, and it is written as  $\det(A)$ :

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

**EXERCISE 4:** Let  $A = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$ . Calculate the determinant of  $A$ .

Let us use  $0$  to denote a zero-matrix or zero-vector (i.e. a matrix all of whose entries are zero). The dimension of  $0$  will be

discernible from context. Observe that the equation  $AX = 0$  always has a solution;  $X = 0$ . This is called the **trivial solution** of the equation. Many times we will want to know if there is a non-trivial solution.

**EXERCISE 5:** Find at least one non-trivial solution of  $AX = 0$ , where  $A = \begin{bmatrix} 2 & -1 \\ -4 & 2 \end{bmatrix}$ .

We will also be interested in whether there are any non-trivial solutions of the equation

$$AX = \lambda X.$$

Here,  $\lambda$  is a scalar. Again, note that the trivial solution  $X = 0$  always works. If there is a non-trivial solution, it is called an **eigenvector** for the matrix  $A$ , and  $\lambda$  is called an associated **eigenvalue**.

**EXERCISE 6:** Let  $A = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$ . Verify that  $X = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  is an eigenvector for  $A$  with associated eigenvalue  $\lambda = 2$ . Verify that the matrix  $Y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  is not an eigenvector for any possible eigenvalue.

Next, we turn to the problem of finding eigenvalues and eigenvectors for a given matrix.

**EXERCISE 7:** Verify that the equation  $AX = \lambda X$  is equivalent to  $(A - \lambda I)X = 0$ .

Because an eigenvector of  $A$  is a non-trivial solution of  $(A - \lambda I)X = 0$ , we know that there will be such a solution if and only if  $\det(A - \lambda I) = 0$ . For otherwise, if  $\det(A - \lambda I) \neq 0$ , then the equation  $(A - \lambda I)X = 0$  has exactly one solution, meaning that the trivial solution  $X = 0$  is the only one. Therefore we can use the determinant of  $A - \lambda I$  to find all the eigenvalues of a matrix. The equation  $\det(A - \lambda I) = 0$  is called the **characteristic equation** for  $A$ .

**EXAMPLE 6:** Let  $A = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$ . Let us find all the eigenvalues of  $A$ . For any eigenvalue  $\lambda$ , we will have

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ &= \det \left( \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) \\ &= \det \begin{bmatrix} 2 - \lambda & 4 \\ 0 & 3 - \lambda \end{bmatrix} \\ &= (2 - \lambda)(3 - \lambda) - (4)(0) \\ &= (2 - \lambda)(3 - \lambda). \end{aligned}$$

The solutions of the characteristic equation  $(2 - \lambda)(3 - \lambda) = 0$  are  $\lambda = 2$  and  $\lambda = 3$ , so these are the eigenvalues of  $A$ .  $\square$

Once we know the eigenvalues, we can insert them for  $\lambda$ , and then finding eigenvectors to go with them is just a matter of solving a system of linear equations.

**EXAMPLE 7:** Let us find the eigenvalues and eigenvectors of  $A = \begin{bmatrix} 11 & -5 \\ -18 & 24 \end{bmatrix}$ . The characteristic equation is

$$\begin{aligned} 0 &= \det \left( \begin{bmatrix} 11 & -5 \\ -18 & 24 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det \begin{bmatrix} 11 - \lambda & -5 \\ -18 & 24 - \lambda \end{bmatrix} \\ &= (11 - \lambda)(24 - \lambda) - (-5)(-18) \\ &= \lambda^2 - 35\lambda + 264 - 90 \\ &= \lambda^2 - 35\lambda + 174. \end{aligned}$$

We can solve this equation using the quadratic formula:

$$\begin{aligned} \lambda &= \frac{35 \pm \sqrt{(-35)^2 - 4(1)(174)}}{2(1)} \\ &= \frac{35 \pm \sqrt{529}}{2} \\ &= \frac{35 \pm 23}{2}, \end{aligned}$$

so that  $\lambda_1 = 29$  and  $\lambda_2 = 6$  are the eigenvalues.

Next we identify the eigenvectors corresponding to each eigenvalue. Using the eigenvalue  $\lambda_1 = 29$  in the equation  $A - \lambda I = 0$

and writing the corresponding eigenvector as  $\begin{bmatrix} a \\ b \end{bmatrix}$ , we obtain

$$\begin{bmatrix} -18 & -5 \\ -18 & -5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Both lines of this matrix equation correspond to the algebraic equation  $-18a - 5b = 0$ , which implies  $b = -\frac{18a}{5}$ . Therefore, any matrix of the form  $\begin{bmatrix} a \\ -\frac{18a}{5} \end{bmatrix}$  is an eigenvector corresponding to the eigenvalue  $\lambda_1 = 29$ . This can also be stated by saying that any non-zero scalar multiple of  $\begin{bmatrix} 1 \\ -\frac{18}{5} \end{bmatrix}$  is an eigenvector.

Turning to the eigenvalue  $\lambda_2 = 6$ , the matrix equation  $(A - \lambda I)X = 0$  can be written as

$$\begin{bmatrix} 5 & -5 \\ -18 & 18 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and both lines of this matrix equation are algebraically equivalent to the scalar equation  $c - d = 0$ , so that  $c = d$ . Consequently, we see that non-zero vectors of the form  $\begin{bmatrix} a \\ a \end{bmatrix}$  (equivalently, non-zero scalar multiples of  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ) are eigenvectors corresponding to the eigenvalue  $\lambda_2 = 6$ .  $\square$

**EXERCISE 8:** Find the eigenvalues and corresponding eigenvectors for the matrix  $A = \begin{bmatrix} 4 & 0 \\ 6 & -2 \end{bmatrix}$ .

