Contents

1. Math 83 Linear Algebra Notes: October 2, 1998

The following notes on linear algebra are taken from Dan Anderson's class notes from UNC Math 83 taught in the spring of 1998. Three main references have been used: Rabenstein[1], Roberts[2] and Strang[3]. Material taken from these references includes theorems, definitions and examples.

2. Systems of Linear Equations

Suppose we have one linear equation

$$2x_1 + x_2 - 5 = 0. \tag{2.1}$$

In terms of the x_1–x_2 plane, solutions to this equation correspond to points along a line. Suppose now that we have two linear equations

$$\begin{aligned} 2x_1 + x_2 - 5 &= 0, && (2.2a) \\ x_1 - x_2 + 1 &= 0. && (2.2b) \end{aligned}$$

We refer to this as a *system of two linear equations*. If we want to know the solution to this system of equations we can find it graphically by plotting the two lines and finding the point of intersection. In this particular case, there is a unique solution $x_1 = \frac{4}{3}$, $x_2 = \frac{7}{3}$. With this geometrical interpretation in mind, however, it is not hard to imagine that a system of two linear equations could have *no* solutions, *one* solution, or *infinitely many* solutions. Geometrically a system of two linear equations with no solutions would correspond to two parallel lines. An example is

$$\begin{aligned} 2x_1 + x_2 - 5 &= 0, && (2.3a) \\ 2x_1 + x_2 - 6 &= 0. && (2.3b) \end{aligned}$$

A system of two linear equations that has infinitely many solutions corresponds to two lines that coincide at every point. An example is

$$\begin{aligned} 2x_1 + x_2 - 5 &= 0, && (2.4a) \\ 4x_1 + 2x_2 - 10 &= 0. && (2.4b) \end{aligned}$$

This geometrical view of systems of linear equations can be extended to systems with three equations and three unknowns. The equation

$$x_1 + x_2 + x_3 = 5, \tag{2.5}$$

has a solution which corresponds to a plane in three dimensional space. If we have an additional equation we may have, in geometrical terms, two planes intersecting at a line, two planes that are coincident, or two parallel planes which do not intersect. One can also imagine three planes intersecting a single point, a line, not at all, etc.

A general linear system with m equations and n unknowns can be expressed as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ &\vdots\ , \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m, \end{aligned} \tag{2.6}$$

where a_{ij} and b_i are given constants and x_j are the unknowns. This linear system has either zero, one, or infinitely many solutions.

In contrast to linear equations, *nonlinear equations* involve nonlinear combinations of the unknowns. Examples of nonlinear equations are

$$x_1 + x_2^2 = 4, \tag{2.7a}$$

$$x_1x_2 + x_2^3 = 7, \tag{2.7b}$$

$$x_1 + 4\sin x_2 = 0. \tag{2.7c}$$

The first is nonlinear because of the second term on the left-hand side. The second is nonlinear because of either term on the left-hand side. The third is nonlinear because of the second term on the left-hand side. Our main focus here will be on linear equations.

SOLVING LINEAR SYSTEMS: We now outline a systematic approach to finding solutions to linear equations. This approach applies in principal to systems of linear equations of any size. Realistically, it is viable approach only for small systems. Large systems, in practice, are much more efficiently treated with a variety of computational methods.

Suppose we have the following linear system

$$x_1 + 3x_2 - x_3 = 1, \tag{2.8a}$$

$$3x_1 + 6x_2 + 2x_3 = -3, \tag{2.8b}$$

$$3x_1 + 4x_2 - 4x_3 = 7. \tag{2.8c}$$

We wish to have a systematic way to solve this and similar linear systems. In order to do so we define the following *Elementary Operations*.

1. Interchange any two equations.
2. Multiply any equation by a nonzero constant.
3. Replace any equation with the sum of that equation and a multiple of another equation.

Applying these operations to the system of linear equations will not change the *solution* of the system. The idea is to simplify the system of equations, systematically by using the elementary operations, to a point where the solution is easily identified. We outline two such procedures below.

The first method is known as *Row Echelon Reduction* or the *Row Echelon Method*. The general idea of this approach is to eliminate the variables x_1, x_2 and x_3 in as many places as possible. Our first step is to multiply Eq. (2.8a) by -3 and add the resulting equation to Eq. (2.8b). Replace Eq. (2.8b) by the result. Here we have used the third elementary operation. The next step is to multiply Eq. (2.8a) by -3 and add the resulting equation to Eq. (2.8c). Replace Eq. (2.8c) by the result. This gives

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 1, && && (2.9a)\\ -3x_2 + 5x_3 &= -6, && [-3 \times \mathrm{ROW1} + \mathrm{ROW2}] && (2.9b)\\ -5x_2 - x_3 &= 4. && [-3 \times \mathrm{ROW1} + \mathrm{ROW3}] && (2.9c) \end{aligned}$$

In the brackets above we write the steps used in simplying the system. By performing these steps, we have eliminated the variable x_1 in all but the first equation. Our next task is to try to eliminate the variable x_2 in all but the second equation, using the elementary operations. To do this, we replace Eq. (2.9a) by the sum of Eqs. (2.9a) and (2.9b). We then replace Eq. (2.9c) by $-5/3$ times Eq. (2.9b) plus Eq. (2.9c). These operations results in the system

$$\begin{aligned} x_1 \qquad + 4x_3 &= -5, && [\mathrm{ROW2} + \mathrm{ROW1}] && (2.10a)\\ -3x_2 + 5x_3 &= -6, && && (2.10b)\\ -\frac{28}{3}x_3 &= 14 && [-5/3 \times \mathrm{ROW2} + \mathrm{ROW3}]. && (2.10c) \end{aligned}$$

We have now eliminated the variable x_2 in all but the second equation. Before attempting to eliminate the variable x_3 in all but the third equation, we simplify the system by multiplying Eq. (2.10b) by $-1/3$ and Eq. (2.10c) by $-3/28$. This gives

$$\begin{aligned} x_1 \qquad + 4x_3 &= -5, && && (2.11a)\\ x_2 - (5/3)x_3 &= 2, && [-1/3 \times \mathrm{ROW2}] && (2.11b)\\ x_3 &= -3/2. && [-3/28 \times \mathrm{ROW3}] && (2.11c) \end{aligned}$$

We now eliminate the variable x_3 everywhere but in Eq. (2.11c). We replace Eq. (2.11a) by -4 times Eq. (2.11c) plus Eq. (2.11a). We replace Eq. (2.11b) by $5/3$ times Eq. (2.11c) plus Eq. (2.11b). The result is the system

$$\begin{aligned} x_1 &= 1, && [-4 \times \mathrm{ROW3} + \mathrm{ROW1}] && (2.12a)\\ x_2 &= -1/2, && [5/3 \times \mathrm{ROW3} + \mathrm{ROW2}] && (2.12b)\\ x_3 &= -3/2. && && (2.12c) \end{aligned}$$

The solution of the original system of equations is therefore just $(x_1, x_2, x_3) = (1, -\frac{1}{2}, -\frac{3}{2})$, which can be checked by substituting these values back into the original equation. This approach is the basic idea of the *Row Echelon Method*.

A similar approach called *Gaussian Elimination* uses the same Elementary Operations but simplifies the original system of equations given by Eq. (2.8) to the form

$$
\begin{aligned}
x_1 + 3x_2 - x_3 &= 1, & \quad (2.13a)\\
-\,3x_2 + 5x_3 &= -6, & \quad (2.13b)\\
-\frac{28}{3}x_3 &= 14. & \quad (2.13c)
\end{aligned}
$$

That is, one equation involves just x_3, another involves just x_2 and x_3, and the third equation involves all three variables. The solution is then obtained by first solving for x_3 using equation (2.13c), then substituting it into equation (2.13b) and solving for x_2 and then substituting the results for both x_3 and x_2 into equation (2.13a) and solving for x_1; a method referred to as *back substitution*.

3. Homogeneous Systems

The systems considered in the previous section were nonhomogeneous (the system has at least one nonzero constant term). That is, if we examine the system given by Eq. (2.6), the system is *nonhomogeneous* if any one (or more) of the terms b_i is nonzero. A *homogeneous* system is one in which *all* of the constant terms b_i are equal to zero. An example of a homogeneous system is the system

$$
\begin{aligned}
x_1 + 3x_2 - x_3 &= 0, & \quad (3.1a)\\
3x_1 + 6x_2 + 2x_3 &= 0, & \quad (3.1b)\\
3x_1 + 4x_2 - 4x_3 &= 0. & \quad (3.1c)
\end{aligned}
$$

We note that one solution to this equation is $x_1 = x_2 = x_3 = 0$, which is referred to as the *trivial solution* (a solution in which all of the unknowns are identically zero). In fact, the following is always true:

A homogeneous system always has the trivial solution.

On the other hand, a homogeneous system may or may not have nontrivial solutions (a *nontrivial solution* is one in which the unknowns are not all zero). We will return to this issue later, but for now we demonstrate this by considering specific examples.

In the case of the system (3.1), we can apply elementary operations to simplify it to

$$
\begin{aligned}
x_1 + 3x_2 - x_3 &= 0, & \quad (3.2a)\\
-3x_2 + 5x_3 &= 0, & \quad (3.2b)\\
x_3 &= 0. & \quad (3.2c)
\end{aligned}
$$

This system is solved only for $x_1 = x_2 = x_3 = 0$. Therefore, this particular example has only the trivial solution.

Let us consider the following example.

$$x_1 + 5x_2 + x_3 = 0, \tag{3.3a}$$

$$2x_1 + x_2 - x_3 = 0. \tag{3.3b}$$

This is a homogeneous system and therefore has as one solution the trivial solution. Notice that this system has three unknowns and two equations. We can apply elementary operations to this system to obtain

$$x_1 + 5x_2 + x_3 = 0, \tag{3.4a}$$

$$-9x_2 - 3x_3 = 0. \tag{3.4b}$$

The second equation is solved whenever $x_2 = -(1/3)x_3$ and the first is then solved whenever $x_1 = (2/3)x_3$. Therefore, if we choose x_3 to be any constant, a, a solution of the following form can be found $(x_1, x_2, x_3) = (\frac{2}{3}a, -\frac{1}{3}a, a)$. In particular we note that both zero and nonzero values of a are allowed. Therefore, the system (3.3) has both trivial and nontrivial solutions. In fact, since a can be *any* number, there are an infinite number of nontrivial solutions.

