Introduction to Linear Algebra for APPM 2360

by Lora Billings and Peter Staab
modified by J.D. Meiss

November 11, 1999

Contents

Chapter 1

Solving Linear Systems

1.1 Introduction

One of the most important relationships in mathematics is a linear relationship. You are already familiar with this. Suppose that you have the following recipe for Ted's Original Texas Road-Kill Chili (found at *http://www.flash.net/∼rockware/chili.html*):

4 lbs	fresh road-kill	2-4 tbsp	chili powder
1 tsp	black pepper	1 tbsp	salt
2	beers or 750ml zinfindel	4 cloves	crushed garlic
1-3 tsp	oregano	1	16 oz. can Cantadina tomato sauce
2	Jalapeno peppers (chopped)	1 tbsp	finely chopped green onions
1 tbsp	cumin powder	2 tbsp	chopped bell pepper

and the recipe serves six people.

If you want to serve chili for eighteen, you have to triple the recipe (multiply the quantity of each ingredient by three). This is a linear scaling. The following is the mathematical definition of algebraic linearity.

Definition 1.1.1 *An algebraic equation is* **linear** *in the variables $x_1, x_2, \ldots, x_n$ if it has the following form:*

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

where $a_{11}, a_{12}, \ldots, a_{1n}$, and b_1 are constants.

Example 1.1.1 James spends $108.91 on CD's. The prices of the various CD's are $9.99, $11.99, and $14.99. What is the algebraic relationship that describes his music purchase?

Answer:

$$9.99x_1 + 11.99x_2 + 14.99x_3 = 108.91,$$

where x_1, x_2, and x_3 are the number of CD's that he bought at $9.99, $11.99, and $14.99, respectively. The total price of CD's is a linear relationship which equals $108.91.

There are many examples of single linear equations, however as in the example just given, there is not enough information to determine the solution to all the variables. In this case, we would need more information–such as the total number of CD's. Often, there will be multiple equations to describe the relationships between unknown variables. If all the equations are linear, then the collection is called a **linear system**.

Definition 1.1.2 *A* **linear system** *is a collection of linear equations that can be written in the form:*

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

Notice that the number of equations is m and the number of unknowns is n. It is important that both the a_{ij}'s, i.e. the coefficients in the equations, and the b_i's, the right hand sides,be constant.

Linear systems are found throughout engineering and mathematics and their importance is comparable to the importance of algebra and calculus. The following example shows one of a wide range of problems where linear systems arise:

Example 1.1.2 The temperature distribution in a long metal rod can be found by imagining that the rod consists of a series of small pieces of metal connected together. (See Figure 1.1) The rod is in local thermal equilibrium when the temperature of any piece of the rod is the average temperature of its neighbors. In mathematical terms this means:

$$T_i = \frac{T_{i-1} + T_{i+1}}{2}$$

for $i = 2, 3, \ldots, N-1$, where N is the total number of "pieces" of the rod. Because the average temperature is an example of a linear equation, the resulting collection of equations is a linear system.

Let's consider an actual case with a 4-piece rod. If one end is in ice water ($0°$) and the other is boiling water ($100°$), then what is the linear system in standard form (as given in the definition above)?

Answer: The following linear system describes the temperature in the rod,

$$\begin{aligned} 2T_1 - T_2 &= 0, \\ -T_1 + 2T_2 - T_3 &= 0, \\ -T_2 + 2T_3 - T_4 &= 0, \\ -T_3 + 2T_4 &= 100. \end{aligned}$$

In this case, the variables are the temperatures, T_1, T_2, T_3, and T_4. The right hand side of the first equation is 0 because the left end of the rod is in ice water. The right hand side of the last equation is 100 because the right end of the rod is in boiling water. Notice that there are 4 variables and 4 equations. You may remember a saying, "You must have the same number of equations as variables." This is true for this example, but isn't always true. We will discuss this in more detail in Section 2.3. You will be asked to solve this system in problem 11.

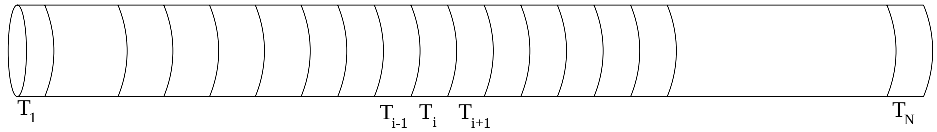


Figure 1.1: To determine the temperature distribution of the rod, the rod can be thought of existing as small distinct pieces, where each has the same temperature. In this figure there are N pieces of rod.

The equation of a plane in $\mathbf{R}^3$ is given by the general form $ax + by + cz = d$, where a, b, c, and d are constants. This is another example of a linear equation. If $x + y + 6z = 17$, $4x - 3y - 2z = 4$, and $2x + 3y + z = -1$ are three planes in 3-space, then the three equations form the linear system:

$$\begin{aligned} x + y + 6z &= 17, \\ 4x - 3y - 2z &= 4, \\ 2x + 3y + z &= -1, \end{aligned} \tag{1.1}$$

where x, y, and z are coordinates in $\mathbf{R}^3$. The solution to this linear system is the intersection of the planes in $\mathbf{R}^3$, which is a point in this case. You may remember from geometry that three planes can intersect in a point, a line, a plane, or have no intersection at all. We will see below how this geometric intuition can be turned into an algebraic statement about linear systems.

1.2 Basic Matrix Operations

One objective of linear algebra is to find the solution of linear systems, like the ones presented in the previous section. Before getting into the details of solving these systems, linear algebra presents us with a shortcut notation for these systems. The notation allows us to simplify the solution process, as you will see in the next section.*The goal of this section is to introduce you to this notation and show you some basic operations.*

The linear system of planes in $\mathbf{R}^3$ in (1.1) can be written in matrix form:

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 6 \\ 4 & -3 & -2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 17 \\ 4 \\ -1 \end{pmatrix}. \\ (3 \text{ by } 3 \text{ matrix})(3 \text{ vector}) &= (3 \text{ vector}) \end{aligned} \tag{1.2}$$

The **matrix** on the left is simply a grid of numbers. In this case it has 3 rows and 3 columns, so its called a 3 by 3 matrix, thus the **size** of the matrix is 3 by 3. If the number of rows equals the number of columns then the matrix is **square**, as in (1.2).

There are two vectors in (1.2) and both of them have length three. The vector on the left side of the equation has length 3 and it contains the unknowns x, y, and z. The vector on the right of (1.2) contains the right hand side of (1.1).

The above matrix is found by taking the coefficients of the variables of (1.1). The rows of the matrix represent each equation from (1.1) and the columns represent the coefficients of each unknown (the elements of the first column of the matrix are the coefficients of x in (1.1), the second column contain the coefficients of y, etc.).

Lastly, it is important to know how to rewrite (1.2) in another form, called an **augmented matrix**, which is found by removing the unknown variables and drawing a vertical line between the coefficient matrix and the right hand side. The matrix equation (1.2) can be written:

$$\left(\begin{array}{rrr|r} 1 & 1 & 6 & 17 \\ 4 & -3 & -2 & 4 \\ 2 & 3 & 1 & -1 \end{array}\right). \tag{1.3}$$

In the general case, where $\mathbf{A}\mathbf{x} = \mathbf{b}$, the augmented matrix is written

$$(\mathbf{A}|\,\mathbf{b})\,,$$

where the unknown vector, $\mathbf{x}$ is not included. (Note: the matrices and vectors in the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ are in bold as are all vectors and matrices throughout this book.) The augmented matrix is the form used to solve linear systems as will be shown in next section.

1.2.1 Review of Vectors

Recall that a vector is used as a way of representing direction and magnitude as in forces. The magnitude or length of a vector

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

is defined

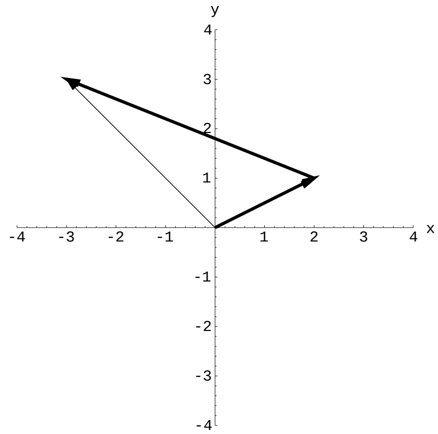
$$|\mathbf{x}| = \sqrt{x_1^2 + \ldots + x_n^2}\ .$$

Notice $|\mathbf{x}|$ is a scalar quantity.

The addition of vectors is performed by summing the individual components. Here's an example in $\mathbf{R}^3$:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 8 \\ 0 \\ -5 \end{pmatrix} = \begin{pmatrix} 9 \\ 2 \\ -2 \end{pmatrix}.$$

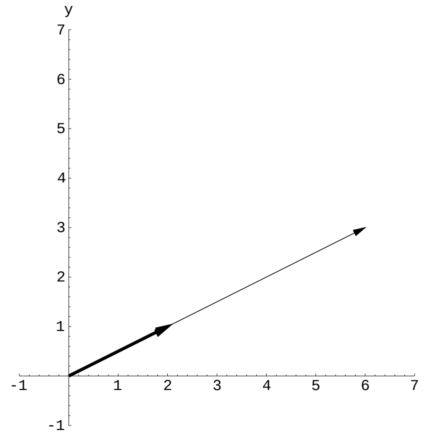
Graphically, addition is performed by graphing the first vector at the origin and graphing the second vector from the endpoint of the first vector. Connecting the origin to the end of the second vector is the resultant vector.



This is an example of vector addition in the plane.

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -5 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$$

The vectors in bold are the original vectors and the non-bold vector is the resultant.



This figure shows the multiplication of a vector by a scalar in the plane.

$$3\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$

The vector in bold is the original vector and the non-bold vector is the resultant.

The **scalar multiplication** of a vector is the product of a number and a vector. In this example

$$2\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix},$$

the 2 multiplies each element of the vector. The resultant is a vector in the same direction but with twice the magnitude.

Another operation on vectors is the **dot product**, which is defined between two vectors $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$. The dot product is defined as

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

where the $\cdot$ is the notation for dot product.

1.2.2 Addition and Scalar Multiplication of Matrices

Matrix addition and scalar multiplication behave the same way as vector addition and scalar multiplication.

$$\begin{array}{ccccc} \text{(2 by 3)} & + & \text{(2 by 3)} & = & \text{(2 by 3)} \\ \begin{pmatrix} 0 & -2 & -2 \\ 1 & -8 & 0 \end{pmatrix} & + & \begin{pmatrix} -10 & 4 & -7 \\ 9 & 0 & 2 \end{pmatrix} & = & \begin{pmatrix} -10 & 2 & -9 \\ 10 & -8 & 2 \end{pmatrix}. \end{array}$$

Here's an example of the multiplication of a scalar by a matrix,

$$\begin{array}{ccc} \text{scalar} \times \text{(2 by 3)} & = & \text{(2 by 3)} \\ 3\begin{pmatrix} 0 & -2 & -2 \\ 1 & -8 & 0 \end{pmatrix} & = & \begin{pmatrix} 0 & -6 & -6 \\ 3 & -24 & 0 \end{pmatrix}. \end{array}$$

Notice that the size of the two matrices in the addition must be the same. *Two matrices of different sizes cannot be added.* Also, the resultant matrix in either the addition or the scalar multiplication is the same size as the matrices on the left side.

1.2.3 Matrix-Vector Multiplication

The matrix equation in (1.2) has a matrix-vector multiplication on the left hand side. The first row of the matrix is multiplied by the column vector in a dot product sense. This gives the first element of the resultant matrix. The second element is found in a similar manner with a dot product of the second row of the matrix with the vector.

