Some Notes on Linear Algebra

prepared for a first course in differential equations

Thomas L. Scofield
Department of Mathematics and Statistics
Calvin College

1998

The purpose of these notes is to present some of the notions of linear algebra that are necessary for the study of systems of linear first-order differential equations. These notes are not to be thought of as a comprehensive presentation of linear algebra, itself quite a broad subject area, nor are those results from linear algebra discussed here accompanied by proofs. Neither do I assert that these notes would serve equally well as the foundation for studying systems of ODEs in all courses and in conjunction with all texts. Rather, this unit has been written as a supplement to the text *Elementary Differential Equations, 6th Edition* by Boyce and DiPrima.

1 Matrices and Vectors

An $m \times n$ matrix is an array (or table) of entries consisting of m rows and n columns. The symbol $\mathcal{M}_{m\times n}(\mathbb{R})$ will be used to denote the set of all $m \times n$ matrices whose entries are real numbers, while $\mathcal{M}_{m\times n}(\mathbb{C})$ will denote the $m \times n$ matrices with entries which are complex numbers. Thus

$$\begin{bmatrix} 2 & 0 \\ -1 & 5 \\ 3 & 0 \end{bmatrix}, \qquad \begin{bmatrix} -1 & 3 & 7 \end{bmatrix}, \qquad \begin{bmatrix} -i & 3+2i & 7+2i \\ -1 & 4-i & -1+i \\ 0 & 0 & -2i \end{bmatrix}$$

are elements in $\mathcal{M}_{3\times 2}(\mathbb{R})$, $\mathcal{M}_{1\times 3}(\mathbb{R})$ and $\mathcal{M}_{3\times 3}(\mathbb{C})$ respectively. Of course, it may also be said that the first two of these are in $\mathcal{M}_{3\times 2}(\mathbb{C})$, $\mathcal{M}_{1\times 3}(\mathbb{C})$, respectively, since the set of real numbers is contained in the set of complex numbers.

We may multiply a matrix by a number (or scalar) in the following way. Three times the first matrix above is defined to be

$$3\begin{bmatrix} 2 & 0 \\ -1 & 5 \\ 3 & 0 \end{bmatrix} := \begin{bmatrix} 6 & 0 \\ -3 & 15 \\ 9 & 0 \end{bmatrix},$$

while i times the last of them is

$$i\begin{bmatrix} -i & 3+2i & 7+2i \\ -1 & 4-i & -1+i \\ 0 & 0 & -2i \end{bmatrix} := \begin{bmatrix} 1 & -2+3i & -2+7i \\ -i & 1+4i & -1-i \\ 0 & 0 & 2 \end{bmatrix}.$$

This process is called *scalar multiplication*. Our scalar may even be a function, so that

$$e^{-3t}\begin{bmatrix} -6 & 1 \\ 2 & 1 \end{bmatrix} := \begin{bmatrix} -6e^{-3t} & e^{-3t} \\ 2e^{-3t} & e^{-3t} \end{bmatrix}.$$

We may also add and subtract two matrices, as long as they have the same number of rows and columns. In such a case we define the sum of two matrices to be the sum of its individual elements, as in

$$\begin{bmatrix} -6 & 1 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ -1 & 5 \end{bmatrix} := \begin{bmatrix} -3 & 1 \\ 1 & 6 \end{bmatrix}.$$

The difference of two matrices is defined similarly, so that

$$\begin{bmatrix} 2 & 3-i & 0 \\ 1+i & 7i & 6+2i \end{bmatrix} - \begin{bmatrix} -6i & -3-i & 2+i \\ 0 & 2i & 1+i \end{bmatrix} := \begin{bmatrix} 2+6i & 6 & -2-i \\ 1+i & 5i & 5+i \end{bmatrix}.$$

Notice that, based upon these definitions for addition and subtraction, it is reasonable to call the $m \times n$ matrix whose entries are all zeros an *additive identity* for $\mathcal{M}_{m\times n}(\mathbb{R})$ (or $\mathcal{M}_{m\times n}(\mathbb{C})$), since adding it to any matrix $\mathbf{A} \in \mathcal{M}_{m\times n}(\mathbb{R})$ results in a sum equal to $\mathbf{A}$.

It is customary to refer to elements in $\mathcal{M}_{m\times n}(\mathbb{R})$ (or $\mathcal{M}_{m\times n}(\mathbb{C})$) by single, boldface capital letters such as $\mathbf{A}$. Nevertheless, we must keep in mind that $\mathbf{A}$ is a table of entries

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

Here the entry of $\mathbf{A}$ in the i^{th}-row, j^{th}-column is denoted by a_{ij}. If this is how we wish to refer to the entries of $\mathbf{A}$, often we say so by indicating that $\mathbf{A} = (a_{ij})$. Thus, the statement $\mathbf{B} = (b_{ij})$ means that we intend to refer to the entry of $\mathbf{B}$ found in the i^{th}-row, j^{th}-column as b_{ij}.

If we rewrite an $m \times n$-matrix with all of its columns now written as rows, we get a new $n \times m$ matrix called the *transpose* matrix. Thus, if

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

then its transpose matrix $\mathbf{A}^T$ is

$$\mathbf{A}^T = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 4 & -3 \end{bmatrix}.$$

Matrices that have only one column are usually referred to as *vectors*. While vectors are simply special matrices, it is customary to use lower-case letters in naming them, rather than upper-case. An arbitrary n-vector $\mathbf{v}$ takes the form

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

Of course, it may be convenient to display a vector horizontally (as a $1 \times n$ matrix) instead of vertically. Literally speaking, this "row" format is the transpose of the column format we listed above, and we should write

$$\mathbf{v}^T = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}.$$

In what follows, you will often see

$$\mathbf{v} = (v_1, v_2, \ldots, v_n).$$

By this, we mean that we are still talking about the column vector (notice it is denoted as $\mathbf{v}$ and not $\mathbf{v}^T$), but wish not to take up as much space on the page as would be required if we *wrote* it in column format. In these cases, the commas will also assist in making the distinction.

