

GEOMETRY OF MATRICES

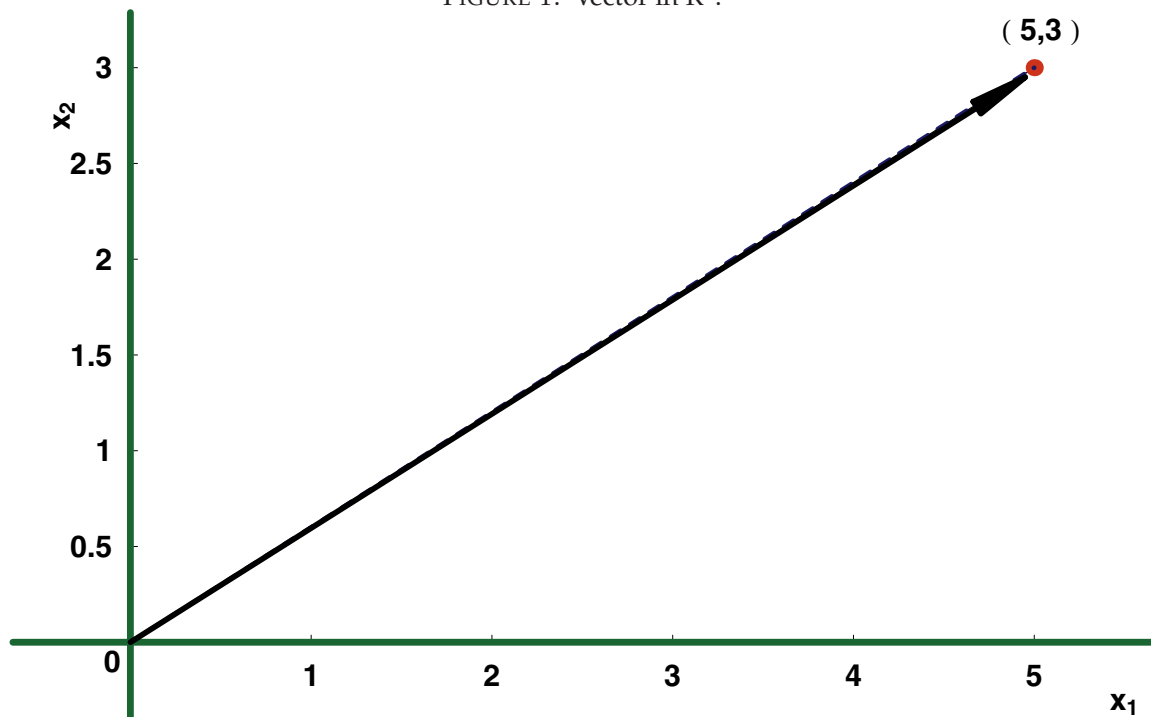
1. SPACES OF VECTORS

1.1. **Definition of R^n .** The space R^n consists of all column vectors with n components. The components are real numbers.

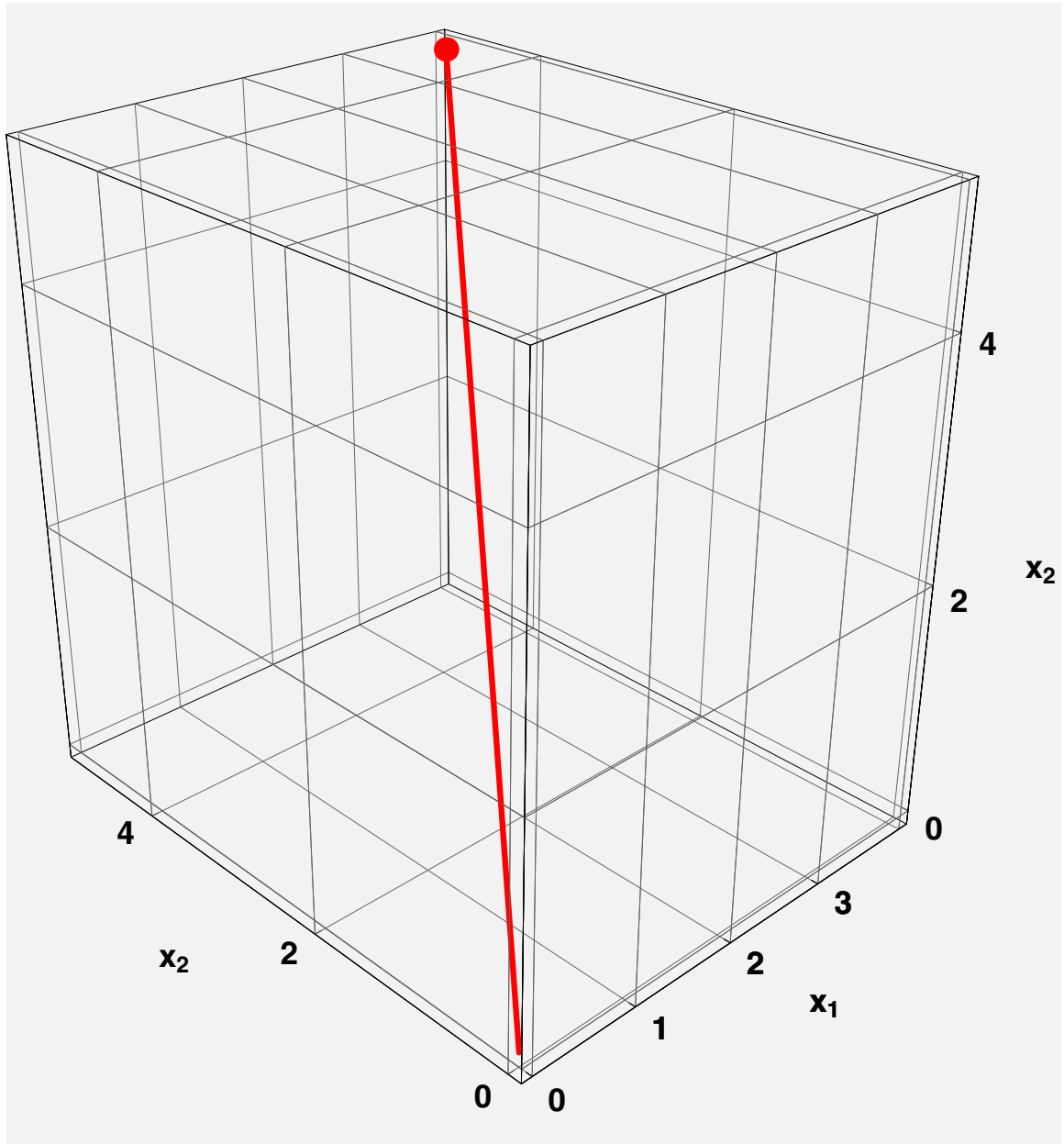
1.2. **Representation of Vectors in R^n .**

1.2.1. R^2 . The space R^2 is represented by the usual x_1, x_2 plane. The two components of the vector give the x_1 and x_2 coordinates of a point and the vector is a directed line segment that goes out from $(0,0)$

FIGURE 1. Vector in R^2 .