THEOREM **3.1** *(Rabenstein[1], Thm. 2.1) A system of linear homogeneous equations in which the number of variables (unknowns) exceeds the number of equations always has an infinite number of nontrivial solutions.*

EXAMPLE **3.1** *(Rabenstein[1], section 2.1, problem 5) Find the solution(s) to the following system using elementary operations.*

$$3x_1 - x_2 = 4, \tag{3.5a}$$

$$6x_1 - 2x_2 = 8, \tag{3.5b}$$

$$-9x_1 + 3x_2 = -12. \tag{3.5c}$$

Note that geometrically these equations correspond to three coincident lines. Using elementary operations we obtain

$$3x_1 - x_2 = 4, \tag{3.6a}$$

$$0 = 0, \quad [-2 \times \text{ROW1} + \text{ROW2}] \tag{3.6b}$$

$$0 = 0. \quad [3 \times \text{ROW1} + \text{ROW3}] \tag{3.6c}$$

The solutions are given by $(x_1, x_2) = (\frac{1}{3}(4 + a), a)$ where a is a constant. Another way to express the solution is just

$$x_1 = \frac{4}{3} + \frac{1}{3}x_2. \tag{3.7}$$

There are an infinite number of nontrivial solutions to this system of equations.

4. Matrices and Vectors

Writing many equations with subscripts, etc. is tedious. One way to simplify things is to use matrix and vector notation. Consider the system of linear equations

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 + 2x_4 &= 3, && (4.1a)\\ 3x_1 - 5x_2 + x_3 + 3x_4 &= -2, && (4.1b)\\ 4x_1 - 3x_2 + x_3 + 2x_4 &= 2. && (4.1c) \end{aligned}$$

We define the <u>*coefficient matrix*</u> (or just matrix) $\boldsymbol{A}$ for the above system by

$$\boldsymbol{A} = \begin{bmatrix} 2 & -3 & 2 & 2 \\ 3 & -5 & 1 & 3 \\ 4 & -3 & 1 & 2 \end{bmatrix} \tag{4.2}$$

This is a 3×4 matrix; it has 3 rows and 4 columns. One can think of the matrix as being made up of <u>*column vectors*</u>

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \tag{4.3}$$

which by themselves can also be thought of as 3×1 matrices, or as made up of <u>*row vectors*</u>

$$\begin{gathered} \begin{bmatrix} 2 & -3 & 2 & 2 \end{bmatrix}, \\ \begin{bmatrix} 3 & -5 & 1 & 3 \end{bmatrix}, \\ \begin{bmatrix} 4 & -3 & 1 & 2 \end{bmatrix}, \end{gathered} \tag{4.4}$$

which can also be thought of as 1×4 matrices. We define the <u>*vectors*</u>

$$\vec{b} = \begin{bmatrix} 3 \\ -2 \\ 2 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \tag{4.5}$$

where we refer to $\vec{b}$ as a three-dimensional vector and $\vec{x}$ as a four-dimensional vector and write the above system (4.1) in <u>*matrix–vector notation*</u> as

$$\boldsymbol{A}\vec{x} = \vec{b}. \tag{4.6}$$

In order to do this, we need to understand how to muliply a matrix and a vector. That is, we need to know what $\boldsymbol{A}\vec{x}$ means. We have

$$\boldsymbol{A}\vec{x} = \begin{bmatrix} 2 & -3 & 2 & 2 \\ 3 & -5 & 1 & 3 \\ 4 & -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2x_1 - 3x_2 + 2x_3 + 2x_4 \\ 3x_1 - 5x_2 + x_3 + 3x_4 \\ 4x_1 - 3x_2 + x_3 + 2x_4 \end{bmatrix}. \tag{4.7}$$

Notice that $\boldsymbol{A}$ is a 3×4 matrix, $\vec{x}$ is 4×1 and the product $\boldsymbol{A}\vec{x}$ is 3×1. The idea of matrix–vector multiplication is that one starts with the first row of the matrix, takes the sum of the products of the terms in that row and the corresponding terms in the vector $\vec{x}$; the result is the first entry in the resulting vector. One performs these steps for each row of the matrix. When one equates the final result in Eq. (4.7) with the right-hand side of the original equation $\vec{b}$ one gets the original system back. Another way to think of this multiplication is as

$$\boldsymbol{A}\vec{x} = \begin{bmatrix} 2 & -3 & 2 & 2 \\ 3 & -5 & 1 & 3 \\ 4 & -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ -5 \\ -3 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}. \tag{4.8}$$

To solve a system in matrix–vector notation one applies the Elementary Operations to the rows of the <u>*augmented matrix*</u>

$$\left[\begin{array}{cccc|c} 2 & -3 & 2 & 2 & 3 \\ 3 & -5 & 1 & 3 & -2 \\ 4 & -3 & 1 & 2 & 2 \end{array}\right], \tag{4.9}$$

where the last column in this matrix represents the right-hand side of Eq. (4.1). The same rules of Gaussian Elimination or Row Echelon reduction apply, only the notation is simplified somewhat.

MATRIX–VECTOR MULTIPLICATION: Not all matrices and vectors can be multiplied together. The dimension of each must be appropriate. Examples are given below that make use of the following matrices and vectors.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix},$$
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{4.10}$$

Based on the above matrices and vectors the following matrix–vector products

$$\boldsymbol{A}\vec{y}, \quad \boldsymbol{B}\vec{x}, \quad \boldsymbol{C}\vec{x}, \quad \text{are all defined,} \tag{4.11}$$

while the following matrix–vector products

$$\boldsymbol{A}\vec{x}, \quad \boldsymbol{B}\vec{y}, \quad \boldsymbol{C}\vec{y}, \quad \text{are not defined.} \tag{4.12}$$

MATRIX–MATRIX MULTIPLICATION: Matrices can also be multiplied together if their dimensions are appropriate. For example, using the matrices defined in Eq. (4.10), we have that

$$\begin{aligned}
\boldsymbol{AC} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \\
&= \begin{bmatrix} (a_{11}c_{11} + a_{12}c_{21}) & (a_{11}c_{12} + a_{12}c_{22}) & (a_{11}c_{13} + a_{12}c_{23}) \\ (a_{21}c_{11} + a_{22}c_{21}) & (a_{21}c_{12} + a_{22}c_{22}) & (a_{21}c_{13} + a_{22}c_{23}) \end{bmatrix}, \qquad (4.13)
\end{aligned}$$

so that the matrix product $\boldsymbol{AC}$ is a 2×3 matrix. In a similar fashion, one can define the matrix product $\boldsymbol{CB}$. We note, however, that $\boldsymbol{CA}$, $\boldsymbol{AB}$, $\boldsymbol{BA}$, and $\boldsymbol{BC}$ are not defined. In general, matrix–matrix multiplication is not commutative. In fact, as we see in the above examples, $\boldsymbol{AC}$ can be defined while $\boldsymbol{CA}$ is not!

EXAMPLE 4.1 *Find the following matrix–matrix product.*

$$\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 4(1) + 2(-2) & 4(-1) + 2(2) \\ 1(1) + 3(-2) & 1(-1) + 3(2) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -5 & 5 \end{bmatrix} \qquad (4.14)$$

SOME MATRIX PROPERTIES: Suppose we are given the following 2×2 matrices

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}. \qquad (4.15)$$

The two matrices are equal to each other, $\boldsymbol{A} = \boldsymbol{B}$, if $a_{11} = b_{11}$, $a_{12} = b_{12}$, etc.

We can add and subtract matrices of the same size

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}. \qquad (4.16)$$

We can multiply any matrix by a scalar r

$$r\boldsymbol{A} = \begin{bmatrix} ra_{11} & ra_{12} \\ ra_{21} & ra_{22} \end{bmatrix}. \qquad (4.17)$$

Also, as noted previously, matrix multiplication is not commutative in general

$$\boldsymbol{AB} \neq \boldsymbol{BA}, \qquad (4.18)$$

but it is associative

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}), \qquad (4.19)$$

as long as each matrix multiplication is defined.

SPECIAL MATRICES: The <u>*identity matrix*</u> is a square matrix (that is, $n \times n$) given by (for the 3×3 case)

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (4.20)$$

In general, it has the entry 1 along its diagonal and 0 everywhere else. The identity matrix has the properties that

$$\boldsymbol{A}\boldsymbol{I} = \boldsymbol{I}\boldsymbol{A} = \boldsymbol{A}, \tag{4.21a}$$

$$\vec{b}\boldsymbol{I} = \boldsymbol{I}\vec{b} = \vec{b}. \tag{4.21b}$$

The matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \tag{4.22}$$

is *symmetric* if $a_{12} = a_{21}$, $a_{13} = a_{31}$ and $a_{23} = a_{32}$.

The *transpose of a matrix* (referring to the above matrix $\boldsymbol{A}$) is

$$\boldsymbol{A}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}. \tag{4.23}$$

That is, it has its elements "flipped" about the diagonal. We also have that $(\boldsymbol{A}^T)^T = \boldsymbol{A}$, $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$ and $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$. Note that if $\boldsymbol{A}$ is symmetric then $\boldsymbol{A}^T = \boldsymbol{A}$.

A *diagonal matrix* has all of its off-diagonal entries equal to zero, e.g.

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}. \tag{4.24}$$

INVERSE OF A MATRIX: Another special matrix is the inverse matrix. Suppose we have the linear system

$$\boldsymbol{A}\vec{x} = \vec{b}, \tag{4.25}$$

where $\boldsymbol{A}$ is an $n \times n$ matrix and $\vec{x}$ and $\vec{b}$ are n-dimensional vectors. If $n = 1$, so the system is just the single equation, then what we really have is

$$ax = b. \tag{4.26}$$

In this case, the way we solve for x is to divide both sides by a, or equivalently multiply both sides by $1/a$. Then the solution is $x = b/a$. This works as long as $a \neq 0$. *Is there an analogous operation for a general $n \times n$ matrix $\boldsymbol{A}$?* The answer is sometimes. For those cases when the answer is yes, it involves the inverse of the matrix.

We say that a matrix $\boldsymbol{A}$ is *invertible* if there exists a matrix $\boldsymbol{B}$ such that $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$ and $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{I}$. There is at most one such, $\boldsymbol{B}$, called the *inverse of* $\boldsymbol{A}$ and is denoted by $\boldsymbol{A}^{-1}$ (note that if $\boldsymbol{A}$ is $n \times n$ then so is $\boldsymbol{A}^{-1}$). So if $\boldsymbol{A}$ is invertible, then

$$\begin{aligned} \boldsymbol{A}^{-1}(\boldsymbol{A}\vec{x}) &= \boldsymbol{A}^{-1}\vec{b}, \\ (\boldsymbol{A}^{-1}\boldsymbol{A})\vec{x} &= \boldsymbol{A}^{-1}\vec{b}, \\ \boldsymbol{I}\vec{x} &= \boldsymbol{A}^{-1}\vec{b}, \\ \vec{x} &= \boldsymbol{A}^{-1}\vec{b}, \end{aligned} \tag{4.27}$$

where the bottom line represents the solution to our system of linear equations. That is, $\vec{x}$ can be obtained by a matrix–vector multiplication of $\boldsymbol{A}^{-1}$ and $\vec{b}$. For future reference we say that $\boldsymbol{A}$ is <u>*invertible*</u> iff it is <u>*nonsingular*</u> and $\boldsymbol{A}$ is <u>*not invertible*</u> iff it is <u>*singular*</u>.