We define the *dot product* of two n-vectors

$$\mathbf{u} = (u_1, u_2, \ldots, u_n) \qquad \text{and} \qquad \mathbf{v} = (v_1, v_2, \ldots, v_n)$$

to be

$$\mathbf{u} \cdot \mathbf{v} := u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{j=1}^{n} u_j v_j.$$

Notice that for this definition to make sense it is necessary that the two vectors have the same number of elements. Thus

$$(2, 4, -1, 0) \cdot \begin{bmatrix} -1 \\ 5 \\ 0 \\ 3 \end{bmatrix} = (2)(-1) + (4)(5) + (-1)(0) + (0)(3) = 18.$$

While it may seem odd that the first vector in the example above was written in row format and the second in column format, we did this because this leads to the way we wish to define multiplication between matrices. When we multiply two matrices, the product is a matrix whose elements arise from dot products between the rows of the first (matrix) factor and columns of the second. An immediate consequence of this: if $\mathbf{A}$ and $\mathbf{B}$ are matrices, the product $\mathbf{AB}$ makes sense precisely when the number of columns in $\mathbf{A}$ is equal to the number of rows in $\mathbf{B}$. To be clearer about how such a matrix product is achieved, suppose $\mathbf{A}$ is an $m \times n$ matrix while $\mathbf{B}$ is an $n \times p$ matrix. If we write

$$\mathbf{A} = \begin{bmatrix} \mathbf{r_1} & \rightarrow \\ \mathbf{r_2} & \rightarrow \\ \vdots & \\ \mathbf{r_m} & \rightarrow \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_2} & \cdots & \mathbf{c_p} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix},$$

with each of the rows $\mathbf{r_i}$ of $\mathbf{A}$ (considered as column vectors) having n components and likewise each of the columns $\mathbf{c_j}$ of $\mathbf{B}$, then their product is an $m \times p$ matrix whose entry in the i^{th}-row, j^{th}-column is obtained by taking the dot product of $\mathbf{r_i}$ with $\mathbf{c_j}$. Thus if

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \\ -5 & 1 \\ 7 & -4 \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} 3 & 1 & 0 \\ -2 & 4 & 10 \end{bmatrix},$$

then the product $\mathbf{AB}$ will be the 4×3 matrix

$$\begin{aligned}
\mathbf{AB} &= \begin{bmatrix} (2,-1)\cdot\begin{bmatrix}3\\-2\end{bmatrix} & (2,-1)\cdot\begin{bmatrix}1\\4\end{bmatrix} & (2,-1)\cdot\begin{bmatrix}0\\10\end{bmatrix} \\ (0,3)\cdot\begin{bmatrix}3\\-2\end{bmatrix} & (0,3)\cdot\begin{bmatrix}1\\4\end{bmatrix} & (0,3)\cdot\begin{bmatrix}0\\10\end{bmatrix} \\ (-5,1)\cdot\begin{bmatrix}3\\-2\end{bmatrix} & (-5,1)\cdot\begin{bmatrix}1\\4\end{bmatrix} & (-5,1)\cdot\begin{bmatrix}0\\10\end{bmatrix} \\ (7,-4)\cdot\begin{bmatrix}3\\-2\end{bmatrix} & (7,-4)\cdot\begin{bmatrix}1\\4\end{bmatrix} & (7,-4)\cdot\begin{bmatrix}0\\10\end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 8 & -2 & -10 \\ -6 & 12 & 30 \\ -17 & -1 & 10 \\ 29 & -9 & -40 \end{bmatrix}.
\end{aligned}$$

Notice that if $\mathbf{A} \in \mathcal{M}_{2\times4}(\mathbb{R})$ and $\mathbf{B} \in \mathcal{M}_{4\times3}(\mathbb{R})$ then the product $\mathbf{AB}$ is defined, but the product $\mathbf{BA}$ is not. This is because the number of columns in $\mathbf{B}$ is unequal to the number of rows in $\mathbf{A}$. Thus, for it to be possible to multiply two matrices, one of which is in $\mathcal{M}_{m\times n}(\mathbb{R})$, in either order, it is necessary that the other be in $\mathcal{M}_{n\times m}(\mathbb{R})$. In particular, the products $\mathbf{CD}$ and $\mathbf{DC}$ are well-defined (though, in general, not equal) if $\mathbf{C}$ and $\mathbf{D}$ are both square of the same dimension — that is, both are in $\mathcal{M}_{n\times n}(\mathbb{R})$. It is interesting to note that the matrix

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

is a *multiplicative identity* for $\mathcal{M}_{n\times n}(\mathbb{R})$; that is, given any matrix $\mathbf{C} \in \mathcal{M}_{n\times n}(\mathbb{R})$,

$$\mathbf{CI_n} = \mathbf{I_nC} = \mathbf{C}.$$

Our method for defining products of matrices lends itself well to writing algebraic systems of linear equations as matrix equations. Consider the system of equations

$$\begin{aligned} 2x + 5y &= 11 \\ 7x - y &= -1. \end{aligned}$$

We can think of the solution (x, y) of this system as a point of intersection between two lines. The matrix equation

$$\begin{bmatrix} 2 & 5 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 11 \\ -1 \end{bmatrix}$$

expresses the exact same equations as appear in the system. Its solution is the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ whose entries match the x and y coordinates for the point of intersection.

2 Gaussian Elimination: A New Spin on the Solution of n Linear Equations in n Unknowns

By now systems of two equations in two variables are quite familiar to us. We have studied them in practically every mathematics course prior to the Calculus. We have been taught to find solutions of such systems via *elimination* or *substitution*, and we understand these solutions to be points of intersection between two graphs. In particular, when the two equations are *linear* in two variables,

$$\begin{aligned} ax + by &= c \\ \alpha x + \beta y &= \gamma. \end{aligned}$$

the solution(s) (assuming there are any) are points of intersection between two lines. The solution(s) of three linear equations

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{33}z &= b_2 \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned}$$

correspond to points of intersection of three planes.

Example: Let us explore the method of elimination and see how we might carry it out using matrices. Consider, for example, the problem of finding intersections between the two lines:

$$\begin{aligned} 2x - 3y &= 7 \\ 3x + 5y &= 1 \end{aligned} \qquad \text{or, in matrix form} \qquad \begin{bmatrix} 2 & -3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}.$$

We have the sense that, with two unknowns x and y, two equations are enough to determine their values. The method of elimination taught in elementary algebra courses goes as follows. In each step, we write what a high-school algebra student might write, and then a matrix formulation of the same set of equations.