1.2.2. R^3 . We can think of this geometrically as the x_1 and x_2 axes being laid out on a table top with the x_3 axis being perpendicular to this horizontal surface as in figure 2. The vector is a directed line segment from $(0,0,0)$ to the point (x_1, x_2, x_3) . We frequently attach an arrow to the end of the line segment, at other times we simply connect the line segment to the dot representing the point.

FIGURE 2. Vector in \mathbb{R}^3 .

1.3. Definition of a vector space. A **vector space** consists of a set V of objects (called vectors) together with two operations, vector addition (assigning a vector $\vec{x} + \vec{y}$ to any vectors \vec{x} and \vec{y} in V) and scalar multiplication (assigning a vector $\alpha \vec{x}$ or $\alpha \cdot \vec{x}$ to any scalar α and a vector \vec{x} in V) such that following eight axioms are satisfied.

V1: $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ (commutative law of vector addition)

- V2:** $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$ (associative law of vector addition)
- V3:** There is an element of V , denoted by $\vec{0}$ or $\vec{0}$ such that $\vec{0} + \vec{x} = \vec{x} + \vec{0} = \vec{x}$.
- V4:** For each \vec{x} in V , there is an element $-\vec{x}$ in V , such that $\vec{x} + -\vec{x} = -\vec{x} + \vec{x} = \vec{0}$.
- V5:** $(\alpha + \beta) \vec{x} = \alpha \vec{x} + \beta \vec{x}$
- V6:** $\alpha (\vec{x} + \vec{y}) = \alpha \vec{x} + \alpha \vec{y}$
- V7:** $(\alpha \beta) \vec{x} = \alpha (\beta \vec{x})$
- V8:** $1 \cdot \vec{x} = \vec{x}$ where 1 is a scalar.

Note that the definitions of $+$ in V.5 are different and that there are two types of multiplication in V.7.

1.4. Properties of a vector space.

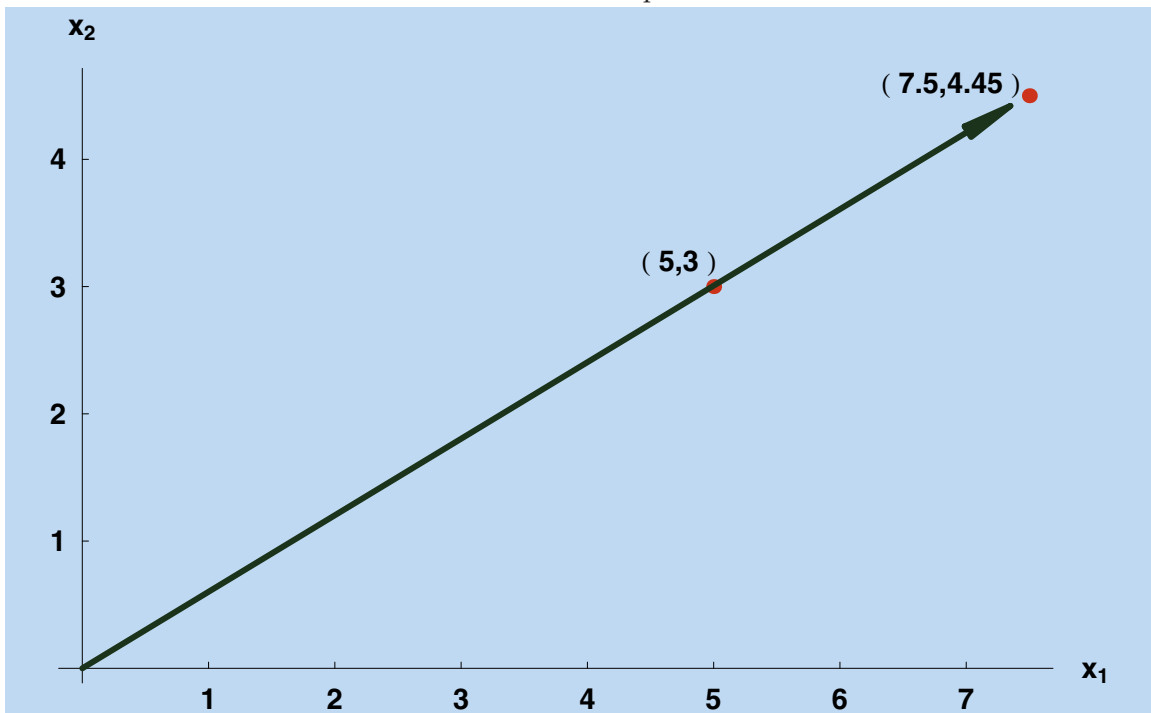
1.4.1. *Geometric representation of scalar multiplication.* A scalar multiple of a vector a , is another vector, say u , whose coordinates are the scalar multiples of a 's coordinates, $u = ca$. Consider the example below.