Example 1.2.1 Multiply the matrices $\mathbf{A}$ and $\mathbf{B}$ given by

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 9 \\ -1 & 2 & -4 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 5 \\ 3 \\ -2 \end{pmatrix}.$$

Answer:

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

Notice that the sizes of the matrix and vector have to be compatible. Because of the way the multiplication takes place, *the number of columns of the matrix must* **equal** *the number of rows of the vector.* In this case the matrix is a two by three and the vector has length three. Because there are two rows in the matrix, the resultant vector has length two.

In general, the result of the product of an m by n matrix, $\mathbf{A}$, with an n-vector, $\mathbf{x}$, is an m-vector, as shown here,

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

1.2.4 Matrix-Matrix Multiplication

To find the product of two matrices, the columns of the second matrix are treated like vectors and the matrix-vector multiplication is done as above. Notice that the dimensions act the following way:

$$\begin{array}{l} \downarrow \qquad\qquad\qquad\qquad \downarrow \ \text{the dimensions of the product} \\ (m \ \text{by} \ \ n)(n \ \ \text{by} \ \ p) = (m \ \ \text{by} \ \ p) \\ \qquad\qquad \underline{\uparrow \ \ \uparrow} \ \ \text{these must match} \end{array}$$

Example 1.2.2 Multiply

$$\begin{pmatrix} 2 & 4 \\ 3 & -8 \\ -1 & 5 \\ 4 & 4 \end{pmatrix} \begin{pmatrix} -5 & -6 & 3 \\ -2 & 4 & 5 \end{pmatrix}.$$

Answer:

$$= \begin{pmatrix} (2)(-5)+(4)(-2) & (2)(-6)+(4)(4) & (2)(3)+(4)(5) \\ (3)(-5)+(-8)(-2) & (3)(-6)+(-8)(4) & (3)(3)+(-8)(5) \\ (-1)(-5)+(5)(-2) & (-1)(-6)+(5)(4) & (-1)(3)+(5)(5) \\ (4)(-5)+(4)(-2) & (4)(-6)+(4)(4) & (4)(3)+(4)(5) \end{pmatrix}$$

$$= \begin{pmatrix} -18 & 4 & 26 \\ 1 & -50 & -31 \\ -5 & 26 & 22 \\ -28 & -8 & 32 \end{pmatrix}.$$

You can think of the matrix-matrix multiplication as some number of matrix-vector multiplications. In this case, the second matrix has three columns, therefore you can this of this product as three matrix-vector products, with the result having three columns.

1.2.5 Transposes

Another basic matrix operation is the matrix transpose–this exchanges the rows and columns of the matrix. If the matrix $\mathbf{A}$ is given by

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

then the transpose is given by

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

Notice that the transpose of an m by n matrix is an n by m matrix. For a shorthand notation, the transpose of the matrix $\mathbf{A}$ is written $\mathbf{A}^T$.

Example 1.2.3 Find the transpose of

$$\mathbf{A} = \begin{pmatrix} 3 & -7 & 1 & 1 \\ 2 & 9 & -10 & 25 \end{pmatrix}.$$

Answer:

$$\mathbf{A}^T = \begin{pmatrix} 3 & 2 \\ -7 & 9 \\ 1 & -10 \\ 1 & 25 \end{pmatrix}.$$

A vector can either be a column vector or a row vector. A **column vector** has the form

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

and a **row vector** has the form

$$(x_1, x_2, \ldots, x_n).$$

If the type of vector is not specified then it is assumed that it is a column vector. Notice that all of the vectors so far in this book have been column vectors. *It is important for matrix-vector multiplications that the vectors are column vectors (or the dimensions won't match up).* Often times, column vectors are written in the form $(x_1, \ldots, x_n)^T$ because they take up less space in printed notation. The transpose takes a row vector to a column vector.

Lastly, the dot product that you saw in section 1.2.1 can be written as

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y},$$

where $\mathbf{y}^T$ can be thought of as a 1 by n matrix instead of row vector of length n.

1.2.6 The Identity Matrix

It's every sports fan's dream to have a team they admire be number one. Well, mathematicians get a little teary-eyed with the number one also, for numerous reasons. A particular reason is that when you multiply any number by one it is not changed. That is, $1 \cdot x = x$, the number 1 is called the multiplicative identity. There is also a matrix, $\mathbf{I}$, that has the property that $\mathbf{IA} = \mathbf{A}$, and as you can guess, it's called the **identity matrix**. This matrix is square, has ones along the diagonal, and zeros above and below the diagonal. The 3 by 3 identity matrix is

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1.3 Gaussian Elimination

In the previous section, we saw the linear system representing three planes in $\mathbf{R}^3$:

$$\begin{aligned} x + y + 6z &= 17, \\ 4x - 3y - 2z &= 4, \\ 2x + 3y + z &= -1. \end{aligned} \tag{1.4}$$

We can solve this system by solving for one of the variables in the first equation. Solving for x in the first equation gives $x = 17 - y - 6z$, and then substituting into the second and third equations of (1.4) gives,

$$\begin{aligned} 4(17 - y - 6z) - 3y - 2z &= 4, \\ 2(17 - y - 6z) + 3y + z &= -1, \end{aligned}$$

or

$$\begin{aligned} -7y - 26z &= -64, \\ y - 11z &= -35. \end{aligned} \tag{1.5}$$

Notice that after one step, we have reduced a 3 by 3 system (3 equations and 3 variables) to a 2 by 2 system, which is easier to solve. From the second equation of (1.5), we can solve for y, or $y = -35 + 11z$, and substitute it into the first equation of (1.5), yielding

$$-7(-35 + 11z) - 26z = -64,$$

or $z = 3$. Once we have found z, we can substitute it into $y = -35 + 11z$ or $y = -2$. Lastly, using y and z, we can get x from $x = 17 - y - 6z$, or $x = 1$. Therefore the solution to the linear system in (1.4) is

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

Though what we have just done is reasonably easy for three variables and three equations,it becomes considerably more tedious for four variables and four equations, and unbearably difficult for 100 variables and 100 equations. In this section, you will learn how to solve matrix equations (linear systems) by using a simple algorithm, called **Gaussian elimination**.

> Gaussian elimination is a series of **elementary row operations** performed on a matrix (usually an augmented matrix). The goal is to reduce the original matrix to a matrix with zeros below the diagonal.

We start by rewriting (1.4) as an augmented matrix as was shown in section 1.2,

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 4 & -3 & -2 & 4 \\ 2 & 3 & 1 & -1 \end{array}\right).$$

The main reason for this is that the augmented matrix contains all the information about our system of equations (except for the names of the variables), but is more compact to write down. To develop the solution algorithm we first note that there are some simple manipulations that one can carry out on the augmented matrix without changing the answer. These are called **elementary row operations**.

For example, since each row of the augmented matrix stands for one of our three equations, it should make no difference in which order these equations are written. So the augmented matrix

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & -1 \\ 4 & -3 & -2 & 4 \\ 1 & 1 & 6 & 17 \end{array}\right).$$

is just as valid as the previous one, since we have just written the first equation last and the last equation first. Thus, for augmented matrices it is legitimate to exchange any two rows without changing the problem we are solving, and therefore without changing the solution. This is the first of our elementary row operations. We denote it by S_{13}, which means swap rows 1 and 3.

There are two more elementary row operations that we are going to need. Notice that if we take the first equation, $x + y + 6z = 17$ and multiply the entire equation by any nonzero scalar, c, it does not change the equation, since $cx + cy + 6cz = 17c$ has exactly the same solutions as before. On the augmented matrix, this means we can multiply any row by a constant without changing the answer. We denote this operation by $cR_1 \to R_1$, i.e. mulitply row 1 by 1 and return it to row 1.

Finally we can add any equation to any other in our system without changing the answer. Even more generally, we can add, say, 3 times the first equation to the second and do no harm. This replaces the second equation by $4x - 3y - 2z + 3(x + y + 6z) = 4 + 3(17)$, or $7x + 0y + 16z = 55$. In this case the augmented matrix would become

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 7 & 0 & 16 & 55 \\ 2 & 3 & 1 & -1 \end{array}\right).$$

We will see that this is the most useful row operation. We denote it by $3R_1 + R_2 \to R_2$, i.e. multiply row 1 by 3 and add it to row 2, then return it to row 2.

To summarize, there are three elementary row operations:

S_{ab}	Swap rows a and b.
$cR_a \to R_a$	Multiply row a by c and return it to row a.
$cR_a + R_b \to R_b$	Multiply row a by c and add it to row b then return it to row b.

The process called "Gaussian Elimination" is simply a combination of elementary row operations designed to make the solution of the linear system easier.

The first objective of Gaussian elimination is to get zeros into the second and third row of the first column, by taking linear combinations of the rows. For instance if we multiply each element of the first row of (1.3) by -4 and add to each element of the second row then (1.3) becomes

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & -7 & -26 & -64 \\ 2 & 3 & 1 & -1 \end{array}\right).$$

To eliminate the 2 in the first column of the third row, we multiply the top row by -2 and add to the last row. The matrix then becomes:

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & -7 & -26 & -64 \\ 0 & 1 & -11 & -35 \end{array}\right). \tag{1.6}$$

Notice that the smaller 2 by 3 augmented matrix in the lower right hand corner represents the equations in (1.5). The next step is to get a 0 into the second column of the third row. To do this we multiply the second row by $\frac{1}{7}$ and add to the third row (we do this because this will eliminate the one in the second column). The resulting matrix is:

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & -7 & -26 & -64 \\ 0 & 0 & -\frac{103}{7} & -\frac{309}{7} \end{array}\right). \tag{1.7}$$

This is the result of Gaussian elimination. Now that you see the results, you can see that we desire to get a matrix in **upper triangular form,** which means that every element below the main diagonal is 0.

The claim is that the system of equations represented by this augmented matrix has the same solution as the original system, and, moreover is easier to solve. To solve it, we perform what is is called **back substitution**. This is precisely what we did when we solved (1.4), solving first for z then finding y, then finding x. Thus the third equation says $\frac{-103}{7}z = \frac{-309}{7}$, which means that $z = 3$. Whew, that was easy, wasn't it! Now that we know z, we look at the second equation, which becomes

$$-7y + -26(3) = -64$$

This is also easy, since we knew z, and gives $y = -2$. Finally, the first equation is

$$x + 1(-2) + 6(3) = 17$$

Again, since we are doing things backwards, there is only one unknown, x, so we easily get the solution, $x = 1$.

If you like, you can also do back substitution on the augmented matrix. The goal here is to apply the elementary row operations to eliminate the entries above the diagonal. First, solve for z. We do this by multiplying the last row of (1.7) by $-\frac{7}{103}$, which results in:

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & -7 & -26 & -64 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

Next, we can get a zero in the third column of the second row by multiplying the third row of (1.3) by 26 and adding to the second row. The result is:

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & -7 & 0 & 14 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

We now want a 1 on the second diagonal element (that is, second column of the second row), so that the value of y is explicit. To obtain this, we multiply the second row by $-\frac{1}{7}$. The result is:

$$\left(\begin{array}{ccc|c} 1 & 1 & 6 & 17 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

Next we get 0's in the second and third columns of the first row. We can do this by multiplying the third row by -6 and adding to the first row to get

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

And finally we can get a 0 in the second column of the first row by multiplying the second row by -1 and adding to the first row. The final result is:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right). \tag{1.8}$$

We can now get the solution to the linear system by reading it off from the matrix. Remember that the augmented matrix represents a linear system. Therefore we can write down the linear system of (1.8),

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

or $x = 1, y = -2$, and $z = 3$.