Step 1. Leaving the top equation alone, we multiply through the bottom one by (-2/3):

$$\begin{aligned} 2x - 3y &= 7 \\ -2x - (10/3)y &= -2/3. \end{aligned} \qquad \text{matrix form:} \quad \begin{bmatrix} 2 & -3 \\ -2 & -10/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ -2/3 \end{bmatrix}.$$

Step 2. We add the two equations, eliminating the x term. (Of course, we could have performed Step 1 differently so that, after this step, the y terms have been eliminated.) What has always been tacitly assumed (but perhaps not overtly mentioned) is that our previous equations do not get thrown away. (We always return to at least one of them once we know the value of y so as to get the value of x.) Nevertheless, as we accumulate more equations (the original two, the "rescaled" version of the second equation and, now, another one resulting from adding the previous two), we have the sense that the problem still has the same number of degrees of freedom (the unknowns) and the same number of constraints (equations) as it had when we began. The new equations do not add new information over what was built into the original problem (i.e., they do not change the solutions in any way). Thus, along with the equation we get by adding the previous two, we carry forward the first equation once again:

$$\begin{aligned} 2x - 3y &= 7 \\ -(19/3)y &= 19/3 \end{aligned} \qquad \text{matrix form:} \quad \begin{bmatrix} 2 & -3 \\ 0 & -19/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 19/3 \end{bmatrix}.$$

Step 3. Now that the x-term has been eliminated in the second equation, we may get the value of y by dividing this equation by (-19/3):

$$\begin{aligned} 2x - 3y &= 7 \\ y &= -1 \end{aligned} \qquad \text{matrix form:} \quad \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}.$$

Step 4. At some stage we probably divide through the first equation by 2 making it easier to determine the x value from the (now known) y value.

$$\begin{aligned} x - (3/2)y &= 7/2 \\ y &= -1 \end{aligned} \qquad \text{matrix form:} \quad \begin{bmatrix} 1 & -3/2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7/2 \\ -1 \end{bmatrix}.$$

Gaussian elimination is a method for solving systems of linear equations which carries out the same process as above, keeping track only of what happens to the coefficient matrix and the vector on the right-hand side. It may appear strange at first look, but compare the following string of matrices to those which appeared above. (These matrices are not equal, so we do not use the symbol "=" between them. Nevertheless, they do represent equivalent systems of linear equations — ones that have the same point(s) of intersection. The symbol $\sim$ indicates this.)

$$\text{(augmented matrix)} \qquad \left[\begin{array}{cc|c} 2 & -3 & 7 \\ 3 & 5 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 2 & -3 & 7 \\ -2 & -10/3 & -2/3 \end{array}\right]$$
$$\sim \left[\begin{array}{cc|c} 2 & -3 & 7 \\ 0 & -19/3 & 19/3 \end{array}\right]$$
$$\sim \left[\begin{array}{cc|c} 2 & -3 & 7 \\ 0 & 1 & -1 \end{array}\right]$$
$$\sim \left[\begin{array}{cc|c} 1 & -3/2 & 7/2 \\ 0 & 1 & -1 \end{array}\right]. \qquad \text{(row echelon form)}$$

One carries out Gaussian elimination generally with the goal in mind to arrive at a final matrix which has, to the left of the "divider", a stair-step appearance, known as *row-echelon form*:

> In each row the leading (nonzero) number is a 1, and the leading 1 in the next row appears in a column farther to the right than in the current row.

Check that the "augmented matrix" (the matrix we got by adding a divider and the right-hand vector to the original coefficient matrix) in the above example has been processed until it fits this description. Notice that, from this row-echelon form, we can easily get the values of x and y:

$$y = -1, \qquad \text{and} \qquad x = 7/2 + (3/2)(-1) = 2.$$

If row-echelon form is the goal, what operations are we allowed to perform to get there? They are exact analogs for the kinds of operations we might perform on a system of equations:

(i) Exchange two rows.

(ii) Multiply a row by a nonzero number.

(iii) Add a multiple of one row to another, replacing the latter row with the result.

Example: Use Gaussian elimination to solve the system of equations (representing three planes)

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

First we form the augmented matrix. Then, we perform row operations as described at right to reduce to row-echelon form.

$$\left[\begin{array}{ccc|c} 0 & -3 & 1 & 1 \\ 1 & 0 & 3 & -1 \\ 2 & -1 & 4 & 0 \end{array}\right] \quad \underset{\sim}{\mathbf{r}_1 \leftrightarrow \mathbf{r}_2} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & -3 & 1 & 1 \\ 2 & -1 & 4 & 0 \end{array}\right] \quad \text{(i.e., we exchange } \mathbf{r}_1 \text{ and } \mathbf{r}_2\text{)}$$

$$\underset{\sim}{-2\mathbf{r}_1 + \mathbf{r}_3 \longrightarrow \mathbf{r}_3} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & -3 & 1 & 1 \\ 0 & -1 & -2 & 2 \end{array}\right] \quad \text{(i.e., replace } \mathbf{r}_3 \text{ with } -2\mathbf{r}_1 + \mathbf{r}_3\text{)}$$

$$\underset{\sim}{(-\mathbf{r}_3) \longrightarrow \mathbf{r}_3} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & -3 & 1 & 1 \\ 0 & 1 & 2 & -2 \end{array}\right] \quad \text{(i.e., replace } \mathbf{r}_3 \text{ by } (-\mathbf{r}_3)\text{)}$$

$$\underset{\sim}{\mathbf{r}_2 \leftrightarrow \mathbf{r}_3} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & -3 & 1 & 1 \end{array}\right]$$

$$\underset{\sim}{3\mathbf{r}_2 + \mathbf{r}_3 \longrightarrow \mathbf{r}_3} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 7 & -5 \end{array}\right]$$

$$\underset{\sim}{(1/7)\mathbf{r}_3 \longrightarrow \mathbf{r}_3} \quad \left[\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 1 & -5/7 \end{array}\right].$$

The last of these is in row-echelon form. From this we get that

$$\begin{aligned} z &= -5/7 \\ y &= -2 - (2)(-5/7) = -4/7 \\ x &= -1 - (3)(-5/7) = 8/7. \end{aligned}$$

The point $(-5/7, -4/7, 8/7)$ is the (only) point of intersection between the three planes.