$$a = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, c = 3$$

$$u = ca = \begin{pmatrix} 6 \\ 15 \end{pmatrix} \tag{1}$$

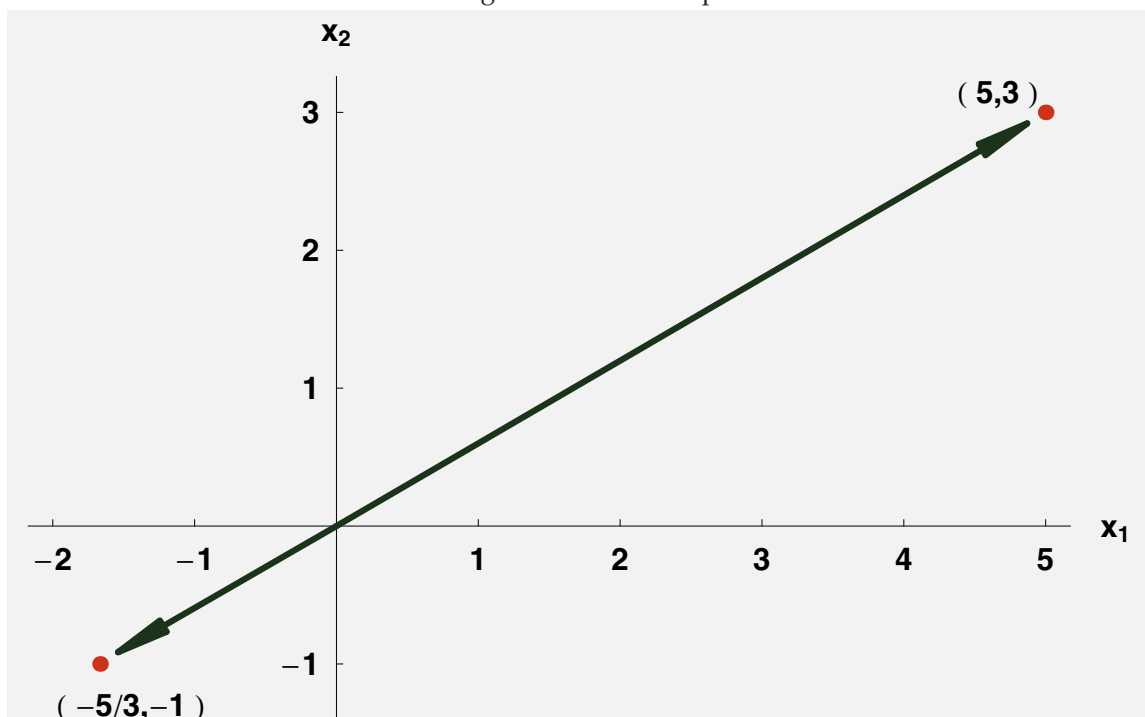
Geometrically, a scalar multiple of a vector a is a segment of the line that passes through 0 and a and continues forever in both directions. For example if we multiply the vector $(5,3)$ by 1.5 we obtain $(7.5, 4.5)$ as shown in figure 3.

FIGURE 3. Scalar Multiple of Vector



If we multiply a vector by a scalar with a negative sign, the direction changes as in figure 4.

FIGURE 4. Negative Scalar Multiple of Vector



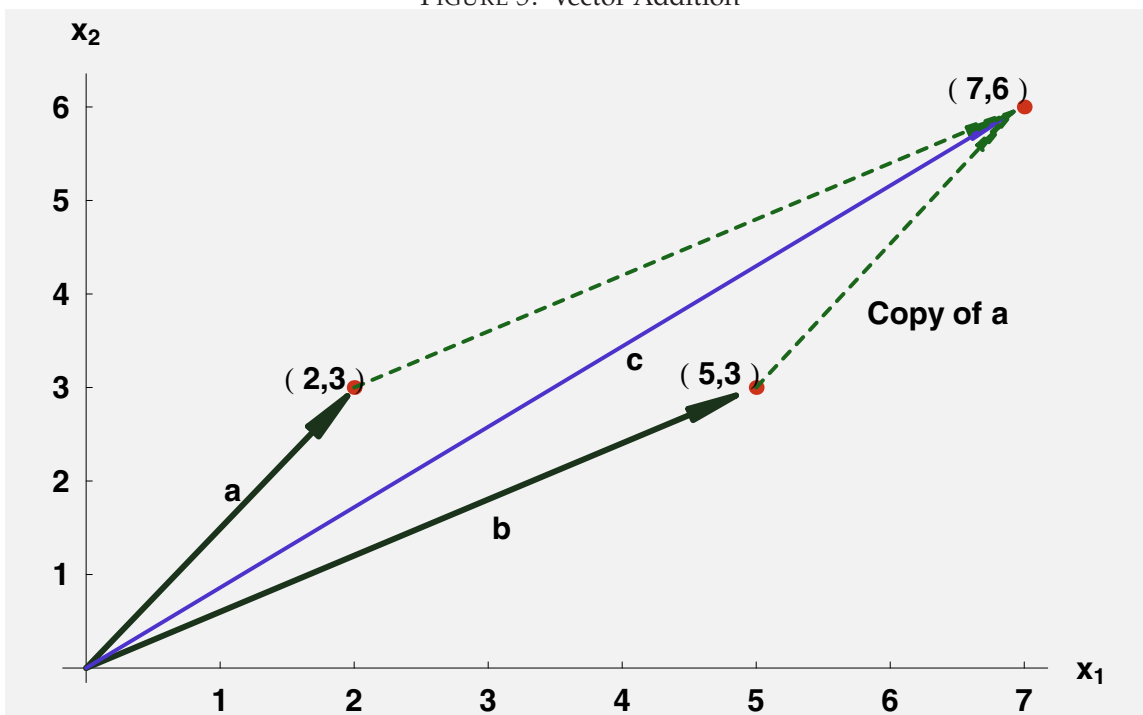
1.4.2. *Geometric representation of addition of vectors.* The sum of two vectors a and b is a third vector whose coordinates are the sums of the corresponding coordinates of a and b . Consider the example below.

$$\begin{aligned} a &= \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ 5 \end{pmatrix} \\ c &= a + b = \begin{pmatrix} -1 \\ 10 \end{pmatrix} \end{aligned} \quad (2)$$

Geometrically, vector addition can be represented by the “parallelogram law” which is that the sum vector $(a+b)$ corresponds to the directed line segment along the diagonal of the parallelogram having a and b as sides. Another way to say this is to move a copy of the vector “ a ” parallel to “ a ” so that its “tail” rests at the head of vector “ b ”, and let the geometrical vector connecting the origin and the head of this shifted “ a ” be the vector $a+b$. Yet another way to think of this is to move in the direction and distance defined by a from the tip of b . Consider figure 5 where

$$\begin{aligned} a &= \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \\ c &= a + b = \begin{pmatrix} 7 \\ 6 \end{pmatrix} \end{aligned} \quad (3)$$

FIGURE 5. Vector Addition



The copy of a extends from the tip of b to the point (7,6). The parallelogram is formed by extending a copy of b from the tip of a which also ends at the point (7,6). The line segment from the origin to this point, which is the diagonal of the parallelogram is the vector c.

1.4.3. *Subspaces.* A subspace of a vector space is a set of vectors (including 0) that satisfies two requirements: If a and b are vectors in the subspace and c is any scalar, then

- 1: $a + b$ is in the subspace
- 2: ca is in the subspace

Note that a subspace necessarily contains the zero vector.

1.4.4. *Subspaces and linear combinations.* A subspace containing the vectors a and b must contain all linear combinations of a and b.