You still may wonder why these steps were easier than the previous method. You are correct in saying we used far more steps in this example of Gaussian elimination than the first method. However, we wrote out each step in great detail. Of course, we can do things much more compactly now that you understand the process.

The following illustrates an example of Gaussian elimination using the elmementary row operation notation:

Example 1.3.1 Solve

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

Answer:

Write the linear system as an augmented matrix.

$$\left(\begin{array}{rrr|r} 2 & -1 & -1 & 0 \\ 1 & 1 & 2 & 6 \\ -3 & 0 & 2 & 3 \end{array}\right).$$

Note that the 0 in the second column of row 3 is because the coefficient of y in the last equation is zero.

$$S_{12}, \quad \left(\begin{array}{rrr|r} 1 & 1 & 2 & 6 \\ 2 & -1 & -1 & 0 \\ -3 & 0 & 2 & 3 \end{array}\right).$$

Swap rows 1 and 2 (this is the easiest way to get a 1 in the first column of the first row).

$$\begin{array}{rcl} -2R_1 + R_2 & \rightarrow & R_2, \\ 3R_1 + R_3 & \rightarrow & R_3, \end{array} \quad \left(\begin{array}{rrr|r} 1 & 1 & 2 & 6 \\ 0 & -3 & -5 & -12 \\ 0 & 3 & 8 & 21 \end{array}\right).$$

To get zeros in the second and third rows of the first columns, we take linear combinations of the first rows.

$$R_2 + R_3 \rightarrow R_3, \quad \left(\begin{array}{rrr|r} 1 & 1 & 2 & 6 \\ 0 & -3 & -5 & -12 \\ 0 & 0 & 3 & 9 \end{array}\right).$$

Add row 2 to row 3. The matrix is in upper triangular form. The remaining steps use back substitution.

$$\frac{1}{3}R_3 \rightarrow R_3, \quad \left(\begin{array}{rrr|r} 1 & 1 & 2 & 6 \\ 0 & -3 & -5 & -12 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

Multiply row 3 by $\frac{1}{3}$.

(When row 3 has zeros in the first two columns and 1 in the third column, we use it to get zeros in the third column of the other two rows.)

$$\begin{array}{rcl} 5R_3 + R_2 & \rightarrow & R_2, \\ -2R_3 + R_1 & \rightarrow & R_1, \end{array} \quad \left(\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 0 & -3 & 0 & 3 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

Multiply row 3 by 5 and add to row 2, then multiply row 3 by -2 and add to row 1.

$$-\frac{1}{3}R_2 \rightarrow R_2, \quad \left(\begin{array}{rrr|r} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

To get the needed 1 in column 2 of row 2, we multiply row 2 by $-\frac{1}{3}$.

$$-R_2 + R_1 \rightarrow R_1, \quad \left(\begin{array}{rrr|r} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right).$$

The final step gets a zero in the second column of the first row, by using the second row. Since we have the identity matrix on the left, we are done.

The solution to the linear system is $x = -1, y = 1$ and $z = 3$.

Although Gaussian elimination seems straight forward, often it isn't. There are cases where the procedure breaks down. Let's take a look at a couple of cases.

Example 1.3.2 Solve

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 5, \\ x_1 + 4x_3 &= -3, \\ 3x_1 + x_2 + 3x_3 &= 6. \end{aligned}$$

Answer:

Write the linear system as an augmented matrix.
$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 0 & 4 & -3 \\ 3 & 1 & 3 & 6 \end{array}\right),$$

$$S_{21}, \quad \left(\begin{array}{ccc|c} 1 & 0 & 4 & -3 \\ 2 & 1 & -1 & 5 \\ 3 & 1 & 3 & 6 \end{array}\right),$$

$$\begin{array}{rcl} -2R_1 + R_2 & \to & R_2, \\ -3R_1 + R_3 & \to & R_3, \end{array} \quad \left(\begin{array}{ccc|c} 1 & 0 & 4 & -3 \\ 0 & 1 & -9 & 11 \\ 0 & 1 & -9 & 15 \end{array}\right),$$

$$-R_2 + R_3 \to R_3, \quad \left(\begin{array}{ccc|c} 1 & 0 & 4 & -3 \\ 0 & 1 & -9 & 11 \\ 0 & 0 & 0 & 4 \end{array}\right).$$

Remember that our goal is to get 1's on the diagonal. However, it is impossible to get a 1 in the third column of the third row. So what do we do? Is there a solution? Remember that the matrix represents a linear system and we can return to the linear system to try to understand this problem. The last matrix can be written as equations as

$$\begin{aligned} x_1 + 4x_3 &= -3, \\ x_2 - 9x_3 &= 11, \\ 0 &= 4. \end{aligned}$$

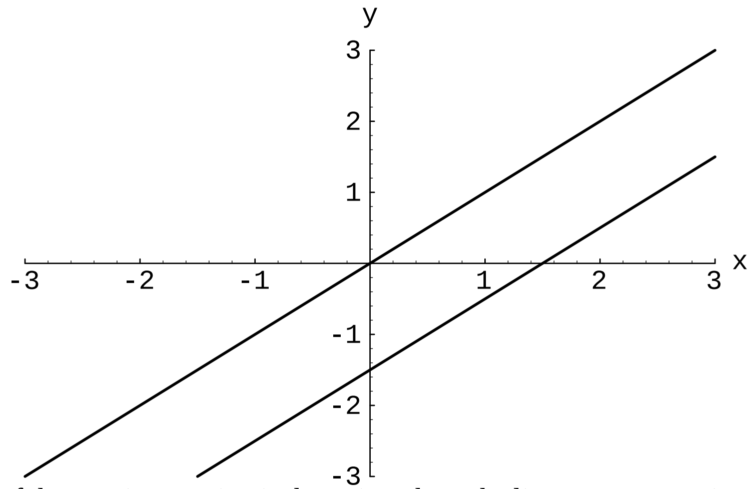
Immediately, you should begin to worry about the last equation, $0 = 4$. Because there are no values of x_1, x_2, and x_3 that can satisfy the last equation, there cannot be a solution to this system. We say that the linear system is **inconsistent**

The following is another example approaching this from a geometric point of view. Solve

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 2 & -2 & 3 \end{array}\right).$$

With one step (multiplying the first row by -2 and adding to row 2) we can reduce the matrix to

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 3 \end{array}\right).$$



The solution of the matrix equation is the point where the lines cross. But since the lines are parallel (each has a slope of 1), there is no solution. This agrees with the results of Gaussian elimination.

In this example, the last equation is $0 = 3$, which presents the same problem as the last example. Let's look at the original matrix in equation form:

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

This system has only two variables, and a two variable linear equation represents a line in the plane. The graphs of the two equations are:

The following example shows that there is a possibility that the solution is **not unique**.

Example 1.3.3 Solve

$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 0, \\ 5x_1 - 3x_2 + x_3 &= 6, \\ 3x_1 - 4x_2 + 4x_3 &= 6. \end{aligned}$$

Answer:

Write the linear system as an augmented matrix.
$$\left(\begin{array}{ccc|c} 2 & 1 & -3 & 0 \\ 5 & -3 & 1 & 6 \\ 3 & -4 & 4 & 6 \end{array}\right),$$

$$\frac{1}{2}R_1 \to R_1, \quad \left(\begin{array}{ccc|c} 1 & 1/2 & -3/2 & 0 \\ 5 & -3 & 1 & 6 \\ 3 & -4 & 4 & 6 \end{array}\right),$$

$$\begin{array}{rcl} -5R_1 + R_2 & \to & R_2, \\ -3R_1 + R_3 & \to & R_3, \end{array} \quad \left(\begin{array}{ccc|c} 1 & 1/2 & -3/2 & 0 \\ 0 & -11/2 & 17/2 & 6 \\ 0 & -11/2 & 17/2 & 6 \end{array}\right),$$

$$-R_2 + R_3 \to R_3, \quad \left(\begin{array}{ccc|c} 1 & 1/2 & -3/2 & 0 \\ 0 & -11/2 & 17/2 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

This situation is similar to the last one at first glance. It is impossible to get 1's on the diagonal. However, this differs in a key way: the last equation is $0 = 0$. This **isn't** the same as the previous example. Let's look at the last matrix in equation form (the last row of zeros doesn't give any additional information, so it isn't included):

$$\begin{array}{rcl} x_1 + \frac{1}{2}x_2 - \frac{3}{2}x_3 & = & 0, \\ -\frac{11}{2}x_2 + \frac{17}{2}x_3 & = & 6. \end{array}$$

The first thing you should notice is that there are three variables and only two equations. Therefore, we cannot get a unique solution as in example 1.3.1. We can, however, solve for x_1 and x_2 in terms of x_3.

First let $x_3 = t$, where t is any real number. Then we can solve for x_1 and x_2 in terms of t.

$$\begin{array}{rcl} x_1 & = & -\frac{25}{11}t + \frac{3}{11}, \\ x_2 & = & \frac{17}{11}t - \frac{6}{11}. \end{array}$$

What kind of geometric object is the solution? Think about what the original equations and the solution represent geometrically.

The last example shows that the numbers of variables and equations do not need to be equal.

Example 1.3.4 Solve

$$\begin{array}{rcl} x + y & = & 6, \\ 2x - y & = & 3, \\ 7x - 2y & = & 15. \end{array}$$

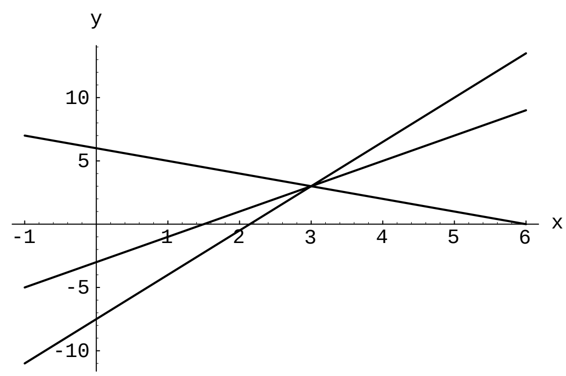
Answer:

$$\left(\begin{array}{cc|c} 1 & 1 & 6 \\ 2 & -1 & 3 \\ 7 & -2 & 15 \end{array}\right) \rightarrow \quad \begin{array}{rcl} -2R_1 + R_2 & \rightarrow & R_2, \\ -7R_1 + R_3 & \rightarrow & R_3, \end{array} \quad \left(\begin{array}{cc|c} 1 & 1 & 6 \\ 0 & -3 & -9 \\ 0 & -9 & -27 \end{array}\right),$$

$$\begin{array}{rcl} -\frac{1}{3}R_2 & \rightarrow & R_2, \\ -3R_2 + R_3 & \rightarrow & R_3, \end{array} \quad \left(\begin{array}{cc|c} 1 & 1 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right),$$

$$-R_2 + R_1 \rightarrow R_1, \quad \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right).$$

The solution to this is $x = 3$, $y = 3$. The geometric interpretation of this problem is that there are three lines crossing in the plane at a point.



1.4 Homogeneous and Non-homogeneous Systems

In example 1.3.1, we presented the system

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

which can be expressed as $\mathbf{A}\mathbf{x} = \mathbf{b}$. Since $\mathbf{b} \neq (0,0,0)^T$, this is classified as a **non-homogeneous system**. As you probably have already guessed, when $\mathbf{b} = (0,0,0)^T$, it is a **homogeneous system**. This property has nothing to do with the solvability of a system.