The next question we ask is that *if* $\boldsymbol{A}$ is invertible, how do we find $\boldsymbol{A}^{-1}$? To address the above question, suppose we have the 2×2 matrix

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}. \tag{4.28}$$

The inverse has the property that $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$. Think of $\boldsymbol{A}^{-1}$ as made up of column vectors

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \uparrow & \uparrow \\ \vec{c}_1 & \vec{c}_2 \\ \downarrow & \downarrow \end{bmatrix}. \tag{4.29}$$

Now imagine the equation $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$ as

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow \\ \vec{c}_1 & \vec{c}_2 \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.30}$$

One column at a time this is

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \vec{c}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{4.31a}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \vec{c}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{4.31b}$$

where we would like to solve for $\vec{c}_1$ and $\vec{c}_2$. We can do so by using the following technique. First form the extended matrix

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]. \tag{4.32}$$

Next, use elementary operations on the rows of this matrix to reduce the left two columns to be the identity matrix, if possible. That is, form the matrix

$$\left[\begin{array}{cc|cc} 1 & 0 & i_{11} & i_{12} \\ 0 & 1 & i_{21} & i_{22} \end{array}\right], \tag{4.33}$$

where $i_{11}, i_{12}, i_{21}, i_{22}$ are determined as the elementary row operatrions are performed. It turns out, then, that

$$\boldsymbol{A}^{-1} = \begin{bmatrix} i_{11} & i_{12} \\ i_{21} & i_{22} \end{bmatrix} \tag{4.34}$$

So for the above example, we have

$$
\begin{aligned}
&\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right], \\
&\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right], \quad [-3 \times \text{ROW1} + \text{ROW2}] \\
&\left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right], \quad [\text{ROW2} + \text{ROW1}] \\
&\left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right], \quad [-\tfrac{1}{2} \times \text{ROW2}]
\end{aligned}
\tag{4.35}
$$

The claim is now that

$$
\boldsymbol{A}^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}. \tag{4.36}
$$

We can check this result by noting that

$$
\boldsymbol{A}\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}. \tag{4.37}
$$

This method works, at least in principle, for any invertible $n \times n$ matrix. A general result for the inverse of a 2×2 matrix is given as follows:

THEOREM **4.1**

$$
\textit{If} \quad \boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \textit{then} \quad \boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},
$$

as long as $ad - bc \neq 0$.

The number $ad - bc$ is actually the *determinant* of this 2×2 matrix. As we shall see in the next section, the value of the determinant tells us whether or not a general $n \times n$ matrix has an inverse. If we now go back to the simplest case in which we are solving the equation $ax = b$, we notice that the matrix analog of $a = 0$ (the inverse $1/a$ is not defined) is that the determinant of the matrix $\boldsymbol{A}$ is equal to zero (the inverse matrix $\boldsymbol{A}^{-1}$ does not exist).

EXAMPLE **4.2** *Find the inverse of the matrix*

$$
\begin{bmatrix} 1 & 0 & 2 \\ 2 & -3 & 4 \\ 0 & 2 & 1 \end{bmatrix}. \tag{4.38}
$$

Start by forming the extended matrix

$$
\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -3 & 4 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right], \tag{4.39}
$$

and simplify using elementary (row) operations,

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right], \quad [-2 \times \text{ROW1} + \text{ROW2}]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -4/3 & 2/3 & 1 \end{array}\right], \quad [2/3 \times \text{ROW2} + \text{ROW3}]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 11/3 & -4/3 & -2 \\ 0 & 1 & 0 & 2/3 & -1/3 & 0 \\ 0 & 0 & 1 & -4/3 & 2/3 & 1 \end{array}\right], \quad \begin{array}{c}[-2 \times \text{ROW3} + \text{ROW1}] \\ [-1/3 \times \text{ROW2}]\end{array}$$

so

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 11/3 & -4/3 & -2 \\ 2/3 & -1/3 & 0 \\ -4/3 & 2/3 & 1 \end{bmatrix}. \tag{4.40}$$

This result can be checked by confirming that $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

5. Determinants

The *determinant* is a property of an $n \times n$ matrix $\boldsymbol{A}$. It is denoted by $\det\boldsymbol{A}$. Before we describe *how* $\det\boldsymbol{A}$ can be calculated, we will state *why* it is worth calculating!

THEOREM **5.1** *(Rabenstein[1], Thm. 2.10) If* $\det\boldsymbol{A} \neq 0$*, then the matrix* $\boldsymbol{A}$ *has an inverse; that is,* $\boldsymbol{A}$ *is nonsingular. Furthermore, the* $n \times n$ *system of equations* $\boldsymbol{A}\vec{x} = \vec{b}$ *has a unique solution.*

THEOREM **5.2** *(Rabenstein[1], Thm. 2.12) The homogeneous linear system of equations* $\boldsymbol{A}\vec{x} = \vec{0}$ *has nontrivial solutions if and only if* $\det\boldsymbol{A} = 0$*. When* $\det\boldsymbol{A} = 0$*,* $\boldsymbol{A}$ *is singular, and the matrix* $\boldsymbol{A}$ *does not have an inverse.*

Determinants play key roles in numerous situations that arise in the context of linear algebra and differential equations. As can be seen from the above two theorems, the determinant of a matrix provides information about a system of linear equations before one actually tries to solve the system. We now focus on how to calculate $\det\boldsymbol{A}$.

We will start by giving an operational definition of the determinant of 2×2 and 3×3 systems before stating a general method for finding $\det\boldsymbol{A}$ for $n \times n$ systems.

• If $\boldsymbol{A}$ is a 2×2 matrix

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{5.1}$$

then

$$\det\boldsymbol{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{5.2}$$

- If $\boldsymbol{A}$ is a 3×3 matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \tag{5.3}$$

then

$$\begin{aligned} \det\boldsymbol{A} &= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}). \end{aligned} \tag{5.4}$$

Here M_{ij} is called the <u>*minor*</u> of element a_{ij}. It is the determinant of the $(n-1) \times (n-1)$ submatrix obtained from $\boldsymbol{A}$ by deleting the i^{th} row and the j^{th} column (i.e. the row and column occupied by the element a_{ij}). In the above formula we chose to expand in minors by going across the top row. The above formula is equivalent to that obtained by expanding down any column or across any row. Examples include

$$\det\boldsymbol{A} = a_{11}M_{11} - a_{21}M_{21} + a_{31}M_{31}, \tag{5.5a}$$

$$\det\boldsymbol{A} = -a_{21}M_{21} + a_{22}M_{22} - a_{23}M_{23}. \tag{5.5b}$$

Hopefully now you will have noticed that there are in some cases plus signs and other cases minus signs in these formulas. We can see where these come from by considering the general $n \times n$ case.

- Consider the general case where $\boldsymbol{A}$ is an $n \times n$ matrix,

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}. \tag{5.6}$$

Next imagine that each entry has associated with it a plus or minus sign in the following way

$$\begin{bmatrix} + & - & + & \dots & (-1)^{n+1} \\ - & + & - & \dots & -(-1)^{n+1} \\ + & - & + & \dots & (-1)^{n+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (-1)^{n+1} & -(-1)^{n+1} & (-1)^{n+1} & \dots & + \end{bmatrix}, \tag{5.7}$$

which is essentially a checkerboard pattern starting with a plus sign in the upper left corner. Now, to write down the determinant for this general case, we again start by choosing any row or any column (as we'll see, it is often advantageous to choose a row or column with lots or zeros in it, if such a row or column is present). If we choose the first row, say, the determinant is given by

$$\det \boldsymbol{A} = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} + \ldots + (-1)^{n+1}a_{1n}M_{1n}, \tag{5.8}$$

where M_{11} through M_{1n} are minors (e.g. M_{11} is the determinant of the submatrix formed by deleting row 1 and column 1 from the original matrix, while M_{1n} is the determinant of the submatrix formed by deleting row 1 and column n of the original matrix). This is a *recursive* formula. That is, the M_{ij} that appear in this formula are themselves determinants of smaller matrices, which can be expressed by formulas similar to equation (5.8). The following definition and theorem state this result.

DEFINITION **5.1** *The cofactor of element a_{ij} is $(-1)^{i+j}M_{ij}$, where M_{ij} is the determinant of the submatrix formed by deleting the row and column occupied by the element a_{ij}.*

THEOREM **5.3** *(Rabenstein[1], Thm 2.8) If $\boldsymbol{A}$ is an $n \times n$ matrix, det$\boldsymbol{A}$ may be found by choosing any row or column vector of $\boldsymbol{A}$, multiplying each component by its cofactor and adding these products.*

It is worth pointing out that this method in principal allows one to hand calculate the determinant of any $n \times n$ matrix. In reality, however, most matrices larger than 3×3, depending on the exact nature of the matrix, should probably be calculated using software such as Matlab or Mathematica. Huge matrices, which appear frequently in applications, almost always require computational tools.

EXAMPLE **5.1** *(Rabenstein[1] section 2.10, problem 2a) Find det$\boldsymbol{A}$ of*

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 & 2 \\ 1 & 1 & -3 \end{bmatrix}. \tag{5.9}$$

Here we choose the first row to expand with cofactors.

$$\begin{aligned} \det \boldsymbol{A} &= (1)\begin{vmatrix} 1 & 2 \\ 1 & -3 \end{vmatrix} - (2)\begin{vmatrix} 4 & 2 \\ 1 & -3 \end{vmatrix} + (-1)\begin{vmatrix} 4 & 1 \\ 1 & 1 \end{vmatrix} \\ &= -5 - 2(-14) - (3) = 20. \end{aligned} \tag{5.10}$$

Since the determinant is nonzero, the matrix $\boldsymbol{A}$ is nonsingular and has an inverse.

EXAMPLE 5.2 *Find det$\boldsymbol{A}$ of*

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & -1 \\ 4 & 0 & 2 \\ 1 & 1 & -3 \end{bmatrix}. \tag{5.11}$$

Here we choose the second column since it has two zeros.

$$\det\boldsymbol{A} = -(1)\begin{vmatrix} 1 & -1 \\ 4 & 2 \end{vmatrix} = -6. \tag{5.12}$$

Again since the determinant is nonzero, the matrix $\boldsymbol{A}$ is nonsingular and has an inverse.

EXAMPLE 5.3 *Find det$\boldsymbol{A}$ of*

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 4 & 1 & 2 \\ 1 & 1 & 1 & -3 \end{bmatrix}. \tag{5.13}$$

Here we choose the first row since it has two zeros (the first column also would be a good choice).

$$\begin{aligned} \det\boldsymbol{A} &= (1)\begin{vmatrix} 1 & 2 & -1 \\ 4 & 1 & 2 \\ 1 & 1 & -3 \end{vmatrix} + (1)\begin{vmatrix} 0 & 1 & -1 \\ 0 & 4 & 2 \\ 1 & 1 & -3 \end{vmatrix} \\ &= 20 + (1)\begin{vmatrix} 1 & -1 \\ 4 & 2 \end{vmatrix} = 20 + 6 = 26. \end{aligned} \tag{5.14}$$

The first 3×3 determinant appearing here we recognize from example 5.1.