1.4.5. *Column space of a matrix.* The column space of the $m \times n$ matrix A consists of all linear combinations of the columns of A. The combinations can be written as Ax . The column space of A is a subspace of R^m because each of the columns of A has m rows. The system of equations $Ax = b$ is solvable if and only if b is in the column space of A. What this means is that the system is solvable if there is some way to write b as a linear combination of the columns of A. Consider the following example where A is a 3×2 matrix, x is a 2×1 vector and b is a 3×1 vector.

$$A = \begin{pmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{4}$$

The matrix A has two columns. Linear combinations of these columns will lie on a plane in \mathbb{R}^3 . Any vectors b which lie on this plane can be written as linear combinations of the columns of A . Other vectors in \mathbb{R}^3 cannot be written in this fashion. For example consider the vector $b' = [1, 10, 8]$. This can be written as follows

$$A = \begin{pmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ 8 \end{pmatrix} \quad (5)$$

$$(1) \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + (2) \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ 8 \end{pmatrix}$$

where the vector $x' = [1, 2]$.

The vector $b' = [3, 15, 9]$ can also be written as a linear combination of the columns of A . Specifically,

$$A = \begin{pmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \\ 9 \end{pmatrix} \quad (6)$$

$$(3) \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + (1) \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \\ 9 \end{pmatrix}$$

where $x' = [3, 1]$.

To find the coefficients x that allow us to write a vector b as a linear combination of the columns of A we can perform row reduction on the augmented matrix $[A, b]$. For this example we obtain

$$\tilde{A} = \begin{pmatrix} 1 & 0 & 3 \\ 4 & 3 & 15 \\ 2 & 3 & 9 \end{pmatrix} \quad (7)$$

Multiply the first row by 4 to yield

$$(4 \quad 0 \quad 12)$$

and subtract it from the second row

$$\begin{array}{ccc} 4 & 3 & 15 \\ 4 & 0 & 12 \\ \hline 0 & 3 & 3 \end{array}$$

This will give a new matrix on which to operate.

$$\tilde{A}_1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 3 & 3 \\ 2 & 3 & 9 \end{pmatrix} \quad (8)$$

Multiply the first row by 2

$$(2 \ 0 \ 6)$$

and subtract from the third row

$$\begin{array}{ccc} 2 & 3 & 9 \\ 2 & 0 & 6 \\ \hline 0 & 3 & 3 \end{array}$$

This will give

$$\tilde{A}_2 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 3 & 3 \\ 0 & 3 & 3 \end{pmatrix} \tag{9}$$

Now divide the second row by 3

$$\tilde{A}_2 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 3 & 3 \end{pmatrix} \tag{10}$$

Now multiply the second row by 3

$$(0 \ 3 \ 3)$$

and subtract from the third row

$$\begin{array}{ccc} 0 & 3 & 3 \\ 0 & 3 & 3 \\ \hline 0 & 0 & 0 \end{array}$$

This will give

$$\tilde{A}_3 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{11}$$

At this point we have an identity matrix in the upper left and know that $x_1 = 3$ and $x_2 = 1$. The third equation implies that $(0)(x_1) + (0)(x_2) = 0$. So the vector $b' = [3, 15, 9]$ can be written as linear combination of the columns of A .

Now consider the vector $b' = [2, 10, 10]$. Write out the augmented matrix system and perform row operations.

$$\tilde{A} = \begin{pmatrix} 1 & 0 & 2 \\ 4 & 3 & 10 \\ 2 & 3 & 10 \end{pmatrix} \tag{12}$$

Multiply the first row by 4

$$(4 \ 0 \ 8)$$

and subtract from the second row

$$\begin{array}{r}
 4 \quad 3 \quad 10 \\
 4 \quad 0 \quad 8 \\
 \hline
 0 \quad 3 \quad 2
 \end{array}$$

This will give

$$\tilde{A}_1 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 2 \\ 2 & 3 & 10 \end{pmatrix} \quad (13)$$

Multiply the first row by 2

$$(2 \quad 0 \quad 4)$$

and subtract from the third row

$$\begin{array}{r}
 2 \quad 3 \quad 10 \\
 2 \quad 0 \quad 4 \\
 \hline
 0 \quad 3 \quad 6
 \end{array}$$

This will give

$$\tilde{A}_2 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 2 \\ 0 & 3 & 6 \end{pmatrix} \quad (14)$$

Now divide the second row by 3

$$\tilde{A}_3 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{2}{3} \\ 0 & 3 & 6 \end{pmatrix} \quad (15)$$

Now multiply the second row by 3

$$(0 \quad 3 \quad 2)$$

and subtract from the third row

$$\begin{array}{r}
 0 \quad 3 \quad 6 \\
 0 \quad 3 \quad 2 \\
 \hline
 0 \quad 0 \quad 4
 \end{array}$$

This will give

$$\tilde{A}_4 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 4 \end{pmatrix} \quad (16)$$

Now multiply the third row by $1/4$. This will give

$$\tilde{A}_5 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2/3 \\ 0 & 0 & 1 \end{pmatrix} \tag{17}$$

At this point we have a problem because the third row of the matrix implies that $(0)(x_1) + (0)(x_2) = 1$, which is not possible.

Consider now the general case where $b' = [b_1, b_2, b_3]$. The augmented matrix system can be written

$$\tilde{A} = \begin{pmatrix} 1 & 0 & b_1 \\ 4 & 3 & b_2 \\ 2 & 3 & b_3 \end{pmatrix} \tag{18}$$

Multiply the first row by 4

$$(4 \quad 0 \quad 4b_1)$$

and subtract from the second row

$$\begin{array}{r} \begin{array}{ccc} 4 & 3 & b_2 \\ 4 & 0 & 4b_1 \\ \hline 0 & 3 & b_2 - 4b_1 \end{array} \end{array}$$

This will give

$$\tilde{A}_1 = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 3 & b_2 - 4b_1 \\ 2 & 3 & b_3 \end{pmatrix} \tag{19}$$

Multiply the first row by 2

$$(2 \quad 0 \quad 2b_1)$$

and subtract from the third row

$$\begin{array}{r} \begin{array}{ccc} 2 & 3 & b_3 \\ 2 & 0 & 2b_1 \\ \hline 0 & 3 & b_3 - 2b_1 \end{array} \end{array}$$

This will give

$$\tilde{A}_2 = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 3 & b_2 - 4b_1 \\ 0 & 3 & b_3 - 2b_1 \end{pmatrix} \tag{20}$$