Example 1.4.1 Solve the corresponding homogeneous system of example 1.3.1,

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

Answer: We can solve the augmented matrix,

$$\left(\begin{array}{rrr|r} 2 & -1 & -1 & 0 \\ 1 & 1 & 2 & 0 \\ -3 & 0 & 2 & 0 \end{array}\right),$$

using the same steps as in example 1.3.1. The resulting augmented matrix is

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

(Try it!) The solution to the linear system is $x = 0$, $y = 0$, and $z = 0$.

Notice how the vector on the right side of the augmented matrix never changed. This is true for all homogeneous systems. Now we are going to present a system that does not have a unique solution.

Example 1.4.2 Using the matrix **A** from example 1.3.3, solve

$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 0, \\ 5x_1 - 3x_2 + x_3 &= 0, \\ 3x_1 - 4x_2 + 4x_3 &= 0. \end{aligned}$$

Answer: Again, we start by forming the augmented matrix

$$\left(\begin{array}{rrr|r} 2 & 1 & -3 & 0 \\ 5 & -3 & 1 & 0 \\ 3 & -4 & 4 & 0 \end{array}\right)$$

and follow the same steps as in example 1.3.3. That results in

$$\left(\begin{array}{ccr|c} 1 & 0 & -8/11 & 0 \\ 0 & 1 & -17/11 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

If we choose $x_3 = c$, the solution is $x_1 = (8/11)c$, $x_2 = (17/11)c$, and $x_3 = c$, where c is a constant. There are infinitely many solutions.

Some people prefer not to write down the zero vector on the right side of the augmented matrix throughout the Gaussian elimination because it does not change. This is perfectly acceptable.

1.5 Problems

1. This is an extension of example 1.1.1. If you also know that James bought 2 CD's at \$11.99 and 10 total CD's for \$108.91, write down the linear system describing his purchase and then determine how many of each CD he bought.

2. (a) For the matrices **A**, **B**, and **C** given by

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 0 & -2 \\ 1 & 1 \\ -2 & 0 \end{pmatrix}$$

Compute the products $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{AC}$, $\mathbf{CA}$, $\mathbf{BC}$, and $\mathbf{CB}$ if they exist, otherwise explain why not.

(b) For arbitrary matrices, why is the equation $\mathbf{AB} = \mathbf{BA}$ false? Discuss.

3. Solve

$$\begin{aligned} 3x_1 - 2x_2 + x_3 &= 6, \\ x_1 + 7x_2 - x_3 &= -2, \\ 6x_1 - 5x_2 - 4x_3 &= -6, \end{aligned}$$

and show that the solution is correct by substitution back into the equations.

4. Solve

$$\begin{aligned} 5x - 15y - 25z &= 0, \\ -2x + 3y + 10z &= 15, \\ 6x - 9y + 5z &= 20, \end{aligned}$$

and show that the solution is correct by substitution back into the equations.

5. Solve

$$\begin{aligned} x + 2y + 8z &= 1, \\ 2x + 4y + 16z &= 2, \\ 4x + 8y + 32z &= 4, \end{aligned}$$

and show that the solution is correct by substitution back into the equations.

6. Solve

$$\begin{aligned} 3x + 5y &= 13, \\ 2x - 3y &= -4, \\ 5x + 2y &= 9, \end{aligned}$$

and show that the solution is correct by substitution back into the equations.

7. Solve

$$\begin{aligned} 8x_1 + 5x_2 - 4x_3 &= 1, \\ 3x_1 + 2x_2 - x_3 &= 5, \end{aligned}$$

and show that the solution is correct by substitution back into the equations.

8. What kind of geometric objects are the solutions in problems 1–4?

9. Solve

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

where a, b, c, d, e, f are constants. Are there any restrictions on the constants to ensure a solution? Is the solution unique?

10. What if instead we have the homogeneous system

$$\begin{aligned} ax + by &= 0, \\ cx + dy &= 0, \end{aligned}$$

where $a, b, c,$ and d are constants. What condition is needed for the constants to ensure a solution? Is the solution unique?

11. Solve the system given in example 1.1.2. Recall that the system came from finding temperatures based on the average of its neighbors. Is your solution correct? Explain.

12. We can solve the same linear system with different right hand sides by making a larger augmented matrix with an extra columns for each non-homogeneous vectors. For instance, we can write the two linear systems:

$$\begin{aligned} 3x + 9y - 6z &= -24, & \quad 3x + 9y - 6z &= -21, \\ -2x + y - 3z &= -5, & \quad -2x + y - 3z &= -14, \\ 5x + 2y - z &= 7, & \quad 5x + 2y - z &= -7, \end{aligned}$$

as the following augmented matrix:

$$\left(\begin{array}{ccc|cc} 3 & 9 & -6 & -24 & -21 \\ -2 & 1 & -3 & -5 & -14 \\ 5 & 2 & -1 & 7 & -7 \end{array} \right).$$

Use Gaussian elimination and back substitution to solve the above augmented matrix.

The matrix form of this system can be written

$$\mathbf{AX} = \mathbf{B},$$

where **X** and **B** are 3 by 2 matrices. The **X** matrix is the unknown matrix and the **B** matrix is the non-homogeneous matrix. Your solution should be written as a 3 by 2 matrix.

13. Using the method of problem 12, solve the following three systems simultaneously.

$$\begin{aligned} x + 2y + 2z &= 1, & \quad x + 2y + 2z &= 0, & \quad x + 2y + 2z &= 0, \\ 2x + y + 2z &= 0, & \quad 2x + y + 2z &= 1, & \quad 2x + y + 2z &= 0, \\ 2x + 2y + z &= 0, & \quad 2x + 2y + z &= 0, & \quad 2x + y + 2z &= 1. \end{aligned}$$

You solution should be three three-dimensional vectors. You can write this as a 3 by 3 matrix. The matrix found in this problem is special. It is called the **inverse** matrix and we will talk more about it in the next chapter.

Similar to the above problem, the matrix equation that you will solve has the form

$$\mathbf{AX} = \mathbf{I}.$$

Your solution should be the 3 by 3 matrix **X**.

14. Suppose we had the homogeneous system

$$\begin{aligned}(1-\lambda)x + 2y &= 0,\\ 2x + (1-\lambda)y &= 0,\end{aligned}$$

and we want a solution other than the trivial solution $(x, y)^T = (0, 0)^T$. What value of λ satisfies that condition? (What is the condition on the last row so that the solution for y does not equal 0?) Compare this to problem 10.

Hint: Write the system in augmented matrix form and perform the following first step, $\frac{R_1}{\lambda-1} \to R_1$.

15. Provide a mathematical reason for why the Gaussian elimination steps (as in the box on page 10) are legal operations.

Chapter 2

Operations on Matrices

In the last chapter you learned about the connection between linear systems (collections of linear equations) and matrices, how to write a system in augmented matrix form, and how to solve such a system using Gaussian elimination.

Matrices can arise in other ways as well, for example as an array of numbers that don't have anything to do with linear systems. For example, the array of test scores for each student in a class is a matrix. In this chapter, we will introduce some new operations on matrices, the determinant and inverse. We will also see how they are connected to the problem of solving linear systems.

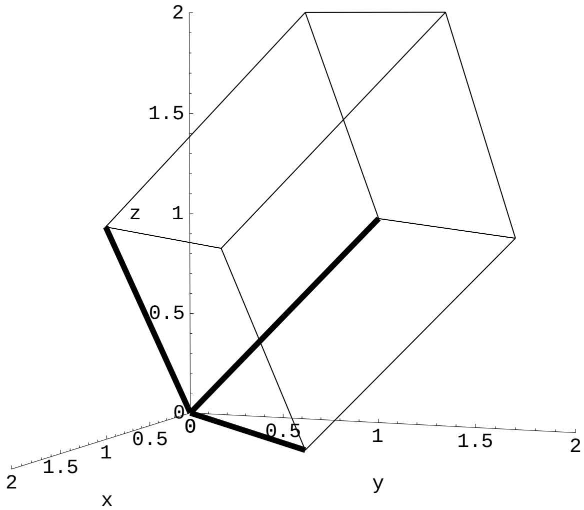
2.1 Determinants

This section gives a brief overview of how to compute determinants. The rules that we will give for computing determinants can be derived, and you can find out how by taking a course in matrix methods, such as APPM 3310.

One definition of *determinant* comes from geometry: a determinant represents the volume of a parallelipiped spanned by the column vectors of a square matrix (it is important that we can define the *determinant* only for square matrices). For example in three dimensions, the triple product $\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})$ represents the volume of the parallelepiped with edges given by the vectors $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$. If the matrix $\mathbf{A}$ is written with $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ in the three columns, with $\mathbf{x} = (x_1, x_2, x_3)^T$, $\mathbf{y} = (y_1, y_2, y_3)^T$ and $\mathbf{z} = (z_1, z_2, z_3)^T$ then the triple product $\mathbf{x} \times \mathbf{y} \cdot \mathbf{z}$ is equal to the determinant of $\mathbf{A}$, denoted det($\mathbf{A}$) or written with vertical bars around the matrix.

$$\det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Example 2.1.1 Write the volume of the parallelepiped given by the vectors $\mathbf{x} = (1, 0, 1)$, $\mathbf{y} = (1, 1, 0)$, and $\mathbf{z} = (0, 1, 1)$ as a determinant.



Answer:

$$\text{Volume of parallelepiped} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix}.$$

We will now learn how to compute this and other determinants. There are numerous other applications of the determinant. A few of them will be shown below, and others you will encounter later on. Our main application will be to determine the solvability of linear systems.

2.1.1 One and Two Dimensional Determinants

Suppose we have a single linear equation in one variable, $ax = b$. We know that we can solve this equation for x whenever $a \neq 0$. The generalization of this condition for solvability to linear systems with n equations in n variables leads to the criterion that the system $\mathbf{Ax} = \mathbf{b}$ is solvable whenever $\det(\mathbf{A}) \neq 0$. Our goal here is to give the rules for the computation of this quantity. In particular, we have just seen that for a 1×1 matrix:

$$\mathbf{A} = (a_{11}), \qquad \det(\mathbf{A}) = a_{11}.$$

If you worked problem 9 in section 1.5, you obtained the condition for solvability of a 2×2 system. This leads to the definition of the determinant of a 2×2 matrix:

For a matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \det(\mathbf{A}) = a_{11}a_{22} - a_{21}a_{12}.$$

This can be easily remembered by multiplying the main diagonal elements and subtracting the product of the off diagonal pieces. This is an important formula to know, because it is needed to compute the determinants of larger matrices. In two dimensions, volume doesn't make sense. Instead, the determinant of a two by two matrix represents the area of the parallelipiped spanned by the two column vectors.

The conditions for solvability of larger systems are most easily derived by breaking the system into smaller pieces. This leads to the "cofactor" method for computing determinants.

2.1.2 The Cofactor Method of Calculating Determinants

Although there are formulas for the determinants of larger matrices, they get complex very quickly. Instead of memorizing formulas, there are properties of matrices that allow easy computation. The cofactor method is a brute force method for obtaining determinants: it is guaranteed to work, but it is not the easiest method to use. We will show in the next subsection that row operations lead to easier methods. However, it sometimes best to use cofactors (for example, when you compute eigenvalues, as you will in Chapter 3).

If a square matrix $\mathbf{A}$ is represented by its elements a_{ij}, then

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{C}_{ij}), \quad \text{for any row } i.$$

where $\mathbf{C}_{ij}$ is the **minor** of element i, j. The minor matrix is found by removing the i^{th} row and j^{th} column of the matrix $\mathbf{A}$.