3 Determinants and the Solvability of Linear Systems

In the last section we learned how to use Gaussian elimination to solve linear systems of n equations in n unknowns. The section completely side-stepped one important question: that of whether a system has a solution and, if so, whether it is unique.

Consider the case of two lines in the plane

$$\begin{aligned} ax + by &= e \\ cx + dy &= f. \end{aligned} \tag{1}$$

In fact, we know that intersections between two lines can happen in any of three different ways:

1. the lines intersect at a unique point (i.e., solution exists and is unique),
2. the lines are coincident (that is, the equations represent the same line and there are infinitely many points of intersection; in this case a solution exists, but is not unique), or
3. the lines are parallel but not coincident (so that no solution exists).

Experience has taught us that it is quite easy to decide which of these situations we are in before ever attempting to solve a linear system of two equations in two unknowns. For instance, the system

$$\begin{aligned} 3x - 5y &= 9 \\ -5x + \frac{25}{3}y &= -15 \end{aligned}$$

obviously contains two representations of the same line (since one equation is a constant multiple of the other) and will have infinitely many solutions. In contrast, the system

$$\begin{aligned} x + 2y &= -1 \\ 2x + 4y &= 5 \end{aligned}$$

will have no solutions. This is the case because, while the left sides of each equation — the sides that contain the coefficients of x and y which determine the slopes of the lines — are in proportion to one another, the right sides are not in the same proportion. As a result, these two lines will have the same slopes but not the same y-intercepts. Finally, the system

$$\begin{aligned} 2x + 5y &= 11 \\ 7x - y &= -1. \end{aligned}$$

will have just one solution (one point of intersection), as the left sides of the equations are not at all in proportion to one another.

What is most important about the preceding discussion is that we can distinguish situation 1 (the lines intersecting at one unique point) from the others simply by looking at the coefficients a, b, c and d from equation (1). In particular, we can determine the ratios $a : c$

and $b : d$ and determine whether these ratios are the same or different. Equivalently, we can look at whether the quantity

$$ad - bc$$

is zero or not. If $ad - bc \neq 0$ then the system has one unique point of intersection, but if $ad - bc = 0$ then the system either has no points or infinitely many points of intersection. If we write equation (1) as a matrix equation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix},$$

we see that the quantity $ad - bc$ is dependent only upon the coefficient matrix. Since this quantity "determines" whether or not the system has a unique solution, it is called the *determinant* of the coefficient matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

and is sometimes abbreviated as $\det(\mathbf{A})$, $|\mathbf{A}|$ or

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

While it is quite easy for us to determine in advance the number of solutions which arise from a system of two linear equations in two unknowns, the situation becomes a good deal more complicated if we add another variable and another equation. The solutions of such a system

$$\begin{aligned} ax + by + cz &= l \\ dx + ey + fz &= m \\ ex + hy + kz &= n \end{aligned} \tag{2}$$

can be thought of as points of intersection between three planes. Again, there are several possibilities:

1. the planes intersect at a unique point,
2. the planes intersect along a line,
3. the planes intersect in a plane, or
4. the planes do not intersect.

It seems reasonable to think that situation 1 can once again be distinguished from the other three simply by performing some test on the numbers a, b, c, d, e, f, g, h and k. As in the case of the system (1), perhaps if we write system (2) as the matrix equation

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} l \\ m \\ n \end{bmatrix},$$

we will be able to define an appropriate quantity $\det(\mathbf{A})$ that depends only on the coefficient matrix

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

in such a way that, if $\det(\mathbf{A}) \neq 0$ then the system has a unique solution (situation 1), but if $\det(\mathbf{A}) = 0$ then one of the other situations (2–4) is in effect.

Indeed it is possible to define $\det(\mathbf{A})$ for a square matrix $\mathbf{A}$ of arbitrary dimension. For our purposes, we do not so much wish to give a rigorous definition of such a determinant as we do wish to be able to find it. As of now, we do know how to find it for a 2×2 matrix:

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

For square matrices of dimension larger than 2 we will find the determinant using *cofactor expansion.*

Let $\mathbf{A} = (a_{ij})$ be an arbitrary $n \times n$ matrix; that is,

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

We define the (i,j)-*minor* of $\mathbf{A}$, M_{ij}, to be the determinant of the matrix resulting from crossing out the i^{th} row and the j^{th} column of $\mathbf{A}$. Thus, if

$$\mathbf{B} = \begin{bmatrix} 1 & -4 & 3 \\ -3 & 2 & 5 \\ 4 & 0 & -1 \end{bmatrix},$$

we have nine possible minors M_{ij} of $\mathbf{B}$, two of which are

$$M_{21} = \begin{vmatrix} -4 & 3 \\ 0 & -1 \end{vmatrix} = 4 \qquad \text{and} \qquad M_{33} = \begin{vmatrix} 1 & -4 \\ -3 & 2 \end{vmatrix} = -10.$$

A concept that is related to the (i,j)-minor is the (i,j)-*cofactor*, C_{ij}, which is defined to be

$$C_{ij} := (-1)^{i+j} M_{ij}.$$

Thus, the matrix $\mathbf{B}$ above has 9 cofactors C_{ij}, two of which are

$$C_{21} = (-1)^{2+1} M_{21} = -4 \qquad \text{and} \qquad C_{33} = (-1)^{3+3} M_{33} = -10.$$

Armed with the concept of cofactors, we are prepared to say how the determinant of an arbitrary square matrix $\mathbf{A} = (a_{ij})$ is found. It may be found by expanding in cofactors along the i^{th} row:

$$\det(\mathbf{A}) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = \sum_{k=1}^{n} a_{ik}C_{ik};$$

or, alternatively, it may be found by expanding in cofactors along the j^{th} column:

$$\det(\mathbf{A}) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} = \sum_{k=1}^{n} a_{kj}C_{kj}.$$

Thus, $\det(\mathbf{B})$ for the 3×3 matrix $\mathbf{B}$ above is

$$\begin{aligned}
\det(\mathbf{B}) = \begin{vmatrix} 1 & -4 & 3 \\ -3 & 2 & 5 \\ 4 & 0 & -1 \end{vmatrix} &= 4(-1)^{3+1}\begin{vmatrix} -4 & 3 \\ 2 & 5 \end{vmatrix} \\
&\quad +(0)(-1)^{3+2}\begin{vmatrix} 1 & 3 \\ -3 & 5 \end{vmatrix} + (-1)(-1)^{3+3}\begin{vmatrix} 1 & -4 \\ -3 & 2 \end{vmatrix} \\
&= 4(-20-6) + 0 - (2-12) \\
&= -94.
\end{aligned}$$