Now divide the second row by 3

$$\tilde{A}_3 = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 1 & \frac{b_2 - 4b_1}{3} \\ 0 & 3 & b_3 - 2b_1 \end{pmatrix} \tag{21}$$

Now multiply the second row by 3

$$(0 \quad 3 \quad b_2 - 4b_1)$$

and subtract from the third row

$$\begin{array}{ccc} 0 & 3 & b_3 - 2b_1 \\ 0 & 3 & b_2 - 4b_1 \\ \hline 0 & 0 & b_3 - b_2 + 2b_1 \end{array}$$

This will give

$$\tilde{A}_4 = \begin{pmatrix} 1 & 0 & b_1 \\ 0 & 1 & \frac{b_2 - 4b_1}{3} \\ 0 & 0 & b_3 - b_2 + 2b_1 \end{pmatrix} \quad (22)$$

We have an identity matrix in the upper left hand corner. The general solution is of the form

$$\begin{aligned} x_1 &= b_1 \\ x_2 &= \frac{b_2 - 4b_1}{3} \\ -b_2 + b_3 + 2b_1 &= 0 \end{aligned} \quad (23)$$

Consider first the case from equation 5. For this system

$$b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ 8 \end{pmatrix}, \quad x = \begin{pmatrix} b_1 \\ \frac{b_2 - 4b_1}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{10 - 4}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad (24)$$

If we then substitute we obtain

$$-b_2 + b_3 + 2b_1 = -10 + 8 + 2 = 0 \quad (25)$$

Now consider the case from equation 6

$$b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \\ 9 \end{pmatrix}, \quad x = \begin{pmatrix} b_1 \\ \frac{b_2 - 4b_1}{3} \end{pmatrix} = \begin{pmatrix} 3 \\ \frac{15 - 12}{3} \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad (26)$$

If we then substitute we obtain

$$-b_2 + b_3 + 2b_1 = -15 + 9 + 6 = 0 \quad (27)$$

So for the systems in equations 5 and 6, the given b vector is in the column space of the matrix A and can be written as a linear combination of these columns with coefficients (1,2) and (3,1) respectively. For the system in equation 12 we have

$$b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 10 \\ 10 \end{pmatrix}, \quad x = \begin{pmatrix} b_1 \\ \frac{b_2 - 4b_1}{3} \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{10 - 8}{3} \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{2}{3} \end{pmatrix} \quad (28)$$

If we then substitute we obtain

$$-b_2 + b_3 + 2b_1 = -10 + 10 + 4 \neq 0 \quad (29)$$

The condition on the elements of b is not satisfied and so b cannot be written as a linear combination of the column space of A . The first two elements are correctly written as combinations of the first two elements of each column of A , but the last element of the linear combination is 6 rather than 10. That is

$$(2) \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + (2/3) \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix} \neq \begin{pmatrix} 2 \\ 10 \\ 10 \end{pmatrix}$$

1.4.6. *Nullspace of a matrix.* The nullspace of an $m \times n$ matrix A consists of all solutions to $Ax = 0$. These vectors x are in \mathbb{R}^n . The elements of x are the multipliers of the columns of A whose weighted sum gives the zero vector. The nullspace containing the solutions x is denoted by $N(A)$. The null space is a subspace of \mathbb{R}^n while the column space is a subspace of \mathbb{R}^m . For many matrices, the only solution to $Ax = 0$ is $x = 0$. If $n > m$, the nullspace will always contain vectors other than $x = 0$. Consider a few examples.

1: Consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Multiply the first row by 3

$$(3 \quad 6)$$

and subtract it from the second row

$$\begin{array}{r} 3 \quad 6 \\ 3 \quad 6 \\ \hline 0 \quad 0 \end{array}$$

This will give the new matrix.

$$\tilde{A}_2 = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

This implies that x_2 can be anything. The convention is to set it equal to one. If $x_2 = 1$, then we have

$$\begin{aligned} x_1 + (2)(1) &= 0 \\ \Rightarrow x_1 &= -2 \end{aligned}$$

so

$$(-2) \begin{pmatrix} 1 \\ 3 \end{pmatrix} + (1) \begin{pmatrix} 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The null space of the matrix A is the vector $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$. All vectors of this form where the first element is -2 times the second element will make the equation $Ax = 0$ true.

2: Consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Multiply the first row by 3

$$(3 \quad 6)$$

and subtract it from the second row

$$\begin{array}{r} 3 \quad 6 \\ 3 \quad 4 \\ \hline 0 \quad 2 \end{array}$$

This will give the new matrix.

$$\tilde{A}_2 = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

This implies that x_2 must be zero. If $x_2 = 0$, then we have

$$\begin{aligned} x_1 + (2)(0) &= 0 \\ \Rightarrow x_1 &= 0 \end{aligned}$$

so

$$(0) \begin{pmatrix} 1 \\ 3 \end{pmatrix} + (0) \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The null space of the matrix A is the vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. The only vector \mathbf{x} that will make the equation $A\mathbf{x} = 0$ is $\mathbf{x} = 0$.

3: Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 8 & 10 \\ 3 & 6 & 11 & 14 \end{pmatrix}$$

Multiply the first row by 2

$$(2 \quad 4 \quad 6 \quad 8)$$

and subtract it from the second row

$$\begin{array}{r} 2 \quad 4 \quad 8 \quad 10 \\ 2 \quad 4 \quad 6 \quad 8 \\ \hline 0 \quad 0 \quad 2 \quad 2 \end{array}$$

This will give the new matrix.

$$\tilde{A}_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 11 & 14 \end{pmatrix}$$

Now multiply the first row by 3

$$(3 \quad 6 \quad 9 \quad 12)$$

and subtract it from the third row

$$\begin{array}{cccc} 3 & 6 & 11 & 14 \\ 3 & 6 & 9 & 12 \\ \hline 0 & 0 & 2 & 2 \end{array}$$

This will give the new matrix.

$$\tilde{A}_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

Now subtract the second row from the third row

$$\begin{array}{cccc} 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ \hline 0 & 0 & 0 & 0 \end{array}$$

This will give the new matrix.

$$\tilde{A}_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

If we write the system using \tilde{A}_4 we obtain

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

The variables x_2 and x_4 can be any values because of the zeroes in the pivot columns. The convention is to set them equal to zero and one. First write the system with $x_2 = 1$ and $x_4 = 0$. This will give

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ 1 \\ x_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This implies that x_3 must be zero. If $x_3 = 0$, then we have