SOME PROPERTIES OF DETERMINANTS:

• If $\boldsymbol{A}$ is diagonal then $\det\boldsymbol{A} = a_{11}a_{22}a_{33}\cdots a_{nn}$ (i.e. the determinant is the product of the diagonal elements).

• If $\boldsymbol{A}$ has a row (or column) with all zero entries, then $\det\boldsymbol{A} = 0$.

• If two rows (or two columns) of $\boldsymbol{A}$ are interchanged, then the determinant of the resulting matrix is $-\det\boldsymbol{A}$.

• If two rows (or two columns) of $\boldsymbol{A}$ are identical, then $\det\boldsymbol{A} = 0$.

As an exercise, write down some 2×2, 3×3 and 4×4 matrices and convince yourself of the above properties.

6. Vector Spaces

We have looked at linear systems, how to solve them, how to write them using matrices and vectors, and we have identified some properties of the systems by computing determinants.

In the following section we shall describe the "space" in which the solutions to these problems reside.

Two of the most familiar "spaces" are $\Re^2$ and $\Re^3$. One can think of $\Re^2$ mathematically as the set of ordered pairs (x, y), or geometrically as the space defined by the $x-y$ plane. Here we can define vectors $\vec{x}_1 = (x_1, y_1)$ and $\vec{x}_2 = (x_2, y_2)$ which "reside" in $\Re^2$. Two important points to notice about $\Re^2$ are that if $\vec{x}_1$ and $\vec{x}_2$ are elements in this space, then

- their sum $\vec{x}_1 + \vec{x}_2$ is also a member of this space (*Closure Under Addition*), and
- any scalar multiple $r\vec{x}_1$ is also a member of this space (*Closure Under Multiplication*).

This is saying that if two vectors in $\Re^2$ are added together, or if a vector in $\Re^2$ is multiplied by a scalar, the result is still a vector in $\Re^2$.

The situation is similar if we consider the three-dimensional space $\Re^3$. Here we have vectors $\vec{x}_1 = (x_1, y_1, z_1)$, $\vec{x}_2 = (x_2, y_2, z_2)$, etc. Again, $\vec{x}_1 + \vec{x}_2$ is in $\Re^3$ and $r\vec{x}_1$ is in $\Re^3$. The same arguments extend to vectors in $\Re^n$, although the familiar geometrical view begins to break down when $n > 3$.

$\Re^1$, $\Re^2$, $\Re^3$, ..., $\Re^n$ are examples of *vector spaces*. As we shall see below in the following formal definition, a vector space is defined not so much by the nature of its elements, but rather by the nature of the *operations applied to its elements.*

DEFINITION **6.1** *Vector Space (from Roberts[2]) Let V be a nonempty collection of elements that can be added together or multiplied by scalars. V is a vector space (or linear space) if for arbitrary elements $\vec{u}$, $\vec{v}$, $\vec{w}$ in V and arbitrary scalars a, b, c the following properties are satisfied:*

1. *$\vec{u} + \vec{v}$ is in V (closure under addition)*
2. *$a\vec{u}$ is in V (closure under multiplication)*
3. *$\vec{u} + \vec{v} = \vec{v} + \vec{u}$ (commutative property)*
4. *$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (associative property)*
5. *There is a zero element $\vec{0}$ in V such that $\vec{u} + \vec{0} = \vec{u}$*
6. *For every $\vec{u}$ in V there is another element $-\vec{u}$ in V such that $\vec{u} + (-\vec{u}) = \vec{0}$*
7. *$c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$*
8. *$(a + b)\vec{u} = a\vec{u} + b\vec{u}$*
9. *$a(b\vec{u}) = (ab)\vec{u}$*
10. *$1\vec{u} = \vec{u}$.*

If you go through this list with $\Re^2$, $\Re^3$ or in general $\Re^n$ in mind, for example, you should be able to convince yourself that each of the ten properties hold.

The notion of a vector space is not limited to sets whose elements are *geometrical vectors* (in the usual way of thinking of vectors). The following examples show some other collections of elements which are vector spaces and some collections of elements which are not.

EXAMPLE **6.1** *The set of all polynomials forms a vector space. Examples of elements in the set of all polynomials are*

$$u(x) = 7 - 3x + x^3,$$

$$w(x) = 5x - x^2 + 2x^3 + x^5.$$

If we add any two polynomials together we get another polynomial (so point 1 is satisfied). If we multiply any polynomial by a scalar we get a polynomial (so point 2 is satisfied). The commutative and associative properties hold for polynomials (points 3 and 4). There is a zero element, namely 0 (point 5). Taking the negative of a polynomial is a polynomial and adding these together gives 0 (point 6). Properties 7 through 10 can also be confirmed for polynomials.

EXAMPLE **6.2** *The set of all real-valued functions that are defined and have a continuous derivative on an interval (a, b) forms a vector space. Here we think of the elements $\vec{u}$, $\vec{v}$ and $\vec{w}$ as representing the functions $f(x)$, $g(x)$ and $h(x)$, for example. Again, a mental check confirms the above 10 conditions.*

EXAMPLE **6.3** *The set of odd integers does not form a vector space. To see this note that* 1 *and* 3 *are both odd integers, but their sum* $1 + 3 = 4$ *is not an odd integer. This set is not closed under addition and is therefore not a vector space.*

Why do we worry about vector spaces? They can be used to describe the space where solutions to linear systems (algebraic or differential systems) reside. In the case of algebraic systems, if we know that the solution to a problem resides in $\Re^2$ (the solutions are vectors in a plane) then we know not to look for solutions in $\Re^3$ (which include vectors that can point out of the above plane). In the case of differential equations, knowing the "size" or *dimension* of a vector space essentially tells us when we can stop looking for solutions (i.e. when we have found all that exist). Often solving differential equations involves making good guesses about the solution so it is advantageous to have all the clues we can get.

7. Subspaces

Basically, subspaces are vector spaces, only smaller.

DEFINITION **7.1** *(Roberts[2]) A nonempty subcollection U of elements in a vector space V is called a subspace of V if the elements of U are themselves a vector space.*

While this is a formal definition of a subspace, it turns out that all one really needs to show is *closure under addition* and *closure under multiplication* for the elements in U. Note that it *is required* that the elements already be elements of a known vector space V. If $\vec{u}$, $\vec{v}$ and $\vec{w}$ are in U, then properties 3,4,7-10 in the definition of vector space hold automatically (since $\vec{u}$, $\vec{v}$ and $\vec{w}$ are also in V). Further, once the properties of closure under addition and multiplication are shown, properties 5 and 6 in the definition of vector space follow. We can state this as a theorem.

THEOREM **7.1** *(Rabenstein[1], Thm. 3.1) If U is a nonempty subcollection of elements of a vector space V, then U is a* <u>*subspace of V*</u> *if and only if U is closed under addition and multiplication. That is,*

- *if $\vec{u}_1$ and $\vec{u}_2$ are in U, then so is $\vec{u}_1 + \vec{u}_2$, and*
- *if $\vec{u}$ is in U, then so is $r\vec{u}$.*

EXAMPLE **7.1** *Show that the set of polynomials of degree 2 or less forms a subspace of the vector space of all polynomials.* Note that the set of all polynomials is a vector space and the set of polynomials of degree 2 or less are included in that vector space.

• Closure under addition: We first identify two general elements in the subcollection as

$$u(x) = a_2x^2 + a_1x + a_0, \qquad v(x) = b_2x^2 + b_1x + b_0.$$

Now note that if we add these together

$$u(x) + v(x) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0).$$

we end up with a polynomial that is degree two or less (it could be a polynomial of degree one, for example, if $a_2 + b_2 = 0$). Therefore this subcollection is closed under addition.

• Closure under multiplication (by a scalar): Suppose we take the element $u(x)$ and multiply it by the scalar r

$$ru(x) = (ra_2)x^2 + (ra_1)x + (ra_0),$$

which is also a polynomial of degree two or less. Therefore this subcollection is closed under multiplication. We then conclude that the set of polynomials of degree 2 or less forms a subspace of the vector space of all polynomials.

EXAMPLE **7.2** *(Rabenstein[1], section 3.2, problem 2) Show that the set of all elements of $\Re^2$ of the form $(1, a)$, where a is any constant, is not a subspace of $\Re^2$.*

• Closure under addition: We first identify two general elements in the subcollection as

$$\vec{u}_1 = (1, a_1), \qquad \vec{u}_2 = (1, a_2).$$

If we add these together we find that

$$\vec{u}_1 + \vec{u}_2 = (2, a_1 + a_2). \tag{7.1}$$

While the second component of this vector is a scalar (like a), the first component is not equal to 1 and therefore this subcollection is not closed under addition and therefore is not a subspace of $\Re^2$.

EXAMPLE **7.3** *Let $\mathcal{F}^2$ be the vector space of all twice differentiable functions defined on the interval $[a, b]$. Show that the set of all functions satisfying*

$$f''(x) + 2f(x) = 0, \tag{7.2}$$

is a subspace of $\mathcal{F}^2$.

• Closure under addition: We first identify two general elements in the subcollection as $f_1(x)$ and $f_2(x)$ where

$$\begin{aligned} f_1''(x) + 2f_1(x) &= 0, && (7.3a) \\ f_2''(x) + 2f_2(x) &= 0. && (7.3b) \end{aligned}$$

We need to first determine if $f_1 + f_2$ is in this subcollection. If we add the two equations in Eq. (7.3) we find that

$$\begin{aligned} f_1''(x) + 2f_1(x) + f_2''(x) + 2f_2(x) &= 0, \\ f_1''(x) + f_2''(x) + 2f_1(x) + 2f_2(x) &= 0, \\ [f_1(x) + f_2(x)]'' + 2[f_1(x) + f_2(x)] &= 0. \end{aligned} \tag{7.4}$$

This shows that the new function $f_1 + f_2$ satisfies Eq. (7.2) so this subcollection is closed under addition.

• Closure under multiplication: We need to determine whether the function rf_1 satisfies Eq. (7.2). Note

$$\begin{aligned} r[f_1''(x) + 2f_1(x)] &= 0, \\ [rf_1(x)]'' + 2[rf_1(x)] &= 0. \end{aligned}$$

The second equation shows that the function rf_1 satisfies Eq. (7.2). Therefore the subcollection is closed under multiplication. We have therefore shown that the subcollection described above is a subspace of $\mathcal{F}^2$.