Example 2.1.2 Find the determinant of

$$\begin{pmatrix} 1 & 1 & 6 \\ 4 & -3 & -2 \\ 2 & 3 & 1 \end{pmatrix}.$$

Answer: Using the properties of the the cofactors, we can expand the determinant along the top row:

$$\begin{aligned} \begin{vmatrix} \boxed{1 \quad 1 \quad 6} \\ 4 \quad -3 \quad -2 \\ 2 \quad 3 \quad 1 \end{vmatrix} &= 1(-1)^{1+1} \begin{vmatrix} -3 & -2 \\ 3 & 1 \end{vmatrix} + 1(-1)^{1+2} \begin{vmatrix} 4 & -2 \\ 2 & 1 \end{vmatrix} \\ &\quad +6(-1)^{1+3} \begin{vmatrix} 4 & -3 \\ 2 & 3 \end{vmatrix}, \\ &= 1(3) + (-1)(8) + 6(18) = 103. \end{aligned}$$

Example 2.1.3 Find the determinant of the matrix in example 1.1.2,

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

Answer: If we expand across the bottom row, the first two elements are zero, so the determinant reduces to two three by three matrices.

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

$$= \left(2\begin{vmatrix} 2 & 0 \\ -1 & -1 \end{vmatrix} + 1\begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}\right)$$

$$+2\left(2\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} + \begin{vmatrix} -1 & -1 \\ 0 & 2 \end{vmatrix}\right)$$

$$= 2(-2) + (1) + 2(2(3) + (-2)) = 5$$

2.1.3 Properties of the Determinant

Elementary row operations can be used to simplify the computation of a determinant. The reason is the following important property of determinant (reminder–only square matrices have determinants):

> 0. If $\mathbf{A}$ is upper or lower triangular, then the determinant is the product of the diagonal elements.

As we have seen, elementary row operations can be used to transform a matrix to upper triangular form. Thus if we know how these operations change the determinant, we can use them to compute it.

> 1. Under the swap operation, S_{ab}, the determinant changes sign, i.e., if two rows of a matrix $\mathbf{A}$ are swapped to give a new matrix $\mathbf{B}$ then det$(\mathbf{A}) = -$det$(\mathbf{B})$.
> 2. Under the operation $cR_a \to R_a$, the determinant is multiplied by c. Alternatively, if a scalar, c is *factored out* of a row of $\mathbf{A}$ leaving a matrix $\mathbf{B}$, then det($\mathbf{A}$)=c det$(\mathbf{B})$.
> 3. The operation $cR_a + R_b \to R_b$ leaves the determinant unchanged.

These properties are useful in computing the determinant in an alternate way which is often easier than cofactors. There are a couple of other interesting, but slightly less useful properties of determinants.

- If the matrix $\mathbf{A}$ has a row of zeros then det($\mathbf{A}$)=0.
- If the matrix $\mathbf{A}$ has two identical rows then det($\mathbf{A}$)=0.
- If any two matrices, $\mathbf{A}$ and $\mathbf{B}$ are the same size then det($\mathbf{AB}$)=det($\mathbf{A}$)det($\mathbf{B}$).
- The transpose operation doesn't change the determinant. det($\mathbf{A}$)=det($\mathbf{A}^T$).

Can you derive these from what you know so far? (Hint: the third property is hard, and the first two follow from row operations).

We now try a few examples.

Example 2.1.4 Find the determinant of

$$\begin{pmatrix} 1 & 1 & 6 \\ 4 & -3 & -2 \\ 2 & 2 & 1 \end{pmatrix}.$$

Answer:
To compute this matrix using row reduction, our goal is to obtain an upper (or lower) triangular system and then calculate the determinant using property 0. The first step is to row reduce the matrix in a Gaussian elimination sense. We use property 3 to accomplish this.

$$\begin{vmatrix} 1 & 1 & 6 \\ 4 & -3 & -2 \\ 2 & 2 & 1 \end{vmatrix} = \begin{matrix} -4R_1 + R_2 \to R_2, \\ -2R_1 + R_3 \to R_3, \end{matrix} \begin{vmatrix} 1 & 1 & 6 \\ 0 & -7 & -26 \\ 0 & 0 & -11 \end{vmatrix}.$$

The two row operations reduced it to an upper triangular matrix without changing the determinant (property 3). By reading off the diagonal elements (property 0), the determinant is (1)(-7)(-11)=77, the same answer that we found in example 2.1.2, and in this case much less work.

Example 2.1.5 Find the determinant of

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

Answer:

$$S_{12}, \quad -1\begin{vmatrix} 3 & 0 & 6 \\ 2 & -1 & -7 \\ -5 & -4 & 2 \end{vmatrix}.$$

The first and second rows are swapped (property 1). This is why a -1 appears out front.

$$\frac{1}{3}R_1 \to R_1, \quad -3\begin{vmatrix} 1 & 0 & 2 \\ 2 & -1 & -7 \\ -5 & -4 & 2 \end{vmatrix}.$$

We factor out a 3 from the first row of the matrix in order to get a row with a 1 in the first column (property 2), The factor of 3 and the previous step is why a -3 appears out front.

$$\begin{matrix} -2R_1 + R_2 & \to & R_2, \\ 5R_1 + R_3 & \to & R_3, \end{matrix} \quad -3\begin{vmatrix} 1 & 0 & 2 \\ 0 & -1 & -11 \\ 0 & -4 & 12 \end{vmatrix}.$$

These two steps (property 3) are used to get zeros in the first column.

$$-4R_2 + R_3 \to R_3, \quad -3\begin{vmatrix} 1 & 0 & 2 \\ 0 & -1 & -11 \\ 0 & 0 & 56 \end{vmatrix}.$$

This step zeroed out the second column of the third row leaving it in upper triangular form. Now, using property 0, we can read off the determinant.

$$(-3)(1)(-1)(56) = -168.$$

Notice that the steps involved are precisely that of Gaussian elimination (without back substitution). However, unlike Gaussian elimination, to compute determinants we have to keep track of the rows that we swap, since each gives a factor of -1, and any constants that we multiply by, since these also change the determinant.

2.2 Inverses

When you studied high school algebra, you learned how to solve simple equations like $2x = 6$. This was done by dividing both sides by 2 to get $x = 3$. Why is it that we have to do Gaussian elimination to get the solution to a linear system? If we have a matrix equation $\mathbf{Ax} = \mathbf{b}$, why isn't $\mathbf{x} = \mathbf{b}/\mathbf{A}$?

The answer lies in the difference between numbers and matrices. You learned a long time ago how to divide numbers, but there isn't a division operation for matrices.

However, we can look at the equation $2x = 6$ in a different light. We can solve this by finding the reciprocal of 2 and multiplying by 6 to get $x = \frac{1}{2}(6) = 3$. The factor $\frac{1}{2}$ can be thought of as the inverse of the number 2.

This is how we can treat matrices. Let's assume that there is a matrix $\mathbf{B}$ that has the form $\mathbf{x} = \mathbf{Bb}$, where $\mathbf{x}$ is the solution to $\mathbf{Ax} = \mathbf{b}$. What can we find out about $\mathbf{B}$? If we multiply (because we do know how to multiply matrices) $\mathbf{x} = \mathbf{Bb}$ by the matrix $\mathbf{A}$, we get $\mathbf{Ax} = \mathbf{ABb}$.

So we have two equations

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} \qquad &\text{and} \qquad \mathbf{Ax} = \mathbf{ABb}, \\ \Rightarrow \quad \mathbf{b} &= \mathbf{ABb}, \\ \Rightarrow \quad \mathbf{I} &= \mathbf{AB}. \end{aligned}$$

This means that the matrix $\mathbf{B}$ that we are looking for multiplies $\mathbf{A}$ to give the identity matrix. We call $\mathbf{B}$ the **inverse** of $\mathbf{A}$ and denote it by $\mathbf{A}^{-1}$.

2.2.1 Calculating the Inverse matrix

We want to know how to find $\mathbf{A}^{-1}$. The only property that we know about the inverse is that $\mathbf{AA}^{-1} = \mathbf{I}$. Recall that in problems 8 and 9 of chapter 1, we learned how to solve systems simultaneously for different right hand sides. Take a few minutes to look back at these problems.

Example 2.2.1 Find $\mathbf{A}^{-1}$ when $\mathbf{AA}^{-1} = \mathbf{I}$ and

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

Answer:

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 3 & 1 & 0 & 0 \\ -2 & 1 & -5 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right).$$

The matrix equation $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ is rewritten as an augmented matrix.

$$\begin{array}{rcl} 2R_1 + R_2 & \to & R_2, \\ -3R_1 + R_3 & \to & R_3, \end{array} \qquad \left(\begin{array}{rrr|rrr} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 & 0 \\ 0 & 1 & -8 & -3 & 0 & 1 \end{array}\right).$$

These row operations get a 0 in the first column of the second and third rows.

$$-R_2 + R_3 \to R_3, \qquad \left(\begin{array}{rrr|rrr} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 & 0 \\ 0 & 0 & -9 & -5 & -1 & 1 \end{array}\right).$$

The matrix is now in upper triangular form.

$$-\frac{1}{9}R_3 \to R_3, \qquad \left(\begin{array}{rrr|rrr} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 5/9 & 1/9 & -1/9 \end{array}\right).$$

Now that the matrix on the matrix has ones on the diagonal, we work to get zeros above the diagonal.

$$\begin{array}{rcl} -R_3 + R_2 & \to & R_2, \\ -3R_3 + R_2 & \to & R_1, \end{array} \qquad \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -2/3 & -1/3 & 1/3 \\ 0 & 1 & 0 & 13/9 & 8/9 & 1/9 \\ 0 & 0 & 1 & 5/9 & 1/9 & -1/9 \end{array}\right).$$

The last two row operations give the identity matrix on the left side, therefore the right side is the inverse matrix.

Notice that the coefficients of the inverse matrix are not the inverses of the coefficients of the matrix, i.e.,

$$a_{ij}^{-1} \neq \frac{1}{a_{ij}}.$$

Example 2.2.2 Solve the system of equations given in example 1.3.1,

$$\begin{array}{rcl} 2x - y - z & = & 0, \\ x + y + 2z & = & 6, \\ -3x + 2z & = & 3, \end{array}$$

by using the inverse of the coefficient matrix.

Answer:

$$\left(\begin{array}{rrr|rrr} 2 & -1 & -1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ -3 & 0 & 2 & 0 & 0 & 1 \end{array}\right),$$

We rewrite the matrix equation $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ as an augmented matrix

$$S_{21}, \quad \left(\begin{array}{rrr|rrr} 1 & 1 & 2 & 0 & 1 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ -3 & 0 & 2 & 0 & 0 & 1 \end{array}\right),$$

Swap rows one and two to get a 1 in the upper left corner

$$\begin{array}{rcl} -2R_1 + R_2 & \to & R_2, \\ 3R_1 + R_3 & \to & R_3, \end{array} \quad \left(\begin{array}{rrr|rrr} 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -5 & 1 & -2 & 0 \\ 0 & 3 & 8 & 0 & 3 & 1 \end{array}\right),$$

These row operations get a 0 in the first column of the second and third rows.

$$R_2 + R_3 \to R_3, \quad \left(\begin{array}{rrr|rrr} 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -5 & 1 & -2 & 0 \\ 0 & 0 & 3 & 1 & 1 & 1 \end{array}\right),$$

These row operations get a 0 in the first column of the second and third rows.

$$\begin{array}{rcl} -\frac{1}{3}R_2 & \to & R_2, \\ \frac{1}{3}R_3 & \to & R_3, \end{array} \quad \left(\begin{array}{rrr|rrr} 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 5/3 & -1/3 & 2/3 & 0 \\ 0 & 0 & 1 & 1/3 & 1/3 & 1/3 \end{array}\right),$$

The matrix is now in upper triangular form.

$$\begin{array}{rcl} -\frac{3}{5}R_3 + R_2 & \to & R_2, \\ -2R_3 + R_1 & \to & R_1, \end{array} \quad \left(\begin{array}{rrr|rrr} 1 & 1 & 0 & -2/3 & 1/3 & -2/3 \\ 0 & 1 & 0 & -8/9 & 1/9 & -5/9 \\ 0 & 0 & 1 & 1/3 & 1/3 & 1/3 \end{array}\right),$$

Now that the matrix on the matrix has ones on the diagonal, we work to get zeros above the diagonal.

$$-R_2 + R_1 \to R_1, \quad \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 2/9 & 2/9 & -1/9 \\ 0 & 1 & 0 & -8/9 & 1/9 & -5/9 \\ 0 & 0 & 1 & 1/3 & 1/3 & 1/3 \end{array}\right),$$

The last two row operations give the identity matrix on the left side, therefore the right side is the inverse matrix.