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow x_1 + 2 = 0$$

$$\Rightarrow x_1 = -2$$

The vector $x_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ is in the nullspace of the matrix A as can be seen by writing out the system as follows

$$(-2) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + (1) \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + (0) \begin{pmatrix} 3 \\ 8 \\ 11 \end{pmatrix} + (0) \begin{pmatrix} 4 \\ 10 \\ 14 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now write the system with $x_2 = 0$ and $x_4 = 1$. This will give

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ 0 \\ x_3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This implies that $2x_3 + 2$ must be zero. This implies that $x_3 = -1$. Making this substitution we have

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow x_1 - 2 + 4 = 0$$

$$\Rightarrow x_1 = -1$$

The vector $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ is in the nullspace of the matrix A as can be seen by writing out the system as follows

$$(-1) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + (0) \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + (-1) \begin{pmatrix} 3 \\ 8 \\ 11 \end{pmatrix} + (1) \begin{pmatrix} 4 \\ 10 \\ 14 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We can then write a general \mathbf{x} as follows

$$\mathbf{x} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2x_2 - x_4 \\ x_2 \\ -x_4 \\ x_4 \end{pmatrix}$$

4: Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 10 \\ 3 & 6 & 9 & 11 \\ -1 & -2 & -3 & 7 \end{pmatrix}$$

Multiply the first row by 2

$$(2 \quad 4 \quad 6 \quad 2)$$

and subtract it from the second row

$$\begin{array}{cccc}
 2 & 4 & 6 & 10 \\
 2 & 4 & 6 & 2 \\
 \hline \hline
 0 & 0 & 0 & 8
 \end{array}$$

This will give the new matrix.

$$\tilde{A}_2 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 3 & 6 & 9 & 11 \\ -1 & -2 & -3 & 7 \end{pmatrix}$$

Now multiply the first row by 3

$$(3 \ 6 \ 9 \ 3)$$

and subtract it from the third row

$$\begin{array}{cccc}
 3 & 6 & 9 & 11 \\
 3 & 6 & 9 & 3 \\
 \hline \hline
 0 & 0 & 0 & 8
 \end{array}$$

This will give the new matrix.

$$\tilde{A}_3 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 8 \\ -1 & -2 & -3 & 7 \end{pmatrix}$$

Now multiply the first row by negative one and subtract it from the fourth row

$$\begin{array}{cccc}
 -1 & -2 & -3 & 7 \\
 -1 & -2 & -3 & -1 \\
 \hline \hline
 0 & 0 & 0 & 8
 \end{array}$$

This will give the new matrix.

$$\tilde{A}_4 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Now multiply the second row by one and subtract it from the third row

$$\begin{array}{cccc}
 0 & 0 & 0 & 8 \\
 0 & 0 & 0 & 8 \\
 \hline \hline
 0 & 0 & 0 & 0
 \end{array}$$

This will give the new matrix.

$$\tilde{A}_5 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Now multiply the second row negative one and subtract it from the fourth row

$$\begin{array}{cccc} 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 8 \\ \hline 0 & 0 & 0 & 0 \end{array}$$

This will give the new matrix.

$$\tilde{A}_6 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

If we write the system using \tilde{A}_6 we obtain

$$\begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow x_1 + 2x_2 + 3x_3 + x_4 = 0$$

$$8x_4 = 0$$

This implies that $x_4 = 0$. The system has three variables and only one equation. Assume that x_2 and x_3 are free variables given the zeroes in the pivot columns. Set $x_2 = 1$ and $x_3 = 0$ to obtain

$$\begin{aligned} x_1 + 2x_2 + 0 &= 0 \\ \Rightarrow x_1 + 2 &= 0 \\ \Rightarrow x_1 &= -2 \end{aligned}$$

So the vector $\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ is in the nullspace of the matrix A as can be seen by writing out the system as follows

$$(-2) \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix} + (1) \begin{pmatrix} 2 \\ 4 \\ 6 \\ -2 \end{pmatrix} + (0) \begin{pmatrix} 3 \\ 6 \\ 9 \\ -3 \end{pmatrix} + (0) \begin{pmatrix} 1 \\ 10 \\ 11 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Now set $x_2 = 0$ and $x_3 = 1$ to obtain

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

So the vector $\begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ is in the nullspace of the matrix A as can be seen by writing out the system as follows

$$(-3) \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix} + (0) \begin{pmatrix} 2 \\ 4 \\ 6 \\ -2 \end{pmatrix} + (1) \begin{pmatrix} 3 \\ 6 \\ 9 \\ -3 \end{pmatrix} + (0) \begin{pmatrix} 1 \\ 10 \\ 11 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We can then write a general \mathbf{x} as follows

$$\mathbf{x} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2x_2 - 3x_3 \\ x_2 \\ x_3 \\ 0 \end{pmatrix}$$

1.5. Basis vectors.

1.5.1. *Definition of a basis.* A set of vectors in a vector space is a **basis** for that vector space if any vector in the vector space can be written as a linear combination of them. The minimum number of vectors needed to form a basis for \mathbb{R}^k is k . For example, in \mathbb{R}^2 , we need two vectors of length two to form a basis. One vector would only allow for other points in \mathbb{R}^2 that lie along the line through that vector.

1.5.2. *Example for \mathbb{R}^2 .* Consider the following three vectors in \mathbb{R}^2

$$a_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{30}$$

If a_1 and a_2 are a basis for \mathbb{R}^2 , then we can write any arbitrary vector as a linear combination of them. That is we can write

$$\alpha_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{31}$$

To see the conditions under which a_1 and a_2 form such a basis we can solve this system for α_1 and α_2 as follows. Appending the column vector b to the matrix A, we obtain the augmented matrix for the system. This is written as

$$\tilde{A} = \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{bmatrix} \tag{32}$$

We can perform row operations on this matrix to reduce it to reduced row echelon form. We will do this in steps. The first step is to divide each element of the first row by a_{11} . This will give

$$\tilde{A} = \begin{bmatrix} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \\ a_{21} & a_{22} & b_2 \end{bmatrix} \tag{33}$$

Now multiply the first row by a_{21} to yield

$$\left[\begin{array}{ccc} a_{21} & \frac{a_{21} a_{12}}{a_{11}} & \frac{a_{21} b_1}{a_{11}} \end{array} \right]$$

and subtract it from the second row

$$\begin{array}{ccc} a_{21} & a_{22} & b_2 \\ a_{21} & \frac{a_{21} a_{12}}{a_{11}} & \frac{a_{21} b_1}{a_{11}} \\ \hline 0 & a_{22} - \frac{a_{21} a_{12}}{a_{11}} & b_2 - \frac{a_{21} b_1}{a_{11}} \end{array}$$