Here we found $\det(\mathbf{B})$ via a cofactor expansion along the third row. You should verify that a cofactor expansion along any of the other two rows would also lead to the same result. Had we expanded in cofactors along one of the columns, for instance column 2, we would have

$$\begin{aligned}
\det(\mathbf{B}) &= (-4)(-1)^{1+2}\begin{vmatrix} -3 & 5 \\ 4 & -1 \end{vmatrix} + (2)(-1)^{2+2}\begin{vmatrix} 1 & 3 \\ 4 & -1 \end{vmatrix} + (0)(-1)^{3+2}\begin{vmatrix} 1 & 3 \\ -3 & 5 \end{vmatrix} \\
&= 4(3-20) + 2(-1-12) + 0 \\
&= -94.
\end{aligned}$$

This process can be used iteratively on larger square matrices. For instance

$$\begin{aligned}
\begin{vmatrix} 3 & 1 & 2 & 0 \\ -1 & 0 & 5 & -4 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & -3 & 1 \end{vmatrix} &= (0)(-1)^{4+1}\begin{vmatrix} 1 & 2 & 0 \\ 0 & 5 & -4 \\ 1 & 0 & -1 \end{vmatrix} + (0)(-1)^{4+2}\begin{vmatrix} 3 & 2 & 0 \\ -1 & 5 & -4 \\ 1 & 0 & -1 \end{vmatrix} \\
&\quad +(-3)(-1)^{4+3}\begin{vmatrix} 3 & 1 & 0 \\ -1 & 0 & -4 \\ 1 & 1 & -1 \end{vmatrix} + (1)(-1)^{4+4}\begin{vmatrix} 3 & 1 & 2 \\ -1 & 0 & 5 \\ 1 & 1 & 0 \end{vmatrix},
\end{aligned}$$

where our cofactor expansion of the original determinant of a 4×4 matrix along its fourth row expresses it in terms of the determinants of several 3×3 matrices. We may proceed to find these latter determinants using cofactor expansions as well:

$$\begin{aligned}
\begin{vmatrix} 3 & 1 & 0 \\ -1 & 0 & -4 \\ 1 & 1 & -1 \end{vmatrix} &= (3)(-1)^{1+1}\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} + (1)(-1)^{1+2}\begin{vmatrix} -1 & -4 \\ 1 & -1 \end{vmatrix} + (0)(-1)^{1+3}\begin{vmatrix} -1 & 0 \\ 1 & 1 \end{vmatrix} \\
&= 3(0+4) - (1+4) + 0 \\
&= 7,
\end{aligned}$$

where we expanded in cofactors along the first row, and

$$\begin{aligned}
\begin{vmatrix} 3 & 1 & 2 \\ -1 & 0 & 5 \\ 1 & 1 & 0 \end{vmatrix} &= (1)(-1)^{1+2}\begin{vmatrix} -1 & 5 \\ 1 & 0 \end{vmatrix} + (0)(-1)^{2+2}\begin{vmatrix} 3 & 2 \\ 1 & 0 \end{vmatrix} + (1)(-1)^{3+2}\begin{vmatrix} 3 & 2 \\ -1 & 5 \end{vmatrix} \\
&= -(0-5) + 0 - (15+2) \\
&= -12,
\end{aligned}$$

where this cofactor expansion was carried out along the second column. Thus

$$\begin{vmatrix} 3 & 1 & 2 & 0 \\ -1 & 0 & 5 & -4 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & -3 & 1 \end{vmatrix} = 0 + 0 + (3)(7) + (1)(-12) = 9.$$

The matrices whose determinants we computed in the preceding paragraph are the coefficient matrices for the two linear systems

$$\begin{aligned}
x_1 - 4x_2 + 3x_3 &= b_1 \\
-3x_1 + 2x_2 + 5x_3 &= b_2 \\
4x_1 - x_3 &= b_3
\end{aligned}$$

and

$$\begin{aligned}
3x_1 + x_2 + 2x_3 &= b_1 \\
-x_1 + 5x_3 - 4x_4 &= b_2 \\
x_1 + x_2 - x_4 &= b_3 \\
-3x_3 + x_4 &= b_4
\end{aligned}$$

respectively; that is, if we write each of these systems as a matrix equation of the form $\mathbf{Ax} = \mathbf{b}$, with

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix},$$

then the coefficient matrix $\mathbf{A}$ in each case is one for which we have already computed the determinant and found it to be nonzero. This means that no matter what values are used for b_1, b_2, b_3 (and b_4 in the latter system) there is exactly one solution for the unknown vector $\mathbf{x}$.

4 Linear Independence and Span

We define vectors as being linearly independent in the same way we defined linear independence for functions on an interval. Given a collection of n-vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k}\}$, we say that this collection is *linearly dependent* if there are constants $c_1, c_2, \ldots, c_k$ not all of which are zero such that

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_k\mathbf{v_k} = \mathbf{0}.$$

Here the $c_j\mathbf{v_j}$ represent scalar multiples of the $\mathbf{v_j}$ for $j = 1, 2, \ldots, k$, and the expression

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_l\mathbf{v_k}$$

is called a *linear combination* of the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$. Clearly we can get $\mathbf{0}$ to be a linear combination of $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ by taking $c_1 = c_2 = \cdots = c_k = 0$. If this is the *only* linear combination that results in $\mathbf{0}$, then the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ are said to be *linearly independent*.

Example: The 2-vectors $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are linearly independent.

> To see this, we suppose that we have a linear combination $c_1\mathbf{e_1} + c_2\mathbf{e_2} = \mathbf{0}$; that is,
>
> $$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$
>
> is equal to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Then clearly $c_1 = 0$ and $c_2 = 0$. There are no other possible values for c_1 and c_2.

Example: The set of three 2-vectors $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{v}\}$, where $\mathbf{e_1}$ and $\mathbf{e_2}$ are the same as in the last example and $\mathbf{v}$ is any other 2-vector, is linearly dependent.