From the steps above we know that the inverse matrix is

$$\mathbf{A}^{-1} = \left(\begin{array}{rrr} 2/9 & 2/9 & -1/9 \\ -8/9 & 1/9 & -5/9 \\ 1/3 & 1/3 & 1/3 \end{array}\right).$$

Therefore the solution to the linear system can be found by the matrix-vector multiplication $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

$$\left(\begin{array}{rrr} 2/9 & 2/9 & -1/9 \\ -8/9 & 1/9 & -5/9 \\ 1/3 & 1/3 & 1/3 \end{array}\right)\left(\begin{array}{r} 0 \\ 6 \\ 3 \end{array}\right) = \left(\begin{array}{r} 1 \\ -3 \\ 3 \end{array}\right)$$

This solution, $x = 1, y = -3$, and $z = 3$ agrees with that we found in example 1.3.1.

The above method shows an alternative method to solving linear systems using Gaussian elimination. You should compare this example with example 1.3.1. Both methods involve using Gaussian elimination, but the above method involves first computing the inverse and then computing matrix-vector multiplication, whereas using only Gaussian elimination as in example 1.3.1 takes fewer steps.

There are other methods that are used to solve linear systems, including one that uses determinants. However, when the linear systems get large Gaussian elimination is still one of the most efficient (fewer steps). Many computer routines that solve linear systems are based on Gaussian elimination.

2.2.2 What's the relationship between inverses and determinants?

In problem 6 of this chapter, you are asked to find the determinant of

$$\begin{pmatrix} 2 & 2 & -1 \\ 4 & 0 & 7 \\ -3 & -1 & -3 \end{pmatrix}.$$

If you have not yet done this problem, you should do so. Here we will compute the inverse of this matrix.

$$\begin{array}{rcl} -2R_1 + R_2 & \to & R_2, \\ \frac{3}{2}R_1 + R_3 & \to & R_3, \end{array} \qquad \left(\begin{array}{ccc|ccc} 2 & 2 & -1 & 1 & 0 & 0 \\ 0 & -4 & 9 & -2 & 1 & 0 \\ 0 & 2 & -9/2 & 3/2 & 0 & 1 \end{array}\right),$$

$$\frac{1}{2}R_2 + R_3 \to R_3, \quad \left(\begin{array}{ccc|ccc} 2 & 2 & -1 & 1 & 0 & 0 \\ 0 & -4 & 9 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1/2 & -1/2 & 1 \end{array}\right).$$

As you can see, the last operation yield three zeros in the last row, making it impossible to solve the augmented matrix. Therefore, there is no inverse for this matrix. The determinant for this matrix is zero. This is an example of an important theorem.

Theorem 2.1 *If the determinant of a matrix A is 0, then the inverse of A does not exist. Conversely, if a matrix A has no inverse then its determinant is zero.*

How can we use this theorem? Now you know two ways to decide whether or not a matrix is invertible. The first is taking the determinant of the matrix and seeing if it is zero. The second involves trying to find the inverse using Gaussian elimination. Sometimes one method will be easier than another, so it is always good to know more than one method. This theorem will also be extended in the next section.

2.3 The Existence and Uniqueness Theorem for Linear Systems

This section discusses the existence and uniqueness of solutions of linear systems. It is important to know when a solution will exist. For instance, we would like to have a way of determining whether or not there is a solution without using Gaussian elimination. If we know that a solution exists, we would like to know if it was unique (had only one solution).

We will consider three cases of the linear system $\mathbf{Ax} = \mathbf{b}$, consisting of m equations and n unknowns: $m > n$, $m < n$, and $m = n$. We also discuss the existence and uniqueness of solution of both homogeneous and non-homogeneous systems.

2.3.1 Case 1: $m > n$

Consider the non-homogeneous system $\mathbf{Ax} = \mathbf{b}$ of m equations and n unknown. If $\mathbf{A}$ can be reduced by Gaussian elimination to the form of an n by n identity matrix followed below by $m - n$ rows of zeros, then the system $\mathbf{Ax} = \mathbf{b}$ has a unique solution.

Example 2.3.1 Explain the system in example 1.3.4,

$$\begin{aligned} x + y &= 6, \\ 2x - y &= 3, \\ 7x - 2y &= 15, \end{aligned}$$

using the statement above.

Answer: The augmented matrix on the left was reduced to the augmented matrix on the right of

$$\left(\begin{array}{cc|c} 1 & 1 & 6 \\ 2 & -1 & 3 \\ 7 & -2 & 15 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right).$$

where the $\sim$ means that Gaussian elimination has been done.
In this example, the 3 by 2 matrix reduced to a 2 by 2 identity matrix and $3 - 2 = 1$ row of zeroes. Therefore, we know that there is a unique solution to this problem.

This doesn't mean that every non-homogeneous system where $m > n$ has a unique solution. This is true only if the left hand side of the augmented matrix can be reduced to an n by n identity matrix. Their are two other possibilities: the system of equations may have no solutions (i.e., the system is *inconsistent*), or an infinite number of solutions. There is usually no easy way to check, but the results of Gaussian elimination will let you know.

2.3.2 Case 2: $m < n$

When $m < n$, there are fewer equations than unknowns and if any solution exists, then there are an infinite number of solutions.

Example 2.3.2 Does the system from Problem 7 in Section 1.3,

$$\begin{aligned} 8x_1 + 5x_2 - 4x_3 &= 1 \\ 3x_1 + 2x_2 - x_3 &= 5 \end{aligned}$$

have a solution? If so, is it unique?

Answer: There not enough equations to determine all the variables; however, it is still possible that the system is inconsistent (and has no solution). However, whe we reduce the augmented matrix we find

$$\left(\begin{array}{ccc|c} 8 & 5 & -4 & 1 \\ 3 & 2 & -1 & 5 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -3 & -23 \\ 0 & 1 & 4 & 37 \end{array}\right).$$

This 2 by 3 system is consistent, and has infinitely many solutions.

Can you make an example of a 2×3 system that is not consistent?

2.3.3 Case 3: $m = n$

Unless otherwise indicated, from now on we will only consider the system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is square. In fact this is the most common case–diligent engineers usually find just as many equations as they have unknowns! Moreover, in this case, we can use the theory of determinants and inverses, since we know how to compute these for the square case. The following theorem gives us guidance.

Theorem 2.2 (Existence and Uniqueness Theorem for Non-homogeneous Systems) *Suppose $\mathbf{A}$ is a square matrix, then*

i) if $\det \mathbf{A} \neq 0$, then the inverse of $\mathbf{A}$ exists and the solution of $\mathbf{Ax} = \mathbf{b}$ is unique.

ii) if the $\det \mathbf{A} = 0$, then $\mathbf{A}^{-1}$ does not exist and $\mathbf{Ax} = \mathbf{b}$ either has infinitely many solutions or no solution.

Example 2.3.3 Does the system of example 1.3.2,

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 5, \\ x_1 + 4x_3 &= -3, \\ 3x_1 + x_2 + 3x_3 &= 6. \end{aligned}$$

have a solution?

Answer: First, we check the determinant.

$$\begin{vmatrix} 2 & 1 & -1 \\ 1 & 0 & 4 \\ 3 & 1 & 3 \end{vmatrix} = 0.$$

You should verify this.

Because the determinant is 0, we are assured that a unique solution does not exist, but we still don't know whether there are infinitely many or no solutions. Sometimes, checking does create a bit of extra work.

Example 2.3.4 Does the system in (Eq. 1.4),

$$\begin{aligned} x + y + 6z &= 17, \\ 4x - 3y - 2z &= 4, \\ 2x + 3y + z &= -1, \end{aligned}$$

have a solution?

Answer: Checking the determinant,

$$\begin{vmatrix} 1 & 1 & 6 \\ 4 & -3 & -2 \\ 2 & 3 & 1 \end{vmatrix} = 103 \neq 0$$

shows that since the determinant is nonzero, the solution will have a unique solution. We found the solution in section 1.3

To summarize the results of the theorem, the following statements are equivalent:

- $\det \mathbf{A} \neq 0$,
- $\mathbf{A}$ is invertible ($\mathbf{A}^{-1}$ exists),
- $\mathbf{Ax} = \mathbf{b}$ has a unique solution.

2.3.4 Existence and Uniqueness of Homogeneous Systems

Above, a theorem for the existence and uniqueness of a non-homogeneous matrix equation was presented. There is also a theorem for homogeneous matrix equations.

Theorem 2.3 (Existence Theorem for Homogeneous Systems) *If the determinant of a square matrix* $\mathbf{A}$ *is non-zero, then the only solution of the matrix equation* $\mathbf{Ax} = \mathbf{0}$ *is* $\mathbf{x} = \mathbf{0}$. *Conversely if* $(\det \mathbf{A}) = 0$ *then there is a non-trivial* $(\mathbf{x} \neq \mathbf{0})$ *solution of* $\mathbf{Ax} = \mathbf{0}$.

Example 2.3.5 Does the homogeneous version of the system in example 1.3.3,

$$\begin{aligned} 2x_1 + x_2 - 3x_3 &= 0, \\ 5x_1 - 3x_2 + x_3 &= 0, \\ 3x_1 - 4x_2 + 4x_3 &= 0, \end{aligned}$$

have a solution? If so, is it unique?

Answer: The coefficient matrix is

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

The determinant of $\mathbf{A}$ is zero (you should check this), and since the linear system is homogeneous, then using Theorem 2.3, we can say that a non-zero solution exists to this system.

With the theorems in the section, you should be able to discuss the solutions of both homogeneous and non-homogeneous linear systems, without solving the systems. You should know also know whether or not a solution to a system is unique.

2.4 Problems

1. Determine the volume of the parallelepiped with edges $(2, 0, 1)$, $(0, 3, 1)$ and $(-1, -2, 0)$.
2. Using both the cofactor and row reduction methods, find the determinant of

$$\begin{pmatrix} 2 & 1 & -1 \\ 6 & 5 & -2 \\ 4 & 3 & 1 \end{pmatrix}.$$

3. Using both the cofactor and row reduction methods, find the determinant of

$$\begin{pmatrix} 2 & 2 & -1 \\ 4 & 0 & 7 \\ -3 & -5 & 1 \end{pmatrix}.$$

4. For each of the properties of the determinant give an example using a 2 by 2 matrix.

5. Using the property $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ and the cofactor method, show that you can extend the cofactor method to reducing determinants by expanding along a column.