This will give a new matrix on which to operate.

$$\tilde{A}_2 = \left[\begin{array}{ccc} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \\ 0 & \frac{a_{11} a_{22} - a_{21} a_{12}}{a_{11}} & \frac{a_{11} b_2 - a_{21} b_1}{a_{11}} \end{array} \right]$$

Now multiply the second row by $\frac{a_{11}}{a_{11} a_{22} - a_{21} a_{12}}$ to obtain

$$\tilde{A}_3 = \left[\begin{array}{ccc} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \\ 0 & 1 & \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}} \end{array} \right] \quad (34)$$

Now multiply the second row by $\frac{a_{12}}{a_{11}}$ and subtract it from the first row. First multiply the second row by $\frac{a_{12}}{a_{11}}$ to yield.

$$\left[\begin{array}{ccc} 0 & \frac{a_{12}}{a_{11}} & \frac{a_{12} (a_{11} b_2 - a_{21} b_1)}{a_{11} (a_{11} a_{22} - a_{21} a_{12})} \end{array} \right] \quad (35)$$

Now subtract the expression in equation 35 from the first row of \tilde{A}_3 to obtain the following row.

$$\left[\begin{array}{ccc} 1 & \frac{a_{12}}{a_{11}} & \frac{b_1}{a_{11}} \end{array} \right] - \left[\begin{array}{ccc} 0 & \frac{a_{12}}{a_{11}} & \frac{a_{12} (a_{11} b_2 - a_{21} b_1)}{a_{11} (a_{11} a_{22} - a_{21} a_{12})} \end{array} \right] = \left[\begin{array}{ccc} 1 & 0 & \frac{b_1}{a_{11}} - \frac{a_{12} (a_{11} b_2 - a_{21} b_1)}{a_{11} (a_{11} a_{22} - a_{21} a_{12})} \end{array} \right] \quad (36)$$

Now replace the first row in \tilde{A}_3 with the expression in equation 36 of obtain \tilde{A}_4

$$\tilde{A}_4 = \left[\begin{array}{ccc} 1 & 0 & \frac{b_1}{a_{11}} - \frac{a_{12} (a_{11} b_2 - a_{21} b_1)}{a_{11} (a_{11} a_{22} - a_{21} a_{12})} \\ 0 & 1 & \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}} \end{array} \right] \quad (37)$$

This can be simplified as by putting the upper right hand term over a common denominator, and canceling like terms as follows

$$\begin{aligned}
 \tilde{A}_4 &= \begin{bmatrix} 1 & 0 & \frac{b_1 a_{11}^2 a_{22} - b_1 a_{11} a_{21} a_{12} - a_{12} a_{11}^2 b_2 + a_{11} a_{12} a_{21} b_1}{a_{11}^2 (a_{11} a_{22} - a_{21} a_{12})} \\ 0 & 1 & \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}} \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & \frac{b_1 a_{11}^2 a_{22} - a_{12} a_{11}^2 b_2}{a_{11}^2 (a_{11} a_{22} - a_{21} a_{12})} \\ 0 & 1 & \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}} \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{21} a_{12}} \\ 0 & 1 & \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}} \end{bmatrix}
 \end{aligned} \tag{38}$$

We can now read off the solutions for x_1 and x_2 . They are

$$\begin{aligned}
 x_1 &= \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{21} a_{12}} \\
 x_2 &= \frac{a_{11} b_2 - a_{21} b_1}{a_{11} a_{22} - a_{21} a_{12}}
 \end{aligned} \tag{39}$$

As long as this determinant is not zero, we can write b as a function of a_1 and a_2 . If the determinant of A is not zero, then a_1 and a_2 are linearly independent. Geometrically this means that they do not lie on the same line. So in \mathbb{R}^2 , any two vectors that point in different directions are a basis for the space.

1.5.3. *Linear dependence.* A set of vectors is linearly dependent if any one of the vectors in the set can be written as a linear combination of the others. The largest number of linearly independent vectors we can have in \mathbb{R}^k is k .

1.5.4. *Linear independence.* A set of vectors is linearly independent if and only if the only solution to

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_k a_k = 0 \tag{40}$$

is

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \tag{41}$$

We can also write this in matrix form. The columns of the matrix A are independent if the only solution to the equation $Ax = 0$ is $x = 0$.

1.5.5. *Spanning vectors.* The set of all linear combinations of a set of vectors is the vector space spanned by those vectors. By this we mean all vectors in this space can be written as a linear combination of this particular set of vectors. Consider for example the following vectors in \mathbb{R}^3 .

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \tag{42}$$

The vector a_3 can be written as $a_1 + 2a_2$. But vectors in \mathbb{R}^3 having a non-zero third element cannot be written as a combination of a_1 and a_2 . These three vectors do not form a basis for \mathbb{R}^3 . They span the space that is made up of the “floor” of \mathbb{R}^3 . Similarly the two vectors from equation 4 make up a basis for a space defined as the plane passing through those two vectors.

1.5.6. *Rowspace of a matrix.* The row space of the $m \times n$ matrix A consists of all linear combinations of the rows of A . The combinations can be written as $x'A$ or $A'x$ depending on whether one considers the resulting vectors to be rows or columns. The row space of A is a subspace of \mathbb{R}^n because each of the rows of A has n columns. The **row space** of a matrix is the subspace of \mathbb{R}^n spanned by the rows.

1.5.7. *Linear independence and the basis for a vector space.* A basis for a vector space with k dimensions is any set of k linearly independent vectors in that space.

1.5.8. *Formal relationship between a vector space and its basis vectors.* A **basis** for a vector space is a sequence of vectors that has two properties simultaneously.

- 1: The vectors are linearly independent
- 2: The vectors span the space

There will be one and only one way to write any vector in a given vector space as a linear combination of a set of basis vectors. There are an infinite number of basis vectors for a given space, but only one way to write any given vector as a linear combination of a particular basis.

1.5.9. *Bases and Invertible Matrices.* The vectors $\xi_1, \xi_2, \dots, \xi_n$ are a basis for \mathbb{R}^n exactly when they are the columns of an $n \times n$ invertible matrix. Therefore \mathbb{R}^n has infinitely many bases, one associated with every different invertible matrix.

1.5.10. *Pivots and Bases.* When reducing an $m \times n$ matrix A to row-echelon form, the pivot columns form a basis for the column space of the matrix A . The pivot rows form a basis for the row space of the matrix A .