> To show this, we really need to consider two cases. The first case is the more trivial case, where the vector $\mathbf{v}$ is the zero vector — the 2-vector whose entries are both zero. In this case, we get $c_1\mathbf{e_1} + c_2\mathbf{e_2} + c_3\mathbf{v} = \mathbf{0}$ by taking $c_1 = 0$, $c_2 = 0$ and $c_3 \neq 0$.
>
> In the other case, where $\mathbf{v}$ is not the zero vector, let $v = (v_1 \;\; v_2)^T$ where, as we have assumed, not both of the entries are zero. Taking $c_1 = -v_1$, $c_2 = -v_2$ and $c_3 = 1$ we again get that $c_1\mathbf{e_1} + c_2\mathbf{e_2} + c_3\mathbf{v} = \mathbf{0}$.

What the last two examples show is that the set of vectors $\{\mathbf{e_1}, \mathbf{e_2}\}$ is linearly independent, but if this set is augmented with any additional 2-vector the set becomes linearly dependent. The reason for this is that all 2-vectors can be written as linear combinations of $\mathbf{e_1}$ and $\mathbf{e_2}$. Another way to say this is to say that the set $\{\mathbf{e_1}, \mathbf{e_2}\}$ *spans* the set of 2-vectors, or that every 2-vector is in *span*$\{\mathbf{e_1}, \mathbf{e_2}\}$. It is easy to show that other sets of 2-vectors, such as $\{(1, 1), (1, -1)\}$ for instance, are linearly independent and span the set of 2-vectors.

Example: Characterize the vectors that are in *span*$\{(1, 0, 1), (0, 1, 0)\}$, and propose a 3-vector $\mathbf{v}$ which, when joined to the set above, forms a set that is linearly independent.

The span of $\{(1,0,1),(0,1,0)\}$ is the set of all linear combinations of the form

$$c_1(1,0,1)+c_2(0,1,0)=(c_1,c_2,c_1);$$

that is, it is the set of all vectors whose first and third components are the same. There are many vectors that can be adjoined to this set while maintaining the linear independence that the set now has. Several possible vectors are $(1,0,0)$, $(1,0,2)$ and $(1,1,2)$. The proof of this fact is left as an exercise.

It is not possible to adjoin two vectors to the original set of vectors from the previous example and maintain linear independence. This is due to an important theorem which says:

Theorem: A linearly independent set of n-vectors can contain at most n vectors.

A related theorem discusses the least number of vectors necessary for spanning the n-vectors.

Theorem: No set containing fewer than n vectors can span the set of all n-vectors.

If we have a set of n n-vectors, there is another way besides the method we have used in examples thus far to test whether the set is linearly independent or not. Let us consider the n-vectors $\{\mathbf{v_1},\mathbf{v_2},\ldots,\mathbf{v_n}\}$. In determining whether or not we have linear independence, we consider whether there are constants $c_1,\ldots,c_n$ not all zero for which the linear combination

$$c_1\mathbf{v_1}+c_2\mathbf{v_2}+\cdots+c_n\mathbf{v_n}$$

is $\mathbf{0}$. We can look at this question in another way using matrices. In particular, if we create an $n\times n$ matrix $\mathbf{V}$ whose columns are the individual (column) vectors $\mathbf{v_j}, j=1,\ldots,n$ — that is,

$$\mathbf{V}:=\begin{bmatrix}\mathbf{v_1} & \mathbf{v_2} & \cdots & \mathbf{v_n}\\ \downarrow & \downarrow & & \downarrow\end{bmatrix}$$

— and if $\mathbf{c}$ is the vector of coefficients $\mathbf{c}:=(c_1,c_2,\ldots,c_n)$, then another way to ask the same question is to ask whether the matrix equation $\mathbf{Vc}=\mathbf{0}$ has any other solution besides $\mathbf{c}=\mathbf{0}$. Now we know that if a matrix equation $\mathbf{Ax}=\mathbf{b}$ has a solution at all, it has either one (unique) solution or infinitely many solutions, and we can use the determinant of the coefficient matrix $\mathbf{A}$ to decide which of these is the case. For the equation $\mathbf{Vc}=\mathbf{0}$, then, $\det(\mathbf{V})=0$ tells us there is a nonzero vector $\mathbf{c}$ satisfying $\mathbf{Vc}=\mathbf{0}$, and hence the set of vectors $\{\mathbf{v_1},\mathbf{v_2},\ldots,\mathbf{v_n}\}$ is linearly dependent. If $\det(\mathbf{V})\neq 0$, then the set of vectors is linearly independent.

Example: Determine whether or not the set of vectors $\{(1,2,2),(-1,1,0),(1,8,6)\}$ is linearly independent.

As per the discussion above, the question can be settled by looking at the determinant

$$
\begin{aligned}
\begin{vmatrix} 1 & -1 & 1 \\ 2 & 1 & 8 \\ 2 & 0 & 6 \end{vmatrix} &= 2(-1)^{3+1}\begin{vmatrix} -1 & 1 \\ 1 & 8 \end{vmatrix} \\
&\quad +(0)(-1)^{3+2}\begin{vmatrix} 1 & 1 \\ 2 & 8 \end{vmatrix} + 6(-1)^{3+3}\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} \\
&= 2(-9) + 0 + 6(3) \\
&= 0.
\end{aligned}
$$

Since this determinant is zero, the set of vectors is linearly dependent.

5 Eigenvalues and Eigenvectors

The product $\mathbf{Ax}$ of a matrix $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$ and an n-vector $\mathbf{x}$ is itself an n-vector. Of particular interest in many settings (of which differential equations is one) is the following question:

> For a given matrix $\mathbf{A}$, what are the vectors $\mathbf{x}$ for which the product $\mathbf{Ax}$ is a scalar multiple of $\mathbf{x}$? That is, what vectors $\mathbf{x}$ satisfy the equation
>
> $$\mathbf{Ax} = \lambda\mathbf{x}$$
>
> for some scalar λ?

It should immediately be clear that, no matter what $\mathbf{A}$ and λ are, the vector $\mathbf{x} = \mathbf{0}$ (that is, the vector whose elements are all zero) satisfies this equation. With such a trivial answer, we might ask the question again in another way:

> For a given matrix $\mathbf{A}$, what are the *nonzero* vectors $\mathbf{x}$ that satisfy the equation
>
> $$\mathbf{Ax} = \lambda\mathbf{x}$$
>
> for some scalar λ?