6. Compute the inverse in problem 2 if it exists.

7. Compute the inverse in problem 3 if it exists.

8. How can you determine whether or not the inverse is correct? (Hint: what equation did we use to find the inverse? Does the solution satisfy this?) Show that the inverse in example 2.2.1 and in problem 7 is correct.

9. Compute the inverse of

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

   Is there a condition when the inverse cannot be found? Solve problem 9 of chapter 1 using the inverse. Write down a statement that includes the results of problems 9 and 10 of chapter 1 and this problem using theorems 2.1, 2.2, and 2.3.

10. Does the system

$$\begin{pmatrix} 0 & 0 & 0 & 4 \\ 1 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 3 \\ 6 \end{pmatrix}$$

   have a unique solution? Does the inverse of the coefficient matrix exist?
   (Hint: Don't find the solution!)

11. Show by finding a solution, that the system in example 2.3.5 has a non-zero solution. Is it unique?

12. Discuss the conditions for existence and uniqueness of both homogeneous and non-homogeneous linear systems.

Chapter 3

Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors arise naturally from differential equations and are found throughout engineering and the sciences. This chapter discusses eigenvalues and eigenvectors from a geometric point of view. The eigenvalues and eigenvectors of a matrix provide information about the matrix which we will find of great use when we solve differential equations.

The first section examines linear transformations which are matrix-vector products. Understanding these transformations will provide an understanding of eigenvalues and eigenvectors.

3.1 Linear Transformations

In this section, we would like to think of matrices on a geometric level. More specifically, a matrix can multiply a vector to give another vector. We will examine what happens geometrically during such a transformation.

Example 3.1.1 In one-dimension, consider $L(x) = 2x$, where x is any real number. Then $L(1) = 2(1)$, $L(10) = 20, \ldots$. What is the result of this transformation

Answer: .

This is a seemingly stupid example, but this is similar to the Chili recipe example in section 1.1. Actually, there really aren't any interesting linear transformations in one-dimension a transformation will either move a point towards zero, away from zero, or leave it fixed (the identity transformation).

Example 3.1.2 Is $P(x) = \sin x$ a linear transformation?

Answer: No, because the function $\sin x$ is not linear.

3.1.1 Linear transformations in 2D

All linear transformations have the form $L(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$. In this section, we consider transformations from the plane to the plane (we also say from $\mathbf{R}^2$ to $\mathbf{R}^2$), thus the $\mathbf{A}$ matrices are 2 by 2.

Example 3.1.3 If $L(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 0 & 8 \end{pmatrix}$, then what is $L(\mathbf{x})$, where $\mathbf{x} = (1, 2)^T$?

Answer: $(11, 16)^T$? A linear transformation is found simply by matrix-vector multiplication.

Linear transformations can be visualized easily in the plane if we plot both the original vector and the resultant vector.

Example 3.1.4 What happens to the vector $(2, 1)^T$ with the linear transformation given by the matrix

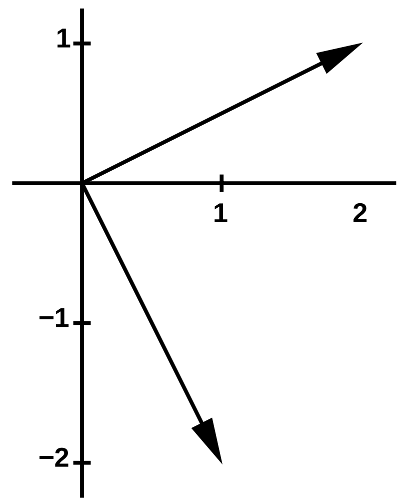
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}?$$

Answer:

The result of the linear transformation is $(1, -2)^T$. This is easy to compute, but it hard to see what is happening.

A graph of the original vector $(2, 1)^T$ and the resultant vector $(1, -2)^T$, will make the form of the transformation more apparent.

Notice that in this example, the original vector has been rotated clockwise by $90°$. The matrix that produced this transformation is part of a larger class of matrices that are called **rotation matrices**. You will investigate them more in Problem 1.



Although every matrix defines a linear transformation, rotation matrices have the important property that they rotate any vector by a given amount. The following example shows another type of linear transformation.

Example 3.1.5 What type of behavior is generated by the linear transformation given by the matrix

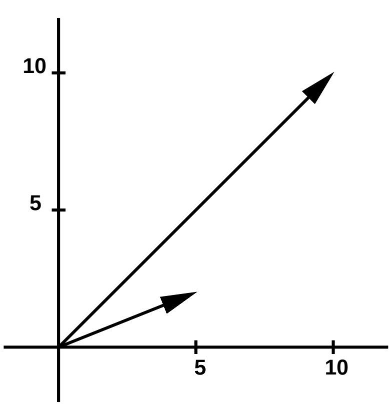
$$\begin{pmatrix} 1/2 & 0 \\ 0 & 1/5 \end{pmatrix}?$$

Answer:

As an example, let's take the vector $(10, 10)^T$ and see what happens under transformation. The resultant vector is $(5, 2)^T$.

The first component is reduced by a half and the second by a fifth.

This matrix is special because it's a diagonal matrix. The element on the diagonal multiplies the first element of the vector. And the second diagonal element of the matrix multiplies the second element of the vector. In this example, both diagonal elements are less than 1. This causes the vector to shrink. An element greater 1 will expand the vector. *What will happen to a vector if a diagonal element is negative? (Try it!)*



Example 3.1.6 Consider the transformation given by the matrix

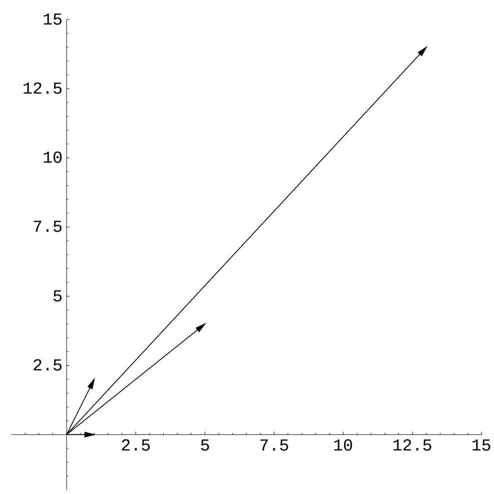
$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

What are the first 4 transformations of the matrix $\mathbf{A}$ with the vector $\mathbf{x}_0 = (1, 0)^T$?

Answer: The first transformation is given by $\mathbf{A}\mathbf{x}_0 = (1, 2)^T$. Let $\mathbf{x}_1 = (1, 2)^T$. We can apply the matrix more times. For the second time $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \mathbf{A}\mathbf{A}\mathbf{x}_0 = \mathbf{A}^2\mathbf{x}_0$. The following table gives the results for the first few applications of the matrix $\mathbf{A}$.

$$\mathbf{x}_2 = \mathbf{A}^2\mathbf{x}_0, \qquad \mathbf{x}_2 = \begin{pmatrix} 5 \\ 4 \end{pmatrix},$$
$$\mathbf{x}_3 = \mathbf{A}^3\mathbf{x}_0, \qquad \mathbf{x}_3 = \begin{pmatrix} 13 \\ 14 \end{pmatrix},$$
$$\mathbf{x}_4 = \mathbf{A}^4\mathbf{x}_0, \qquad \mathbf{x}_4 = \begin{pmatrix} 41 \\ 40 \end{pmatrix}.$$

vector number	length	relative length ($\|\mathbf{x}_{i+1}\|/\|\mathbf{x}_i\|$)
1	2.23	2.23
2	6.40	2.86
3	19.11	2.98
4	57.28	2.998



The figure shows $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_3$, where $\mathbf{x}_0$ is the smallest and $\mathbf{x}_3$ is the largest vector.

It appears that the linear transformation will triple the length of the vector, and the vector is approaching one where both elements are equal, or the vectors approach the $45°$ line. This will be explained in the next section.

3.1.2 Linear Transformations in Higher Dimensions

The same ideas can be applied to higher dimensions. Linear transformations from $\mathbf{R}^3$ to $\mathbf{R}^3$ are represented by 3 by 3 matrices. Among the possibilities are rotations and dilations as before. The major difference is that it is harder to visualize what is going on in 3 dimensions. Of course, if you think this is hard, try 4 or more dimensions!

3.2 Eigenvectors and Eigenvalues

In the previous section, we considered what happens to a typical vector under a linear tranformation. It is also possible to look for special vectors that transform in some simple way. For example, we call a vector whose *direction* stays the same under transformation an *eigenvector*; when we apply the matrix to a such a vector it stretches by a factor which we call the *eigenvalue*. More formally we define these as follows:

> A eigenvector v of a matrix $\mathbf{A}$ is a nonzero vector for which
>
> $$\mathbf{Av} = \lambda \mathbf{v},$$
>
> *The vector* v *cannot be the zero vector.* The scalar λ is called an eigenvalue of $\mathbf{A}$.

3.2.1 Computing Eigenvalues

We start with the problem of finding eigenvalues, i.e., values λ for which the equation above has a solution. Using our linear algebra skills, we can rewrite the equation as

$$\begin{aligned} \mathbf{Av} &= \lambda \mathbf{v} \\ \mathbf{Av} - \lambda \mathbf{v} &= \mathbf{0} \\ \mathbf{Av} - \lambda \mathbf{Iv} &= \mathbf{0} \\ (\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v} &= \mathbf{0} \end{aligned} \tag{3.1}$$

The matrix form of this equation is

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{3.2}$$

We can write the previous equation as the augmented matrix

$$\left(\begin{array}{cccc|c} a_{11} - \lambda & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda & 0 \end{array}\right).$$

Using Theorem 2.3, there is a nonzero solution to this system **only** when

$$\det(\mathbf{A} - \lambda \mathbf{I}) \;=\; 0. \tag{3.3}$$

If we expand the determinant in (3.3), we obtain a polynomial expression $p(\lambda)$ of degree n for which we must find the roots:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = p(\lambda)$$

The polynomial $p(\lambda)$ is called the **characteristic polynomial** of the matrix $\mathbf{A}$.

Example 3.2.1 Find the eigenvalues of the matrix from example 3.1.6

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

Answer: The characteristic polynomial of $\mathbf{A}$ is

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3$$

The characteristic equation is

$$\lambda^2 - 2\lambda - 3 = 0$$

which can be factored to

$$(\lambda - 3)(\lambda + 1) = 0.$$

Therefore $\lambda_1 = -1$ and $\lambda_2 = 3$ are the eigenvalues of $\mathbf{A}$.

Recall the table from example 3.1.6. The relative lengths of the series of vectors seem to approach 3 because the eigenvalue is 3 dominates the other smaller eigenvalues. Notice from the plot in 3.1.6 that each successive vector flips across the $45°$ line because the second eigenvalue is negative.

Example 3.2.2 Find the eigenvalues of the matrix

$$\mathbf{B} = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 3 & 1 \end{pmatrix}.$$

Answer: The characteristic polynomial of $\mathbf{B}$ is

$$p(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 0 \\ -1 & -\lambda & 1 \\ 1 & 3 & 1-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - \lambda - 3) - (\lambda - 2).$$

Hint: if you are careful not to fully expand this polynomial, you can often avoid the problem of having to factor a cubic polynomial.

$$-(\lambda - 2)(\lambda^2 - \lambda - 2) = -(\lambda - 2)(\lambda - 2)(\lambda + 1) = 0.$$

Therefore $\lambda_1 = -1$ and $\lambda_2 = \lambda_3 = 2$ are the eigenvalues of $\mathbf{B}$.