8. Span of a Set, Linear Combination

Consider the set $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ made up of any k elements of a vector space V. This could be a set of vectors, a set of functions, a set of polynomials, etc. A *linear combination* of these elements is defined as

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \ldots + c_k\vec{u}_k, \tag{8.1}$$

where $c_1, \ldots c_k$ are constants. Now consider the set of all linear combinations of these elements (that is, the set of all elements of the form given by equation (8.1)). If we take two such linear combinations $l_1 = a_1\vec{u}_1 + a_2\vec{u}_2 + \ldots + a_k\vec{u}_k$ and $l_2 = b_1\vec{u}_1 + b_2\vec{u}_2 + \ldots + b_k\vec{u}_k$ and add them together we get another linear combination $l_1+l_2 = (a_1+b_1)\vec{u}_1+(a_2+b_2)\vec{u}_2+\ldots+(a_k+b_k)\vec{u}_k$. If we take the first linear combination, say, and multiply it by a scalar r we get another linear combination $rl_1 = (ra_1)\vec{u}_1 + (ra_2)\vec{u}_2 + \ldots + (ra_k)\vec{u}_k$. Therefore, the set of all linear combinations of these elements is closed under addition and multiplication; it is a <u>*subspace*</u> U of V. This subspace U, which is made up of all linear combinations of the elements $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_k$, is said to be <u>*spanned*</u> by $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_k$. That is, the elements $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_k$ <u>*span*</u> the subspace U.

An equivalent way of thinking about the span of a set of elements is the following. If the set $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ <u>*spans*</u> U, then any element of U can be written as a <u>*linear combination*</u> of the elements $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_k$. Note that there could be an infinite number of elements in U (which are all linear combinations of $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_k$). For example, if two functions $y_1(x)$ and $y_2(x)$ span the set of all solutions to the second order homogeneous linear differential equation $y''+p(x)y'+q(x)y = 0$, then any solution to this differential equation can be written as a linear combination of $y_1(x)$ and $y_2(x)$.

The idea here is that if you know that a solution you are seeking (to a differential equation or a linear algebraic equation) resides in a vector space and that you have a set of vectors (or elements) that span that vector space (or subspace), then, you can stop looking for solutions since all other solutions are <u>*linear combinations*</u> of the spanning vectors.

The above description tells how the concept of the span of a set of elements is related to linear combinations of those elements. In the sections that follow the notion of the span of a set will also be related to notions of <u>*dimension*</u>, <u>*basis*</u> and <u>*linear dependence*</u>. For now, we shall consider several examples based on the above definitions of a linear combination and the span of a set.

EXAMPLE **8.1** *(Rabenstein[1], section 3.2, problem 5) Let U be the subspace of $\Re^3$ (a vector space) that is spanned by the vectors $(-2,1,1)$ and $(1,-1,3)$. Interpret this subspace geometrically.* Since these two vectors <u>*span*</u> the subspace, any element of U must have the form

$$\vec{u} = a\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + b\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}. \tag{8.2}$$

You should be able to confirm for yourself that such elements do, in fact, form a subspace (i.e. show closure under addition and multiplication). The geometric interpretation of this subspace is a plane passing through the origin and the points $(-2,1,1)$ and $(1,-1,3)$ (i.e. their x, y, z coordinates). The parametric representation of the plane is

$$\begin{aligned} x &= -2a+b, \\ y &= a-b, \\ z &= a+3b, \end{aligned}$$

where a and b are the "parameters" of the plane (once given, they determine x, y, z). The equation for the plane can also be represented by

$$z + 4x + 7y = 0.$$

EXAMPLE **8.2** *(Rabenstein[1], section 3.2, problem 4) Let U be the subspace of $\Re^3$ (a vector space) that is spanned by the vector $(1, 3, -2)$. Interpret this subspace geometrically.* Since the vector *spans* the subspace, any element of U must have the form

$$\vec{u} = a \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}. \tag{8.3}$$

The geometric interpretation of this subspace is a line passing through the origin and the point $(1, 3, -2)$ in 3D. The parametric representation of the line is

$$x = a, \qquad y = 3a, \qquad z = -2a.$$

This can also be expressed as

$$\frac{x}{1} = \frac{y}{3} = \frac{z}{-2}.$$

EXAMPLE **8.3** *(Rabenstein[1], section 2.5, problem 20) Suppose we have a vector $\vec{v}$. This vector is a linear combination of the vectors $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_m$ if there are numbers $c_1, c_2, \ldots, c_m$ (not all zero) such that*

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_m\vec{v}_m. \tag{8.4}$$

Now consider the linear system $\boldsymbol{A}\vec{x} = \vec{b}$ where $\boldsymbol{A}$ is an $m \times n$ matrix, $\vec{x}$ is an n-dimensional vector and $\vec{b}$ is an m-dimensional vector.

$$\boldsymbol{A}\vec{x} = \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_n \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n = \vec{b}. \tag{8.5}$$

- *If $\boldsymbol{A}\vec{x} = \vec{b}$ has a solution $\vec{x}$ then $\vec{b}$ is a linear combination of the column vectors of the matrix $\boldsymbol{A}$.*
- *If $\vec{b} = c_1\vec{a}_1 + c_2\vec{a}_2 + \ldots + c_n\vec{a}_n$ then $x_1 = c_1$, $x_2 = c_2$, ..., $x_n = c_n$ (i.e. $\vec{x} = \vec{c}$) is a solution (see Theorem 2.4 in Rabenstein[1]).*

9. Linear Dependence

We've been describing sets of "vectors" and that if one takes all linear combinations of a set of vectors in a vector space one gets a subspace spanned by those vectors.

Consider the following situation. Let U be the subspace of $\Re^3$ that is spanned by the vectors

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}. \tag{9.1}$$

Therefore, any element of U must have the form

$$\vec{u} = a \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ b+c \\ a+c \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{b} \\ \tilde{c} \end{bmatrix}. \tag{9.2}$$

In geometrical terms, this is a plane in (x, y, z) space with $x = 0$. So apparently, these three vectors span this plane.

Now consider the subspace of $\Re^3$ spanned by the vectors

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{9.3}$$

We can show, as in the above case that these *two* vectors span the same geometrical space as the previous *three* vectors spanned. Somehow the vectors in Eq. (9.3) more efficiently describes the plane with $x = 0$ than do the vectors in Eq. (9.1). The reason for this greater efficiency has to do with the notion of *linear dependence*. This is also closely related to *dimension and basis* which we shall discuss in section 11.

DEFINITION **9.1** *A finite set of elements $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ of a vector space V is linearly dependent if there exist numbers $c_1, c_2, \ldots, c_m$ (not all zero) such that*

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_m\vec{v}_m = 0. \tag{9.4}$$

The set or the elements themselves are referred to as linearly dependent. If the only way to satisfy this equation is with $c_1 = c_2 = \ldots = c_m = 0$ then the elements (or the set) are linearly independent.

EXAMPLE **9.1** *So if we consider the previous case of the three vectors in Eq. (9.1) and try to satisfy*

$$c_1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{9.5}$$

we find that we can do this by taking the nonzero values $c_1 = 1$*,* $c_2 = 1$ *and* $c_3 = -1$ *(this is just one possible choice). Therefore, the vectors given in Eq. (9.1) are* *linearly dependent.* *The two vectors given in Eq. (9.3) are* *linearly independent* *since the only solution of*

$$c_1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{9.6}$$

is $c_1 = c_2 = 0$*.*

Another way to think of this result is described in following theorem.

THEOREM **9.1** *(Rabenstein*[1]*, Thm. 3.3) The set of elements* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ *is* *linearly dependent* *if and only if at least one of the elements in this set is a linear combination of the others.*

The basic idea behind this theorem is that for a linearly dependent set of vectors, one of the vectors must be redundant.

Now consider n vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in $\Re^m$. Each vector has m components, for example

$$\vec{v}_3 = \begin{bmatrix} v_{31} \\ v_{32} \\ \vdots \\ v_{3m} \end{bmatrix}. \tag{9.7}$$

The linear dependence of these n vectors is determined by finding for which values of c_i the following equation holds

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_n\vec{v}_n = 0. \tag{9.8}$$

If we rewrite this equation using slightly different notation we have

$$c_1 \begin{bmatrix} \uparrow \\ \vec{v}_1 \\ \downarrow \end{bmatrix} + c_2 \begin{bmatrix} \uparrow \\ \vec{v}_2 \\ \downarrow \end{bmatrix} + \ldots + c_n \begin{bmatrix} \uparrow \\ \vec{v}_n \\ \downarrow \end{bmatrix} \quad = \quad 0, \tag{9.9}$$

which we can also express in matrix form as

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} \vec{c} \quad = \quad 0, \tag{9.10}$$

or $\boldsymbol{A}\vec{c} = \vec{0}$ where $\boldsymbol{A}$ is an $m \times n$ matrix made up of the column vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ and $\vec{c}$ is an n-dimensional vector made up of the elements $c_1, c_2, \ldots, c_n$. Recall that we are looking for $\vec{c}$ which satisfy this equation. In particular we want to know if there are any nontrivial solutions (the trivial solution $\vec{c} = \vec{0}$ satisfies this equation).

Q: "When does $\boldsymbol{A}\vec{c} = \vec{0}$ have nontrivial solutions?"

A: <u>If $n > m$</u> there is always a nontrivial solution since the number of unknowns n is greater than the number of equations m (see Theorem 3.1). <u>If $n = m$</u> there exist nontrivial solutions when $\det\boldsymbol{A} = 0$ (see Theorem 5.2). <u>If $n < m$</u> then there may or may not be nontrivial solutions.

THEOREM **9.2** *(Rabenstein[1], Thm. 3.4) Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be elements of $\Re^m$. If $n > m$ then these elements are <u>linearly dependent</u>. If $n = m$ then these elements are <u>linearly dependent</u> if and only if*

$$det \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} = 0 \tag{9.11}$$

EXAMPLE **9.2** *Determine if the following vectors are linearly dependent or linearly independent.*

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

First evaluate the determinant

$$\det \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = 0 - (1)(-1) + (1)(1) = 2 \neq 0. \tag{9.12}$$

Since this determinant is nonzero, these vectors are *<u>linearly independent</u>*. Note that if any additional vector in $\Re^3$ was included in this list, those 4 vectors would automatically be *<u>linearly dependent</u>* as a consequence of the first part of Theorem 9.2.

EXAMPLE **9.3** *Are the following vectors linearly dependent?*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix}, \qquad \vec{v}_3 = \begin{bmatrix} -1 \\ -12 \\ -5 \end{bmatrix}.$$

To check, evaluate the determinant

$$\det \begin{bmatrix} 1 & 2 & -1 \\ 0 & 6 & -12 \\ -1 & 1 & -5 \end{bmatrix} = (1)(-18) - 0 + (-1)(-24 + 6) = 0. \tag{9.13}$$

Therefore, since the determinant vanishes, these vectors are *<u>linearly dependent</u>*. This means that

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = 0, \tag{9.14}$$

can be satisfied by values of c_1, c_2, c_3 that are not all zero. By setting up the system

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 6 & -12 \\ -1 & 1 & -5 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \vec{0} \tag{9.15}$$

and using elementary operations on the rows, show that any vector $\vec{c} = (c_1, c_2, c_3)$ of the form

$$\vec{c} = (-3a, 2a, a), \tag{9.16}$$

where a is a constant, is a solution.