**EXERCISE 9:** Use the definition of eigenvalue and eigenvector to prove that, if  $X$  is an eigenvector corresponding to the eigenvalue  $\lambda$ , then every non-zero scalar multiple of  $X$  is also an eigenvector corresponding to  $\lambda$ .

Eigenvalues and eigenvectors can also be complex, even if the entries of the original matrix are all real.



**EXAMPLE 8:** Let  $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ . The characteristic equation is

$$\begin{aligned} 0 &= \det \left( \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) \\ &= \det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} \\ &= \lambda^2 + 1. \end{aligned}$$

The solutions of the equation  $\lambda^2 + 1 = 0$  are  $\lambda_1 = i$  and  $\lambda_2 = -i$ . For the first case, the eigenvectors  $\begin{bmatrix} a \\ b \end{bmatrix}$  satisfy

$$\begin{bmatrix} -i & 1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

implying the scalar equation  $-ia + b = 0$ , so  $b = ia$ . Thus the eigenvectors corresponding to  $\lambda_1 = i$  are the non-zero scalar multiples of  $\begin{bmatrix} 1 \\ i \end{bmatrix}$ . Turning to the eigenvalue  $\lambda_2 = -i$ , we have the matrix equation for the eigenvectors:

$$\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

implying  $b = -ia$ . Hence the eigenvectors corresponding to  $\lambda_2 = -i$  are the non-zero scalar multiples of  $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ .  $\square$

**EXERCISE 10:** Find the eigenvalues and corresponding eigenvectors for  $A = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix}$ .

We finish this section with some terminology from linear algebra. A collection of vectors  $\{X_1, \dots, X_n\}$  is said to be **linearly independent** if the only scalars  $c_1, \dots, c_n$  for which  $c_1X_1 + \dots + c_nX_n = 0$  are  $c_1 = \dots = c_n = 0$ . Otherwise, the collection is said to be **linearly dependent**.

**EXAMPLE 9:** The vectors  $X_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  and  $X_2 = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$  form a linearly dependent set because there are non-zero scalars  $c_1$  and  $c_2$  such that  $c_1X_1 + c_2X_2 = 0$ . (For example, take  $c_1 = 3$  and  $c_2 = 1$ .)  $\square$

**EXERCISE 11:** Determine whether the vectors  $X_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$  and  $X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  form a linearly independent set.

**EXERCISE 12:** Prove that the collection of vectors  $X_1, X_2$  is linearly dependent if and only if one of the vectors is a scalar multiple of the other.

If  $\{X_1, X_2\}$  is a linearly independent set of  $2 \times 1$  vectors, then any  $2 \times 1$  vector  $Y$  can be written as a **linear combination** of these vectors:  $Y = c_1X_1 + c_2X_2$  for some scalars  $c_1, c_2$ . In fact, these scalars are unique.

It turns out that, if  $X_1$  and  $X_2$  are both eigenvectors corresponding to different eigenvalues, then the collection  $\{X_1, X_2\}$  must be linearly independent.

The characteristic equation for a  $2 \times 2$  matrix  $A$  is a quadratic equation. It therefore has either one or two roots (possibly complex). If there are two roots, and if we find one eigenvector for each root, then that collection of eigenvectors must be linearly independent. On the other hand, if  $A$  has only one eigenvalue, we may not be able to find two eigenvectors that are linearly independent. In fact, the only time we can find two linearly independent eigenvectors corresponding to the same eigenvalue is if  $A$  is a scalar multiple of the identity matrix. A proof of this fact will be explored in the problem set.

## Problems

**PROBLEM 1:** The matrix  $\begin{bmatrix} 1 & r \\ 2 & 4 \end{bmatrix}$  has an eigenvalue  $\lambda_1 = 0$ . What is the value of  $r$ , and what is the other eigenvalue?

**PROBLEM 2:** The trace of a matrix  $A$  is the sum of the entries on the main diagonal (from top left to bottom right):  $\text{tr} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + d$ . Prove that the characteristic equation for a  $2 \times 2$  matrix  $A$  is  $\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$ .

**PROBLEM 3:** Let  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  have non-zero determinant. Define another matrix, denoted as  $A^{-1}$ , by  $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ . Prove that  $A^{-1}A = AA^{-1} = I$ . (The matrix  $A^{-1}$  is called the **multiplicative inverse**, or just **inverse**, of  $A$ .)

**PROBLEM 4:** Solve the system of linear equations

$$2x + y = 1$$

$$x + y = 2$$

by writing it as a matrix equation in the form  $AX = B$  and then multiplying both sides of this equation on their lefts by  $A^{-1}$  (as defined in the previous problem). (*It is important to multiply each side of the equation on the left for two reasons: because matrix multiplication is not commutative, you can simplify  $A^{-1}AX$  but not  $AXA^{-1}$ ; furthermore, because of the dimensions of  $A^{-1}$  and  $B$ , the product  $BA^{-1}$  is not even defined.*)

**PROBLEM 5:** Suppose that  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  has the property that  $AX = \lambda X$  for every  $2 \times 1$  vector  $X$ . Prove that  $A = \lambda I$ . (*Hint: Start by inserting  $X = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ; what does the resulting system of equations tell you about  $a$  and  $c$ ? What can you insert instead to obtain information about  $b$  and  $d$ ?*)