The eigenvalues of a matrix $\mathbf{A}$ are found from the characteristic polynomial. The degree of the polynomial for an $n \times n$ matrix is n. Since there are at most n solutions to a polynomial of degree n, there are at most n distinct eigenvalues. However, some of the roots of the polynomial may be "double" or higher order, so it is not guaranteed that you will find n distinct eigenvalues. For example, in 3.2.2, there are only two distinct eigenvalues.

3.3 Computing Eigenvectors

Having found the eigenvalue λ, we substitute it into (3.1) and solve for $\mathbf{v}$ in the equation given by

$$(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v} \quad = \quad \mathbf{0}.$$

Example 3.3.1 Find the eigenvectors of the matrix in example 3.1.6,

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

Answer:

We know from example 3.2.1 that the eigenvalues of $\mathbf{A}$ are $\lambda_1 = -1$ and $\lambda_2 = 3$. There will be two eigenvectors for this system. Each eigenvector corresponds to one of the eigenvalues. First let's look for the eigenvector associated with λ_1. We do this by substituting λ_1 into the system $(\mathbf{A} - \lambda I)\mathbf{v} = \mathbf{0}$, or

$$\begin{aligned} & \left(\begin{array}{cc|c} 1-\lambda_1 & 2 & 0 \\ 2 & 1-\lambda_1 & 0 \end{array}\right), \\ = & \left(\begin{array}{cc|c} 2 & 2 & 0 \\ 2 & 2 & 0 \end{array}\right), \\ \sim & \left(\begin{array}{cc|c} 2 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right). \end{aligned}$$

this system can be written in terms of the individual vectors:

$$2x + 2y = 0,$$

or

$$x = -y.$$

This system has an infinite number of solutions, in other words the eigenvectors are not unique. We can choose any value for x and y which satisfy the relationship $x = -y$. Let's choose $y = 1$ and $x = -1$. Thus, the first eigenvector is $\mathbf{v}_1 = (-1, 1)^T$.

The second eigenvector is found in a similar manner. Use $\lambda_2 = 3$ instead and proceed as above. Finding the second eigenvector involves reducing the augmented matrix:

$$\begin{aligned} & \left(\begin{array}{cc|c} -2 & 2 & 0 \\ 2 & -2 & 0 \end{array}\right), \\ \sim & \left(\begin{array}{cc|c} -2 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right). \end{aligned}$$

The relationship for this matrix is $x = y$. If we let $y = 1$ and $x = 1$, the second eigenvector is $(1, 1)^T$.

Notice that the plot in example 3.1.6 approaches the $45°$ line. This line corresponds to the eigenvector $(1, 1)^T$. The intitial vector approaches this eigenvector upon transformation because this eigenvector corresponds to $\lambda_1 = 3$, which is the dominant eigenvector.

Also recall that in problem 2, part (a), you found a series of vectors starting with $(-1,1)^T$. Because this vector is an eigenvector of the matrix, it flips between $(-1,1)^T$ and $(1,-1)^T$. (It transforms between these two vectors because $\lambda_1 = -1$.)

The last thing to notice that in problem 2, part (b), is that the transformed vectors go toward the eigenvector $(1,1)^T$ even though the starting vector is close to the eigenvector $(-1,1)^T$. The eigenvector $(1,1)^T$ is the dominant eigenvector, so no matter how close we are to the $(-1,1)^T$ eigenvector, the $(1,1)^T$ eigenvector will dominate.

3.3.1 Finding the Eigenvectors of Real Eigenvalues

Example 3.3.2 In Example 3.2.2, the eigenvalues for the matrix

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

where $\lambda_1 = -1$ and $\lambda_2 = \lambda_3 = 2$. Find the corresponding eigenvectors.

Answer: For $\lambda_1 = -1$, $(\mathbf{B} - \lambda_1 \mathbf{I})\,\mathbf{v} = \mathbf{0}$ simplifies to:

$$\begin{aligned} \begin{pmatrix} 2-(-1) & 1 & 0 \\ -1 & -(-1) & 1 \\ 1 & 3 & 1-(-1) \end{pmatrix} \mathbf{v} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \\ \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & 3 & 2 \end{pmatrix} \mathbf{v} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

To solve, we use the augmented matrix

$$\begin{aligned} (\mathbf{B} - \lambda_1 \mathbf{I} |\, \mathbf{0}) &= \left(\begin{array}{ccc|c} 3 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 1 & 3 & 2 & 0 \end{array}\right), \\ &\sim \left(\begin{array}{ccc|c} 1 & 0 & -1/4 & 0 \\ 0 & 1 & 3/4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right). \end{aligned}$$

The solution of the homogeneous system for any scalar r is

$$\begin{pmatrix} r/4 \\ -3r/4 \\ r \end{pmatrix} = r \begin{pmatrix} 1/4 \\ -3/4 \\ 1 \end{pmatrix}.$$

The eigenvector $\mathbf{v_1}$ is given by this vector for any nonzero scalar r. This solution is not unique because $\det(\mathbf{B} - \lambda_1 \mathbf{I})$ is zero. For the final answer any r can be used. It is perhaps nicest to choose a value that that gets rid of the fractions. For example, if we choose $r = 4$, then we have

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix}.$$

The calulation we have just done must must be repeated for the other eigenvalue $\lambda_2 = 2$:

$$\begin{aligned}(\mathbf{B} - \lambda_2 \mathbf{I} | \, \mathbf{0}) &= \left(\begin{array}{rrr|r} 0 & 1 & 0 & 0 \\ -1 & -2 & 1 & 0 \\ 1 & 3 & -1 & 0 \end{array} \right), \\ &\sim \left(\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).\end{aligned}$$

A solution to this system for any scalar s is

$$\begin{pmatrix} s \\ 0 \\ s \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Here, $s = 1$ is an appropriate choice. The final answer is

$$\mathbf{v_2} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Note: Even though we had a repeated eigenvalues, there is only one corresponding eigenvector. A matrix with fewer eigenvectors than its dimension is called *deficient*.

3.3.2 Eigenvectors of Complex Eigenvalues

The eigenvalues of a system are not necessarily real numbers, the the characteristic polynomial can have complex roots.

If the matrix **A** is real, that is all entries are real, then the coefficients of the characteristic polynomial are real. When this is true, if there are complex roots of the polynomial, then they come in complex conjugate pairs. That is, if λ is a root of $p(\lambda)$ then so is $\overline{\lambda}$. (Recall that the complex conjugate of a number $\lambda = a + bi$ is $\overline{\lambda} = a - bi$.)

Thus, when a real matrix has a complex eigenvalue, λ, then the complex conjugate, $\overline{\lambda}$ is also an eigenvalue. Recall that when **v** is an eigenvector of a real matrix **A** with an eigenvalue λ, it is a solution of

$$(\mathbf{A} - \lambda \mathbf{I}) \, \mathbf{v} = \mathbf{0}.$$

The complex conjugate of the equation is

$$\left(\mathbf{A} - \overline{\lambda} \mathbf{I} \right) \overline{\mathbf{v}} = \mathbf{0},$$

and therefore $\overline{\mathbf{v}}$ is an eigenvector with corresponding eigenvalue $\overline{\lambda}$. (Because $\mathbf{A}$ and $\mathbf{I}$ are real matrices, $\mathbf{A} = \overline{\mathbf{A}}$ and $\mathbf{I} = \overline{\mathbf{I}}$.)

Example 3.3.3 Given the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 2 \\ -1 & -3 \end{pmatrix},$$

find the eigenvalues and eigenvectors.

Answer: The characteristic equation of $\mathbf{A}$ is

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} -1-\lambda & 2 \\ -1 & -3-\lambda \end{vmatrix} = \lambda^2 + 4\lambda + 5 = 0.$$

The eigenvalues are $\lambda_1 = -2+i$ and $\lambda_2 = -2-i$. Next, find the eigenvector of $\lambda_1 = -2+i$, by substituting λ_1 into $(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{z} = \mathbf{0}$,

$$\begin{pmatrix} 1-i & 2 \\ -1 & -1-i \end{pmatrix} \mathbf{z_1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution to this system is

$$\begin{pmatrix} 2s \\ (-1+i)s \end{pmatrix} = s \begin{pmatrix} 2 \\ -1+i \end{pmatrix}$$

Choosing $s = 1$, the final answer is

$$\mathbf{z_1} = \begin{pmatrix} 2 \\ -1+i \end{pmatrix}.$$

You can check that the eigenvector for λ_2 is

$$\mathbf{z_2} = \begin{pmatrix} 2 \\ -1-i \end{pmatrix}.$$

As you can see, finding the eigenvalues and eigenvectors of a matrix is a relatively simple procedure. First, the eigenvalues of the matrix are found by taking the determinant of the matrix $\mathbf{A} - \lambda \mathbf{I}$, and setting it to zero. For an n by n matrix, the resulting characteristic polynomial is of order n. When $n = 2$, the characteristic polynomial is a quadratic and can be solved by either factoring or using the quadratic formula. When n is 3 or greater the characteristic polynomial is a cubic or higher order. There is no simple "cubic formula", so factoring is the key to solving these polynomials. *It is recommended that you don't expand the individual terms of the determinant.*

Lastly, recall that finding the eigenvectors of a matrix involves substituting the eigenvalues into the matrix the solving the homogeneous system using gaussian elimination. We are allowed to choose one solution of this homogeneous equation to fix the eigenvector.

3.4 Problems

1. In section 3.1 we encountered one example of a rotation matrix. Any rotation matrix can be written as

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

   (a) Show that the matrix in example 3.1.4, in fact is a rotation matrix. What is θ?

   (b) For any other angle of your choice, write down the rotation matrix corresponding to the angle. Perform the linear transformation on the vector $\mathbf{x} = (1 \;\; 1)^T$ and graph both the vector $\mathbf{x}$ as well as the resultant vector to show that the matrix does perform a rotation.

(c) What is the determinant of the rotation matrix above?

(d) Find the inverse of the rotation matrix above.

(e) Geometrically, how does the inverse of the rotation matrix transform a vector? Give an example.

2. (a) If $x_0 = (-1, 1)^T$, perform the same steps as in example 3.1.6 to find $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$. Construct a table also as in example 3.1.6.

   (b) If $x_0 = (-1, 1.1)^T$, perform the same steps as in example 3.1.6 to find $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$. Construct a table also as in example 3.1.6.

   (c) Are the results the same for both? If not, how do they differ? This will be explained in the next section.

3. Explain why the eigenvalues of a matrix are found by solving $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

4. Find the eigenvalues and eigenvectors of the rotation matrix in problem 1. What are the magnitudes of the eigenvalues? (hint: the magnitude of a complex number $a + bi$ is $\sqrt{a^2 + b^2}$)

5. For the system

$$\begin{aligned} x_1 + x_3 &= a \\ x_2 &= b \\ x_1 - x_3 &= c \end{aligned}$$

   find the eigenvalues and eigenvectors.

6. For

$$\mathbf{B} = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$$

   find the eigenvalues and eigenvectors of $\mathbf{B}$, $\mathbf{B}^2$, and $\mathbf{B}^3$, where $\mathbf{B}^2 = \mathbf{BB}$. What is the relationship between the eigenvalues and eigenvectors of each matrix?

7. For

$$\mathbf{C} = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}$$

   find the eigenvalues and eigenvectors of $\mathbf{C}$ and $\mathbf{C}^{-1}$. Can you see the relationship between the eigenvalues and eigenvectors of each?

8. Explain why eigenvectors are not unique.

9. Find the eigenvalues and eigenvectors of

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -3 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -5 & 0 \end{pmatrix}.$$