10. Wronskian

We have just described the linear dependence of vectors in $\Re^m$. We now focus on the linear dependence of functions. Consider the functions

$$f(x) = e^x, \qquad g(x) = e^{-x}. \tag{10.1}$$

These two functions are *linearly independent* if the equation

$$c_1 e^x + c_2 e^{-x} = 0, \tag{10.2}$$

holds only for $c_1 = c_2 = 0$. If there exist nonzero c_1 and c_2 such that this equation holds for all x, then the functions are linearly dependent. This statement follows directly from the definition of linear dependence/independence given by Def. 9.1. Given the above Eq. (10.2) it is also true that we can take its derivative to obtain

$$c_1 e^x - c_2 e^{-x} = 0. \tag{10.3}$$

If we write Eqs. (10.2) and (10.3) in matrix notation we have

$$\begin{bmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \vec{0}, \tag{10.4}$$

or $\boldsymbol{A}\vec{c} = \vec{0}$. This has nontrivial solutions $\vec{c}$ if and only if $\det \boldsymbol{A} = 0$ for all values of x (see Thm. 5.2). However, we note that

$$\det \boldsymbol{A} = \begin{vmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{vmatrix} = -1 - 1 = -2 \neq 0 \quad \text{for any } x. \tag{10.5}$$

Therefore $c_1 = c_2 = 0$ is the only solution. Consequently, the functions e^x and e^{-x} are linearly independent. Notice that in the above example, the matrix $\boldsymbol{A}$ is really

$$\boldsymbol{A} = \begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix}. \tag{10.6}$$

The determinant of $\boldsymbol{A}$ is, in general, a function of x and is called the *Wronskian* (this is the 2×2 case) of the functions $f(x)$ and $g(x)$ at the point x. This definition generalizes to any number of functions.

EXAMPLE 10.1 *Are the functions $f(x) = e^x$, $g(x) = e^{-x}$ and $h(x) = e^{2x}$ linearly independent?* To check, consider the solutions $\vec{c}$ of the equation

$$c_1 e^x + c_2 e^{-x} + c_3 e^{2x} = 0. \tag{10.7}$$

In a similar fashion to the previous discussion we can formulate two more equations by taking derivatives of the above equation. In matrix–vector notation we obtain the system

$$\begin{bmatrix} e^x & e^{-x} & e^{2x} \\ e^x & -e^{-x} & 2e^{2x} \\ e^x & e^{-x} & 4e^{2x} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \vec{0}, \tag{10.8}$$

We next evaluate the Wronskian of the above 3×3 matrix (which is the determinant of this matrix) and we find

$$W = \begin{vmatrix} e^x & e^{-x} & e^{2x} \\ e^x & -e^{-x} & 2e^{2x} \\ e^x & e^{-x} & 4e^{2x} \end{vmatrix} = -6e^{2x}, \tag{10.9}$$

where W is the Wronskian. Since $W \neq 0$ for $x = 1$, say (or for any other value of x), then $c_1 = c_2 = c_3 = 0$. Therefore, these functions are linearly independent.

DEFINITION 10.1 *In general if $f_1(x)$, $f_2(x)$, $f_3(x)$, ..., $f_m(x)$ are defined functions with at least $m-1$ derivatives on some interval, we define the Wronskian of these functions to be*

$$W = det \begin{bmatrix} f_1(x) & f_2(x) & f_3(x) & \dots & f_m(x) \\ f_1'(x) & f_2'(x) & f_3'(x) & \dots & f_m'(x) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_1^{(m-1)}(x) & f_2^{(m-1)}(x) & f_3^{(m-1)}(x) & \dots & f_m^{(m-1)}(x) \end{bmatrix}, \tag{10.10}$$

where the superscripts in parentheses in the last row indicate the $(m-1)$ derivative of the function.

THEOREM 10.1 *(Rabenstein[1], Thm. 3.6) Consider the functions $f_1(x)$, $f_2(x)$, $f_3(x)$, ..., $f_m(x)$. If these functions are linearly dependent on some interval, then the Wronskian is zero for all values of x on that interval. THEREFORE, if $W \neq 0$ for at least one value of x on the interval, then the functions are linearly independent on that interval.*
CAUTION: If $W = 0$, it is not necessarily true that the functions are linearly dependent. As an example, consider the two functions $g_1(x) = x^2$ and $g_2(x) = x|x|$ which are defined for all values of x. These have $W = 0$ but are linearly independent on the interval $-\infty < x < \infty$.

EXAMPLE 10.2 *Show that the functions $f(x) = x^4$, $g(x) = x^3$, and $h(x) = x^2$ are linearly independent for all values of x.* Evaluate the Wronskian

$$W = \det \begin{bmatrix} x^4 & x^3 & x^2 \\ 4x^3 & 3x^2 & 2x \\ 12x^2 & 6x & 2 \end{bmatrix} = -2x^6. \tag{10.11}$$

Since we can identify values of x for which $W \neq 0$, these functions are linearly independent.

EXAMPLE **10.3** *Determine the whether the functions $f(x) = x^2 + 2x + 1$, $g(x) = x^2 - 3x$, and $h(x) = 5x + 1$ are linearly independent or linearly dependent.* Evaluate the Wronskian

$$W = \det \begin{bmatrix} x^2+2x+1 & x^2-3x & 5x+1 \\ 2x+2 & 2x-3 & 5 \\ 2 & 2 & 0 \end{bmatrix} = 0. \tag{10.12}$$

Since $W = 0$ Theorem 10.1 does not apply. If we go back to our original definition of linear independence we know that these functions will be linearly dependent if we can find c_1, c_2 and c_3 not all zero such that

$$c_1 f(x) + c_2 g(x) + c_3 h(x) = 0, \tag{10.13}$$

for all x. Notice that for the above functions, this equation is satisfied for $c_1 = 1$, $c_2 = -1$, and $c_3 = -1$; that is, $g(x) + h(x) = f(x)$. Therefore, these functions are linearly dependent.

EXERCISE **10.1** *(Rabenstein[1], section 3.4, problem 1f) Determine whether or not the functions x, e^x and xe^x are linearly dependent or linearly independent.*

11. Dimension, Basis

In the following section we describe a vector space by its size, or *dimension*. When we first introduced the notions of vector space and subspace, we talked about a set of vectors spanning that space. We also showed that it is possible for more than one set of vectors to span the same space and that the larger spanning set was linearly dependent. There are several other theorems and definitions which extend these ideas.

DEFINITION **11.1** *Suppose that $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly independent elements of a vector space V. If every set of more than n elements of V is linearly dependent then V is a vector space of dimension n. THE DIMENSION OF A VECTOR SPACE IS THE NUMBER OF ELEMENTS IN THE LARGEST LINEARLY INDEPENDENT SET IN THAT SPACE.*

DEFINITION **11.2** *Suppose V is a vector space of dimension n. Any set of n linearly independent elements of V is called a basis for V. A BASIS FOR V IS A SET OF VECTORS THAT SPANS V AND IS ALSO LINEARLY INDEPENDENT.*

So one could also think of a basis as a *maximal independent set* or a *minimal spanning set*. In order to describe the elements of a vector space, all one really needs is a basis! Notice that there is not just one basis for a vector space. For example, the two sets of vectors $\{(1,0),(0,1)\}$ and $\{(1,-1),(1,1)\}$ are both bases for $\Re^2$.

THEOREM **11.1** *(Rabenstein[1], Thm. 3.8) If there is a set of n linearly independent elements of V that spans V (i.e. a basis with n elements) then V has dimension n.*

EXAMPLE **11.1** *Use the above theorem to show that $\Re^3$ has dimension* 3. Consider the set of three vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}. \tag{11.1}$$

These vectors are linearly independent since

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1 \neq 0. \tag{11.2}$$

Furthermore, since

$$a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \tag{11.3}$$

represents a general element of $\Re^3$, these three vectors span $\Re^3$. Therefore, by Theorem 11. .1, $\Re^3$ has dimension 3. Similar reasoning shows that $\Re^m$ has dimension m.

EXAMPLE **11.2** *Show that the vector space of all polynomials of degree two or less has dimension* 3. Consider the set of elements

$$\{1, \quad x, \quad x^2\}. \tag{11.4}$$

If we examine the Wronskian $W(1, x, x^2)$

$$W = \det \begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 0 & 2 \end{bmatrix} = 2 \neq 0, \tag{11.5}$$

we can conclude that these three functions are linearly independent. Furthermore, as in the above example, since

$$a(1) + b(x) + c(x^2) = a + bx + cx^2, \tag{11.6}$$

characterizes all polynomials of degree two or less, these three functions span the vector space. Therefore, the vector space of all polynomials of degree two or less has dimension 3. Similar reasoning shows that the vector space of all polynomials of degree $m-1$ or less has dimension m.

THEOREM **11.2** *(Rabenstein[1], Thm. 3.9) If V is a vector space of dimension n, then every set of linearly independent elements that spans V (i.e. every basis) has exactly n elements.*

EXAMPLE **11.3** *(Roberts[2], p. 180, problem 17) Can we find a set of three polynomials $\{p_1(x), p_2(x), p_3(x)\}$ that will form a basis for the vector space of all polynomials of degree three or less?* No. One could argue as we did in the previous example that the vector space of all polynomials of degree three or less has dimension 4. By Theorem 11.2 , in order for a set to span this space, the set must contain exactly 4 elements. The proposed set $\{p_1(x), p_2(x), p_3(x)\}$ has only 3 elements and therefore cannot form a basis for this space.

EXAMPLE **11.4** *What is the dimension of the space spanned by*

$$g_1(x) = \sin x, \qquad g_2(x) = \cos x? \tag{11.7}$$

Elements in this space look like

$$c_1 \sin x + c_2 \cos x, \tag{11.8}$$

where c_1 and c_2 are constants. If these elements are linearly independent (and since they span this space) then by Theorem 11.1 the space would have dimension 2. So, consider the Wronskian

$$W = \det \begin{bmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{bmatrix} = -\sin^2 x - \cos^2 x = -1 \neq 0. \tag{11.9}$$

Therefore these two functions are linearly independent. By Theorem 11.1 the space has dimension 2.

It is also true that the set $\{1, x, x^2, x^3, \ldots\}$ forms a basis for the vector space of all polynomials. This vector space has <u>*infinite dimension*</u>.