To answer this question, we first perform some algebraic manipulations upon the equation $\mathbf{Ax} = \lambda\mathbf{x}$. We note first that, if $\mathbf{I} = \mathbf{I_n}$ (the $n \times n$ multiplicative identity in $\mathcal{M}_{n\times n}(\mathbb{R})$), then we can write

$$
\begin{aligned}
\mathbf{Ax} = \lambda\mathbf{x} \quad &\Leftrightarrow \quad \mathbf{Ax} - \lambda\mathbf{x} = \mathbf{0} \\
&\Leftrightarrow \quad \mathbf{Ax} - \lambda\mathbf{Ix} = \mathbf{0} \\
&\Leftrightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.
\end{aligned}
$$

Remember that we are looking for nonzero $\mathbf{x}$ that satisfy this last equation. But $\mathbf{A} - \lambda\mathbf{I}$ is an $n \times n$ matrix and, should its determinant be nonzero, this last equation will have exactly one solution, namely $\mathbf{x} = \mathbf{0}$. Thus our question above has the following answer:

The equation $\mathbf{Ax} = \lambda\mathbf{x}$ has nonzero solutions for the vector x if and only if the matrix $\mathbf{A} - \lambda\mathbf{I}$ has zero determinant.

As we will see in the examples below, for a given matrix $\mathbf{A}$ there are only a few special values of the scalar λ for which $\mathbf{A} - \lambda\mathbf{I}$ will have zero determinant, and these special values are called the *eigenvalues* of the matrix $\mathbf{A}$. Based upon the answer to our question, it seems we must first be able to find the eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$ of $\mathbf{A}$ and then see about solving the individual equations $\mathbf{Ax} = \lambda_i\mathbf{x}$ for each $i = 1, \ldots, n$.

Example: Find the eigenvalues of the matrix $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix}$.

The eigenvalues are those λ for which $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Now

$$\begin{aligned}
\det(\mathbf{A} - \lambda\mathbf{I}) &= \det\left(\begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \\
&= \begin{vmatrix} 2-\lambda & 2 \\ 5 & -1-\lambda \end{vmatrix} \\
&= (2-\lambda)(-1-\lambda) - 10 \\
&= \lambda^2 - \lambda - 12.
\end{aligned}$$

The eigenvalues of $\mathbf{A}$ are the solutions of the quadratic equation $\lambda^2 - \lambda - 12 = 0$, namely $\lambda_1 = -3$ and $\lambda_2 = 4$.

As we have discussed, if $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ then the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{b}$ has either no solutions or infinitely many. When we take $\mathbf{b} = \mathbf{0}$ however, it is clear by the existence of the solution $\mathbf{x} = \mathbf{0}$ that there are infinitely many solutions (i.e., we may rule out the "no solution" case). If we continue using the matrix $\mathbf{A}$ from the example above, we can expect nonzero solutions $\mathbf{x}$ (infinitely many of them, in fact) of the equation $\mathbf{Ax} = \lambda\mathbf{x}$ precisely when $\lambda = -3$ or $\lambda = 4$. Let us procede to characterize such solutions.

First, we work with $\lambda = -3$. The equation $\mathbf{Ax} = \lambda\mathbf{x}$ becomes $\mathbf{Ax} = -3\mathbf{x}$. Writing

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and using the matrix $\mathbf{A}$ from above, we have

$$\mathbf{Ax} = \begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 \\ 5x_1 - x_2 \end{bmatrix},$$

while

$$-3\mathbf{x} = \begin{bmatrix} -3x_1 \\ -3x_2 \end{bmatrix}.$$

Setting these equal, we get

$$\begin{bmatrix} 2x_1 + 2x_2 \\ 5x_1 - x_2 \end{bmatrix} = \begin{bmatrix} -3x_1 \\ -3x_2 \end{bmatrix} \quad \begin{aligned} &\Rightarrow \quad 2x_1 + 2x_2 = -3x_1 \qquad \text{and} \qquad 5x_1 - x_2 = -3x_2 \\ &\Rightarrow \quad 5x_1 = -2x_2 \\ &\Rightarrow \quad x_1 = -\frac{2}{5}x_2. \end{aligned}$$

This means that, while there are infinitely many nonzero solutions (solution vectors) of the equation $\mathbf{Ax} = -3\mathbf{x}$, they all satisfy the condition that the first entry x_1 is $-2/5$ times the second entry x_2. Thus all solutions of this equation can be characterized by

$$\begin{bmatrix} 2t \\ -5t \end{bmatrix} = t \begin{bmatrix} 2 \\ -5 \end{bmatrix},$$

where t is any real number. The nonzero vectors $\mathbf{x}$ that satisfy $\mathbf{Ax} = -3\mathbf{x}$ are called *eigenvectors* associated with the eigenvalue $\lambda = -3$. One such eigenvector is

$$\mathbf{u_1} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

and all other eigenvectors corresponding to the eigenvalue (-3) are simply scalar multiples of $\mathbf{u_1}$ — that is, $\mathbf{u_1}$ spans this set of eigenvectors.

Similarly, we can find eigenvectors associated with the eigenvalue $\lambda = 4$ by solving $\mathbf{Ax} = 4\mathbf{x}$:

$$\begin{bmatrix} 2x_1 + 2x_2 \\ 5x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 4x_1 \\ 4x_2 \end{bmatrix} \quad \begin{aligned} &\Rightarrow \quad 2x_1 + 2x_2 = 4x_1 \qquad \text{and} \qquad 5x_1 - x_2 = 4x_2 \\ &\Rightarrow \quad x_1 = x_2. \end{aligned}$$

Hence the set of eigenvectors associated with $\lambda = 4$ is spanned by

$$\mathbf{u_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Example: Find the eigenvalues and associated eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} 7 & 0 & -3 \\ -9 & -2 & 3 \\ 18 & 0 & -8 \end{bmatrix}.$$

First we compute $\det(\mathbf{A} - \lambda\mathbf{I})$ via a cofactor expansion along the second column:

$$\begin{aligned} \begin{vmatrix} 7-\lambda & 0 & -3 \\ -9 & -2-\lambda & 3 \\ 18 & 0 & -8-\lambda \end{vmatrix} &= (-2-\lambda)(-1)^4 \begin{vmatrix} 7-\lambda & -3 \\ 18 & -8-\lambda \end{vmatrix} \\ &= -(2+\lambda)[(7-\lambda)(-8-\lambda) + 54] \\ &= -(\lambda+2)(\lambda^2 + \lambda - 2) \\ &= -(\lambda+2)^2(\lambda - 1). \end{aligned}$$

Thus $\mathbf{A}$ has two distinct eigenvalues, $\lambda_1 = -2$ and $\lambda_3 = 1$. (Note that we might say $\lambda_2 = -2$, since, as a root, -2 has multiplicity two. This is why we labelled the eigenvalue 1 as λ_3.)