The following describes a method for finding a basis for a subspace spanned by a set of vectors. Consider the subspace of $\Re^4$ spanned by the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 1 \\ -4 \\ 5 \\ 6 \end{bmatrix}. \tag{11.10}$$

The first thing to notice is that

$$\det \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} = 0. \tag{11.11}$$

Therefore, based on Theorem 9.2, we know that these vectors are linearly dependent. Consequently, these vectors do not form a basis; they are not linearly independent. A method for finding a basis starts by forming the matrix whose *rows* are made up of the above vectors, namely

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & -1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & -4 & 5 & 6 \end{bmatrix}, \tag{11.12}$$

and applying the elementary operations on the rows. So

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & -2 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & -4 & 4 & 4 \end{bmatrix} \qquad \begin{matrix} \\ [-1 \times \text{ROW1} + \text{ROW2}] \\ [-1 \times \text{ROW1} + \text{ROW3}] \\ [-1 \times \text{ROW1} + \text{ROW4}] \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{matrix} \\ [1/2 \times \text{ROW2}] \\ [-1 \times \text{ROW2} + \text{ROW3}] \\ [4 \times \text{ROW2} + \text{ROW4}] \end{matrix}$$

The above matrix is now as simplified as possible. If we extract the two remaining (nonzero) row vectors we have

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \qquad \vec{w}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \end{bmatrix} \tag{11.13}$$

These two vectors are linearly independent since $c_1\vec{w}_1 + c_2\vec{w}_2 = 0$ can only be satisfied by $c_1 = c_2 = 0$. Furthermore, the vectors $\vec{w}_1$ and $\vec{w}_2$ span the same space as did the vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ and $\vec{v}_4$; this result is the consequence of the following theorem.

THEOREM **11.3** *(Rabenstein[1] Thm 3.10) Given an $m \times n$ matrix $\boldsymbol{A}$. Let the matrix $\boldsymbol{B}$ be the matrix formed by performing a finite number of elementary operations on the rows of $\boldsymbol{A}$. Then the subspace of $\Re^n$ that is spanned by the rows vectors of $\boldsymbol{A}$ is the same as the subspace spanned by the row vectors of $\boldsymbol{B}$.*

Therefore, a basis for the space spanned by the vectors $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ and $\vec{v}_4$ is given by the set $\{\vec{w}_1, \vec{w}_2\}$. This subspace therefore has dimension 2.

12. Linear Transformations

A function $y = f(x)$ can be thought of as a *mapping* that starts with a point x and maps it into the point y. We write this as the mapping $f : \Re^1 \to \Re^1$. The function f maps an element of $\Re^1$ into an element in $\Re^1$. Similarly,

$$y_1 = f_1(x_1, x_2), \qquad y_2 = f_2(x_1, x_2), \tag{12.1}$$

is a mapping from $\Re^2$ to $\Re^2$. We say that the coordinate (x_1, x_2) is in the <u>*domain*</u> and the coordinate (y_1, y_2) is in the <u>*range*</u> of the mapping. Another mapping is the definite integral

$$\int_0^1 f(x), \tag{12.2}$$

which takes a function in $\mathcal{F}$ and maps it into $\Re^1$. Of particular interest are mappings that take us from one vector space to another (it could be the same vector space or it could be a different one).

Suppose we have two vector spaces and a mapping $T : U \to W$. This mapping takes elements from U and maps them into W (the range of T is in W but may not be the whole of W). Suppose that $\vec{u}$ and $\vec{v}$ are in U, then $\vec{u} + \vec{v}$ and $r\vec{u}$ are in U (since U is a vector space). Then $T(\vec{u})$, $T(\vec{v})$, $T(\vec{u} + \vec{v})$ and $T(r\vec{u})$ are in the range of U (which is in W). Now, since W is a vector space $T(\vec{u}) + T(\vec{v})$ is in W and $rT(\vec{u})$ is in W. These are not, however, necessarily in the range of T.

DEFINITION **12.1** *The mapping $T : U \to W$ is called a <u>linear transformation</u> (or linear operator, or linear mapping) if for every vector $\vec{u}$ and $\vec{v}$ in U, and for every scalar r, we have*

$$\begin{aligned} T(r\vec{u}) &= rT(\vec{u}), & (12.3\text{a}) \\ T(\vec{u} + \vec{v}) &= T(\vec{u}) + T(\vec{v}). & (12.3\text{b}) \end{aligned}$$

EXAMPLE **12.1** *Is $f : \Re^1 \to \Re^1$ where $f(x) = mx$ for constant m a linear transformation?* To answer this we need to see if (i) $f(rx) = rf(x)$, and (ii) $f(x_1 + x_2) = f(x_1) + f(x_2)$. Part (i) is satisfied since $m(rx) = r(mx)$ and part (ii) is satisfied since $m(x_1 + x_2) = mx_1 + mx_2$. Therefore, this mapping is a linear transformation.

EXAMPLE **12.2** *Is $f : \Re^1 \to \Re^1$ where $f(x) = mx + b$ for constant m and b a linear transformation?* To answer this we need to see if (i) $f(rx) = rf(x)$, and (ii) $f(x_1 + x_2) = f(x_1) + f(x_2)$. For part (i) we require $m(rx) + b = r(mx + b)$ but this only works if $b = rb$ (so $b = 0$ or $r = 1$). For part (ii) we require $m(x_1 + x_2) + b = (mx_1 + b) + (mx_2 + b)$, which again works only if $b = 0$. Consequently since both of these requirements are not satisfied, this mapping (with $b \neq 0$) is not a linear transformation.

EXERCISE **12.1** *Show that the mapping $f : \Re^1 \to \Re^1$ where $f(x) = ax^2$ is not a linear transformation.*

EXAMPLE **12.3** *Is the transformation $I : \mathcal{F} \to \Re^1$, where*

$$I(f) = \int_0^1 f(x)dx, \tag{12.4}$$

a linear transformation? Here $\mathcal{F}$ represents the vector space of all functions defined on $[0,1]$. The first thing we will do is to identify two elements of $\mathcal{F}$, call them $f(x)$ and $g(x)$. We first ask (i) Does $I(f+g) = I(f) + I(g)$?

$$\begin{aligned} I(f+g) &= \int_0^1 [f(x) + g(x)]\, dx, \\ &= \int_0^1 f(x)dx + \int_0^1 g(x)dx, \\ &= I(f) + I(g), \end{aligned}$$

so the first required property holds. Next, we ask (ii) Does $I(rf) = rI(f)$?

$$\begin{aligned} I(rf) &= \int_0^1 rf(x)dx, \\ &= r\int_0^1 f(x)dx, \\ &= rI(f), \end{aligned}$$

so the second required property holds. Therefore, this is a linear transformation.

Note that in Example 123 the functions $f(x)$ and $g(x)$ themselves may be nonlinear functions of x (for example $f(x) = x^2$). A linear transformation reflects how the elements of the domain appear; in this case the elements are f and g. The following example is not a linear transformation.

EXAMPLE **12.4** *Is the mapping $I : \mathcal{F} \to \Re^1$, where*

$$I(f) = \int_0^1 f^2(x)dx, \tag{12.5}$$

a linear transformation? Suppose we take two elements of the domain like we did in the previous example and ask (1) Does $I(f+g) = I(f) + I(g)$?

$$\begin{aligned} I(f+g) &= \int_0^1 (f(x) + g(x))^2 dx, \\ &= \int_0^1 \left[f^2(x)dx + 2f(x)g(x) + g^2(x)\right] dx, \\ &= I(f) + I(g) + \int_0^1 2f(x)g(x)dx. \end{aligned}$$

Since the integral appearing in the last line is not in general equal to zero, the first requirement for a linear transformation does not hold (the second does not either!). Therefore, this mapping is not a linear transformation.

EXERCISE **12.2** *Show that the mapping $T : \Re^n \to \Re^m$ defined by $\boldsymbol{A}\vec{x} = \vec{y}$, where $\boldsymbol{A}$ is an $m \times n$ matrix, $\vec{x}$ is an $n-$dimensional vector and $\vec{y}$ is an $m-$dimensional vector is a linear transformation.* This $m \times n$ matrix takes an $n-$dimensional vector and maps it into an $m-$dimensional vector.

A related result, which we state as a theorem is

THEOREM 12.1 *For any linear transformation $T : U \rightarrow W$, the zero element of U gets mapped to the zero element of W.*

It is also possible that other elements of a vector space, say $\Re^n$, can get mapped to the zero element of another vector space, say $\Re^m$. An example would be a nontrivial solution $\vec{x} \neq \vec{0}$ to the matrix equation $\boldsymbol{A}\vec{x} = \vec{0}$. The set of all elements of U mapped to the zero element of W forms a subspace of U. This subspace is called the *kernel* or *nullspace* of the linear transformation.

EXAMPLE 12.5 *(Rabenstein[1], section 3.7, problem 3) Consider $T : \Re^2 \rightarrow \Re^2$ defined by $T(\vec{x}) = \boldsymbol{A}\vec{x}$, where*

$$\boldsymbol{A} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}. \tag{12.6}$$

Does $\vec{x} = (1, 2)$ belong to the kernel of T? We note that

$$\boldsymbol{A}\vec{x} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{12.7}$$

and so the vector $(1, 2)$ is not in the kernel of T. In fact, if we notice that $\det \boldsymbol{A} \neq 0$, then we can conclude that there are no nontrivial solutions to $\boldsymbol{A}\vec{x} = \vec{0}$ and hence only the trivial solution $(0, 0)$ is in the kernel of T.

EXAMPLE 12.6 *Describe the kernel of the linear transformation $I : \mathcal{F} \rightarrow \Re^1$*

$$I(f) = \int_0^1 f(x)dx. \tag{12.8}$$

The kernel includes any function whose area under the $x-$axis is equal to the area above the $x-$axis between $x = 0$ and $x = 1$. One example would be $f(x) = \sin(2\pi x)$.

EXAMPLE 12.7 *The linear transformation*

$$I(f) = \int_0^x f(t)dt, \tag{12.9}$$

maps a function into a function $I : \mathcal{F} \rightarrow \mathcal{F}$. Compare this with the previous example.

PROPERTIES OF LINEAR TRANSFORMATIONS: There are two properties of linear transformations which we shall point out. They involve sums and products of linear transformations.