Now, to find the associated eigenvectors, we solve the equation $(\mathbf{A} - \lambda_j \mathbf{I})\mathbf{x} = \mathbf{0}$ for $j = 1, 2, 3$. Using the eigenvalue $\lambda_3 = 1$, we have

$$\begin{aligned}(\mathbf{A} - \mathbf{I})\mathbf{x} &= \begin{bmatrix} 6x_1 - 3x_3 \\ -9x_1 - 3x_2 + 3x_3 \\ 18x_1 - 9x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \ x_3 = 2x_1 \qquad \text{and} \qquad x_2 = x_3 - 3x_1 \\ &\Rightarrow \ x_3 = 2x_1 \qquad \text{and} \qquad x_2 = -x_1.\end{aligned}$$

So the eigenvectors associated with $\lambda_3 = 1$ are all scalar multiples of

$$\mathbf{u_3} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}.$$

Now, to find eigenvectors associated with $\lambda_1 = -2$ we solve $(\mathbf{A} + 2\mathbf{I})\mathbf{x} = \mathbf{0}$. We have

$$\begin{aligned}(\mathbf{A} + 2\mathbf{I})\mathbf{x} &= \begin{bmatrix} 9x_1 - 3x_3 \\ -9x_1 + 3x_3 \\ 18x_1 - 6x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \ x_3 = 3x_1.\end{aligned}$$

Something different happened here in that we acquired no information about x_2. In fact, we have found that x_2 can be chosen arbitrarily, and independently of x_1 and x_3 (whereas x_3 cannot be chosen independently of x_1). This allows us to choose two linearly independent eigenvectors associated with the eigenvalue $\lambda = -2$, such as $\mathbf{u_1} = (1, 0, 3)$ and $\mathbf{u_2} = (1, 1, 3)$. It is a fact that all other eigenvectors associated with $\lambda_2 = -2$ are in the span of these two; that is, all others can be written as linear combinations $c_1\mathbf{u_1} + c_2\mathbf{u_2}$ using an appropriate choices of the constants c_1 and c_2.

Example: Find the eigenvalues and associated eigenvectors of the matrix

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

We compute

$$\begin{aligned}\det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} -1 - \lambda & 2 \\ 0 & -1 - \lambda \end{vmatrix} \\ &= (\lambda + 1)^2.\end{aligned}$$

Setting this equal to zero we get that $\lambda = -1$ is a (repeated) eigenvalue. To find any associated eigenvectors we must solve for $\mathbf{x} = (x_1, x_2)$ so that $(\mathbf{A} + \mathbf{I})\mathbf{x} = \mathbf{0}$; that is,

$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad x_2 = 0.$$

Thus, the eigenvectors corresponding to the eigenvalue $\lambda = -1$ are the vectors whose second component is zero, which means that we are talking about all scalar multiples of $\mathbf{u} = (1, 0)$.

Notice that our work above shows that there are no eigenvectors associated with $\lambda = -1$ which are linearly independent of $\mathbf{u}$. This may go against your intuition based upon the results of the example before this one, where an eigenvalue of multiplicity two had two linearly independent associated eigenvectors. Nevertheless, it is a (somewhat disparaging) fact that eigenvalues can have fewer linearly independent eigenvectors than their multiplicity suggests.

Example: Find the eigenvalues and associated eigenvectors of the matrix

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

We compute

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 2-\lambda & -1 \\ 1 & 2-\lambda \end{vmatrix} \\ &= (\lambda - 2)^2 + 1 \\ &= \lambda^2 - 4\lambda + 5. \end{aligned}$$

The roots of this polynomial are $\lambda_1 = 2+i$ and $\lambda_2 = 2-i$; that is, the eigenvalues are not real numbers. This is a common occurrence, and we can press on to find the eigenvectors just as we have in the past with real eigenvalues. To find eigenvectors associated with $\lambda_1 = 2+i$, we look for $\mathbf{x}$ satisfying

$$\begin{aligned} (\mathbf{A} - (2+i)\mathbf{I})\mathbf{x} = \mathbf{0} \quad &\Rightarrow \quad \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \quad \begin{bmatrix} -ix_1 - x_2 \\ x_1 - ix_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \quad x_1 = ix_2. \end{aligned}$$

Thus all eigenvectors associated with $\lambda_1 = 2+i$ are scalar multiples of $\mathbf{u_1} = (i, 1)$. Proceeding with $\lambda_2 = 2-i$, we have

$$\begin{aligned} (\mathbf{A} - (2-i)\mathbf{I})\mathbf{x} = \mathbf{0} \quad &\Rightarrow \quad \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \quad \begin{bmatrix} ix_1 - x_2 \\ x_1 + ix_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow \quad x_1 = -ix_2, \end{aligned}$$

which shows all eigenvectors associated with $\lambda_2 = 2 - i$ to be scalar multiples of $\mathbf{u_2} = (-i, 1)$.

Notice that $\mathbf{u_2}$, the eigenvector associated with the eigenvalue $\lambda_2 = 2 - i$ in the last example, is the complex conjugate of $\mathbf{u_1}$, the eigenvector associated with the eigenvalue $\lambda_1 = 2 + i$. It is indeed a fact that, if $\mathbf{A} \in \mathcal{M}_{n\times n}(\mathbb{R})$ has a nonreal eigenvalue $\lambda_1 = \lambda + i\mu$ with corresponding eigenvector $\xi_\mathbf{1}$, then it also has eigenvalue $\lambda_2 = \lambda - i\mu$ with corresponding eigenvector $\xi_\mathbf{2} = \bar{\xi}_\mathbf{1}$.