- If T_1 and T_2 are linear transformations from V to W we define the operator $T_1 + T_2$

$$(T_1 + T_2)\vec{u} \quad = \quad T_1\vec{u} + T_2\vec{u}. \tag{12.10}$$

For example, we can add two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ if they are the same size,

$$(\boldsymbol{A}+\boldsymbol{B})\vec{x} = \boldsymbol{A}\vec{x} + \boldsymbol{B}\vec{x}. \tag{12.11}$$

So if

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

then

$$(\boldsymbol{A}+\boldsymbol{B})\vec{x} = \left(\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\right)\vec{x} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\vec{x} = \begin{bmatrix} 3x_1 - 2x_2 \\ -x_1 + x_2 \end{bmatrix}$$

while

$$\boldsymbol{A}\vec{x} + \boldsymbol{B}\vec{x} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}\vec{x} + \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\vec{x} = \begin{bmatrix} 2x_1 - 3x_2 \\ -x_1 + 2x_2 \end{bmatrix} + \begin{bmatrix} x_1 + x_2 \\ -x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 - 2x_2 \\ -x_1 + x_2 \end{bmatrix}$$

Another example is given by the tranformations I_1 and I_2 where f is assumed to be defined on $[-1, 1]$.

$$I_1(f) = \int_{-1}^{0} f(x)dx, \quad I_2(f) = \int_{0}^{1} f(x)dx.$$

We then have

$$(I_1 + I_2)f = I_1(f) + I_2(f) = \int_{-1}^{1} f(x)dx.$$

- If $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations we define the operator T_2T_1,

$$(T_2T_1)\vec{u} = T_2(T_1\vec{u}).$$

Note that unless the vector space W is equivalent to the vector space V, we cannot define the operator T_1T_2 (this operator requires that the range of T_2 be in the domain of T_1). For example, consider the linear transformation $T_1 : \Re^2 \to \Re^2$ given by

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

and the linear transformation $T_2 : \Re^2 \to \Re^3$ given by

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 3 & 1 \end{bmatrix}.$$

We can define, for any 2−dimensional vector (lets say $\vec{x} = (1, 1)$),

$$T_2T_1\vec{x} = (\boldsymbol{B}\boldsymbol{A})\vec{x} = \boldsymbol{B}\left(\boldsymbol{A}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \boldsymbol{B}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 \\ 14 \\ 16 \end{bmatrix}. \tag{12.12}$$

However, for the case of these two matrices, the operator T_1T_2, which is the same as the matrix multiplication $\boldsymbol{AB}$, is not defined.

Another example of a linear transformation is a *differential operator*. Suppose we have the n^{th} order linear differential equation

$$a_0(x)\frac{d^n y}{dx^n} + a_1(x)\frac{d^{n-1}y}{dx^n} + \ldots + a_{n-1}(x)\frac{dy}{dx} + a_n(x)y = 0. \tag{12.13}$$

We then define a derivative operator $D = d/dx$ where we also use the notation that $D^n = d^n/dx^n$. This allows us to define the general n^{th} order differential operator

$$L \equiv a_0(x)D^n + a_1(x)D^{n-1} + \ldots + a_{n-1}(x)D + a_n(x), \tag{12.14}$$

so that the above differential equation (12.13) can be represented simply by $Ly = 0$. This differential operator L maps an element from the vector space of functions that have at least n derivatives to an element in the vector space of functions, $L : \mathcal{F}^n \to \mathcal{F}$.

13. Eigenvalues and Eigenvectors

Eigenvalues are an important issue in linear algebra. They will also play a key role in the study of linear systems of differential equations.

Consider the matrix/vector system

$$\boldsymbol{A}\vec{x} = \lambda\vec{x}, \tag{13.1}$$

where $\boldsymbol{A}$ is an $n \times n$ matrix and λ is an unknown constant. In order to solve this problem, we need to find a vector $\vec{x}$ such that when it gets multiplied by $\boldsymbol{A}$ it equals a constant times that vector.

In order to solve this problem, notice that the system can be rewritten as

$$\begin{aligned} \boldsymbol{A}\vec{x} &= \lambda\vec{x} = \lambda\boldsymbol{I}\vec{x}, \\ \boldsymbol{A}\vec{x} - \lambda\boldsymbol{I}\vec{x} &= 0, \\ (\boldsymbol{A} - \lambda\boldsymbol{I})\vec{x} &= 0. \end{aligned} \tag{13.2}$$

The last line represents a homogeneous linear system with a matrix $\boldsymbol{A} - \lambda\boldsymbol{I}$. This matrix is modified from the original matrix $\boldsymbol{A}$ by subtracting λ from each of the diagonal elements. We know from Theorem 52 that if $\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$ then this homogeneous problem has nontrivial solutions – notice that Eq. (13.1) is satisfied by the trivial solution $\vec{x} = \vec{0}$.

Now, since we assume that $\boldsymbol{A}$ is a given matrix, the idea is that we wish to *choose* λ such that $\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$. Such a value is called an *eigenvalue* of $\boldsymbol{A}$ (also called a characteristic value). The point of choosing such a value of λ is that then we know that the homogeneous problem has a nontrivial solution. We call such solutions *eigenvectors* (or characteristic vectors).

There are a wide variety of situations which arise when studying eigenvalue problems. Many of these are described in Boyce and DiPrima[4]. In these notes, we will treat only 2×2 matrices that have either *real and distinct* eigenvalues or *complex eigenvalues.* In each of these cases, we can identify two linearly independent eigenvectors. Mainly, the methods will be demonstrated by way of examples.

EXAMPLE **13.1** *Consider the following system*

$$\begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \vec{x} = \lambda \vec{x}, \tag{13.3}$$

and find the eigenvalues and eigenvectors. We begin by finding the values of λ that satisfy

$$\det \begin{bmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{bmatrix} = 0. \tag{13.4}$$

We find that

$$\begin{aligned} (4-\lambda)(-3-\lambda)+10 &= 0, \\ (\lambda-2)(\lambda+1) &= 0. \end{aligned} \tag{13.5}$$

This equation is sometimes called the <u>*characteristic polynomial equation*</u>. Therefore, the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -1$. Now, for each of these eigenvalues we want to find the corresponding nontrivial solution, or eigenvector. We first consider the case $\lambda = \lambda_1 = 2$. This gives us the system $((\boldsymbol{A} - \lambda \boldsymbol{I})\vec{x} = \vec{0})$

$$\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \vec{x} = \vec{0}. \tag{13.6}$$

This system represents two equations, each of which is $2x_1 - 5x_2 = 0$ (generally, the system may need to be reduced using elementary operations on the rows of the matrix before writing down the solution). Therefore, we have as a solution

$$\vec{x}_1 = a_1 \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \tag{13.7}$$

where a_1 is an arbitrary, but nonzero, constant. This is the eigenvector $\vec{x}_1$ corresponding to the eigenvalue $\lambda_1 = 2$. We next consider the case $\lambda = \lambda_2 = -1$. This gives us the system $((\boldsymbol{A} - \lambda \boldsymbol{I})\vec{x} = \vec{0})$

$$\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \vec{x} = \vec{0}. \tag{13.8}$$

This system represents two equations, each of which is $x_1 - x_2 = 0$. Therefore, we have as a solution

$$\vec{x}_2 = a_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{13.9}$$

where a_2 is an arbitrary, but nonzero, constant. This is the eigenvector $\vec{x}_2$ corresponding to the eigenvalue $\lambda_2 = -1$.

EXAMPLE **13.2** *Find the eigenvalues and eigenvectors for the following matrix*

$$\begin{bmatrix} 2 & 5 \\ -1 & -2 \end{bmatrix}. \tag{13.10}$$

We begin by finding the values of λ that satisfy

$$\det \begin{bmatrix} 2-\lambda & 5 \\ -1 & -2-\lambda \end{bmatrix} = 0. \tag{13.11}$$

This gives

$$\begin{aligned} -(2-\lambda)(2+\lambda)+5 &= 0, \\ \lambda^2+1 &= 0. \end{aligned} \tag{13.12}$$

This equation has no real solutions but has complex solutions given by $\lambda_{1,2} = \pm i$, where the imaginary number i is defined by $i \equiv \sqrt{-1}$. These are the eigenvalues. The procedure to find the corresponding eigenvectors is no different from the previous example, except for the fact that the eigenvectors will in general be complex and we will need to use complex arithmetic to obtain them. Let us first consider the case with $\lambda = i$. This gives us the system $((\boldsymbol{A} - \lambda \boldsymbol{I})\vec{x} = \vec{0})$

$$\begin{bmatrix} 2-i & 5 \\ -1 & -2-i \end{bmatrix} \vec{x} = \vec{0}. \tag{13.13}$$

Here we will use elementary operations on the rows to simplify the problem.

$$\begin{gathered} \begin{bmatrix} 2-i & 5 \\ -1 & -2-i \end{bmatrix}, \\ \begin{bmatrix} 0 & (2-i)(-2-i)+5 \\ -1 & -2-i \end{bmatrix} \qquad [(2-i) \times \text{ROW2} + \text{ROW1}], \\ \begin{bmatrix} 0 & 0 \\ -1 & -2-i \end{bmatrix} \end{gathered} \tag{13.14}$$

This represents the single equation $-x_1 + (-2-i)x_2 = 0$. Therefore, we have as a solution

$$\vec{x}_1 = a_1 \begin{bmatrix} -(2+i) \\ 1 \end{bmatrix}, \tag{13.15}$$

where a_1 is an arbitrary, but nonzero, complex constant. This is the eigenvector $\vec{x}_1$ corresponding to the eigenvalue $\lambda_1 = i$. Let us next consider the case with $\lambda = -i$. This gives us the system $((\boldsymbol{A} - \lambda \boldsymbol{I})\vec{x} = \vec{0})$

$$\begin{bmatrix} 2+i & 5 \\ -1 & -2+i \end{bmatrix} \vec{x} = \vec{0}. \tag{13.16}$$

Here we will use elementary operations on the rows to simplify the problem.

$$\begin{bmatrix} 2+i & 5 \\ -1 & -2+i \end{bmatrix},$$

$$\begin{bmatrix} 0 & (2+i)(-2+i)+5 \\ -1 & -2+i \end{bmatrix} \qquad [(2+i) \times \text{ROW2} + \text{ROW1}] \quad ,$$

$$\begin{bmatrix} 0 & 0 \\ -1 & -2+i \end{bmatrix} \tag{13.17}$$

This represents the single equation $-x_1 + (-2+i)x_2 = 0$. Therefore, we have as a solution

$$\vec{x}_1 = a_2 \begin{bmatrix} -2+i \\ 1 \end{bmatrix}, \tag{13.18}$$

where a_2 is an arbitrary, but nonzero, complex constant. This is the eigenvector $\vec{x}_2$ corresponding to the eigenvalue $\lambda_2 = i$. You may notice that the two eigenvalues are complex conjugates of each other and that the two eigenvectors are complex conjugates of each other as well. So actually, once we found the first eigenvector, we could have just written down the other eigenvector without further calculations.

As was mentioned earlier, there are many different scenarios that can play out in eigenvalue problems. The above examples describe only two such situations. See Boyce and DiPrima[4] for further discussion of eigenvalues and eigenvectors.

References

[1] A.L. Rabenstein, *Elementary Differential Equations With Linear Algebra*, Fourth Edition, (Harcourt Brace Jovanovich, Fort Worth, 1992).

[2] A.W. Roberts, *Elementary Linear Algebra*, Second Edition, (Benjamin/Cummings, Reading, MA, 1985).

[3] G. Strang, *Linear Algebra and its Applications*, Third Edition, (Harcourt Brace Jovanovich, San Diego, 1988).

[4] W.E. Boyce and R.C. DiPrima, *Elementary Differential Equations and Boundary Value Problems* Sixth Edition, (Wiley, New York, 1997).