1.5.11. *Dimension of a vector space.* The dimension of a vector space is the number of vectors in every basis. For example, the dimension of \mathbb{R}^2 is 2, while the dimension of the vector space consisting of points on a particular plane in \mathbb{R}^3 is also 2.

1.5.12. *Dimension of a subspace of a vector space.* The dimension of a subspace space S_n of an n -dimensional vector space V_n is the maximum number of linearly independent vectors in the subspace.

1.5.13. *Rank of a Matrix.*

- 1: The number of non-zero rows in the row echelon form of an $m \times n$ matrix A produced by elementary operations on A is called the rank of A .
- 2: The rank of an $m \times n$ matrix A is the number of pivot columns in the row echelon form of A .
- 3: The column rank of an $m \times n$ matrix A is the maximum number of linearly independent columns in A .
- 4: The row rank of an $m \times n$ matrix A is the maximum number of linearly independent rows in A .
- 5: The column rank of an $m \times n$ matrix A is equal to the row rank of the $m \times n$ matrix A . This common number is called the **rank** of A .
- 6: An $n \times n$ matrix A with rank = n is said to be of full rank.

1.5.14. *Dimension and Rank.* The dimension of the column space of an $m \times n$ matrix A equals the rank of A , which also equals the dimension of the row space of A . The number of independent columns of A equals the number of independent rows of A . As stated earlier, the r columns containing pivots in the row echelon form of the matrix A form a basis for the column space of A .

1.5.15. *Vector spaces and matrices.* An $m \times n$ matrix A with full column rank (i.e., the rank of the matrix is equal to the number of columns) has all the following properties.

- 1: The n columns are independent
- 2: The only solution to $Ax = 0$ is $x = 0$.
- 3: The rank of the matrix = dimension of the column space = n
- 4: The columns are a basis for the column space.

1.5.16. *Determinants, Minors, and Rank.*

Theorem 1. *The rank of an $m \times n$ matrix A is k if and only if every minor in A of order $k + 1$ vanishes, while there is at least one minor of order k which does not vanish.*

Proposition 1. Consider an $m \times n$ matrix A .

- 1: $\det A = 0$ if every minor of order $n - 1$ vanishes.
- 2: If every minor of order n equals zero, then the same holds for the minors of higher order.
- 3 (**restatement of theorem**): The largest among the orders of the non-zero minors generated by a matrix is the rank of the matrix.

1.5.17. *Nullity of an $m \times n$ Matrix A .* The nullspace of an $m \times n$ matrix A is made up of vectors in \mathbb{R}^n and is a subspace of \mathbb{R}^n . The dimension of the nullspace of A is called the nullity of A . It is the the maximum number of linearly independent vectors in the nullspace.

1.5.18. *Dimension of Row Space and Null Space of an $m \times n$ matrix A .* Consider an $m \times n$ Matrix A . The dimension of the nullspace, the maximum number of linearly independent vectors in the nullspace, plus the rank of A , is equal to n . Specifically,

Theorem 2.

$$\text{rank}(A) + \text{nullity}(A) = n \tag{43}$$

Proof. Let the vectors $\xi_1, \xi_2, \dots, \xi_k$ be a basis for the **nullspace** of the $m \times n$ matrix A . This is a subset of \mathbb{R}^n with n or less elements. These vectors are, of course, linearly independent. There exist vectors $\xi_{k+1}, \xi_{k+2}, \dots, \xi_n$ in \mathbb{R}^n such that $\xi_1, \xi_2, \dots, \xi_k, \xi_{k+1}, \dots, \xi_n$ form a **basis** for \mathbb{R}^n . We can prove that

$$A (\xi_{k+1} \ \xi_{k+2} \ \cdots \ \xi_n) = (A\xi_{k+1} \ A\xi_{k+2} \ \cdots \ A\xi_n)$$

is a basis for the column space of A . The vectors $A\xi_1, A\xi_2, \dots, A\xi_n$ span the column space of A because $\xi_1, \xi_2, \dots, \xi_k, \xi_{k+1}, \dots, \xi_n$ form a basis for \mathbb{R}^n and thus certainly span \mathbb{R}^n . Because $A\xi_j = 0$ for $j \leq k$, we see that $A\xi_{k+1}, A\xi_{k+2}, \dots, A\xi_n$ also span the column space of A . Specifically we obtain

$$A (\xi_1 \ \xi_2 \ \cdots \ \xi_k \ \xi_{k+1} \ \cdots \ \xi_n) = (\mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0} \ A\xi_{k+1} \ A\xi_{k+2} \ \cdots \ A\xi_n)$$

We need to show that these vectors are independent. Suppose that they are not independent. Then there exist scalars c_i such that

$$\sum_{i=k+1}^n c_i (A\xi_i) = 0 \tag{44}$$

We can rewrite 44 as follows

$$A \left(\sum_{i=k+1}^n c_i \xi_i \right) = 0 \quad (45)$$

This implies that the vector

$$\xi = \sum_{i=k+1}^n c_i \xi_i \quad (46)$$

is in the nullspace of A. Because $\xi_1, \xi_2, \dots, \xi_k$ form a basis for the nullspace of A, there must be scalars, $\gamma_1, \gamma_2, \dots, \gamma_k$ such that

$$\xi = \sum_{i=1}^k \gamma_i \xi_i \quad (47)$$

If we subtract one expression for ξ from the other, we obtain

$$\xi = \sum_{i=1}^k \gamma_i \xi_i - \sum_{j=k+1}^n c_j \xi_j = 0 \quad (48)$$

and because $\xi_1, \xi_2, \dots, \xi_k, \xi_{k+1}, \dots, \xi_n$ are linearly independent, we must have

$$\gamma_1 = \gamma_2 = \dots = \gamma_k = c_{k+1} = c_{k+2} = \dots = c_n = 0 \quad (49)$$

If r is the rank of A, the fact that $A\xi_{k+1} A\xi_{k+2} \dots A\xi_n$ form a basis for the column space of A tells us that $r = n - k$. Because k is the nullity of A then we have

$$\begin{aligned} \text{rank}(A) + \text{nullity}(A) &= n \\ n - k + k &= n \end{aligned} \quad (50)$$

□

2. PROJECTIONS

2.1. Orthogonal vectors.

2.1.1. *Definition of orthogonal vectors.* Two vectors a and b are said to be orthogonal if their inner product is zero, that is if $a'b = 0$. Geometrically two vectors are orthogonal if they are perpendicular to each other.

2.1.2. *Example in R^2 .* Consider the two unit vectors

$$a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (51)$$

$$a'b = [1 \ 0] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = (1)(0) + (0)(1) = 0$$

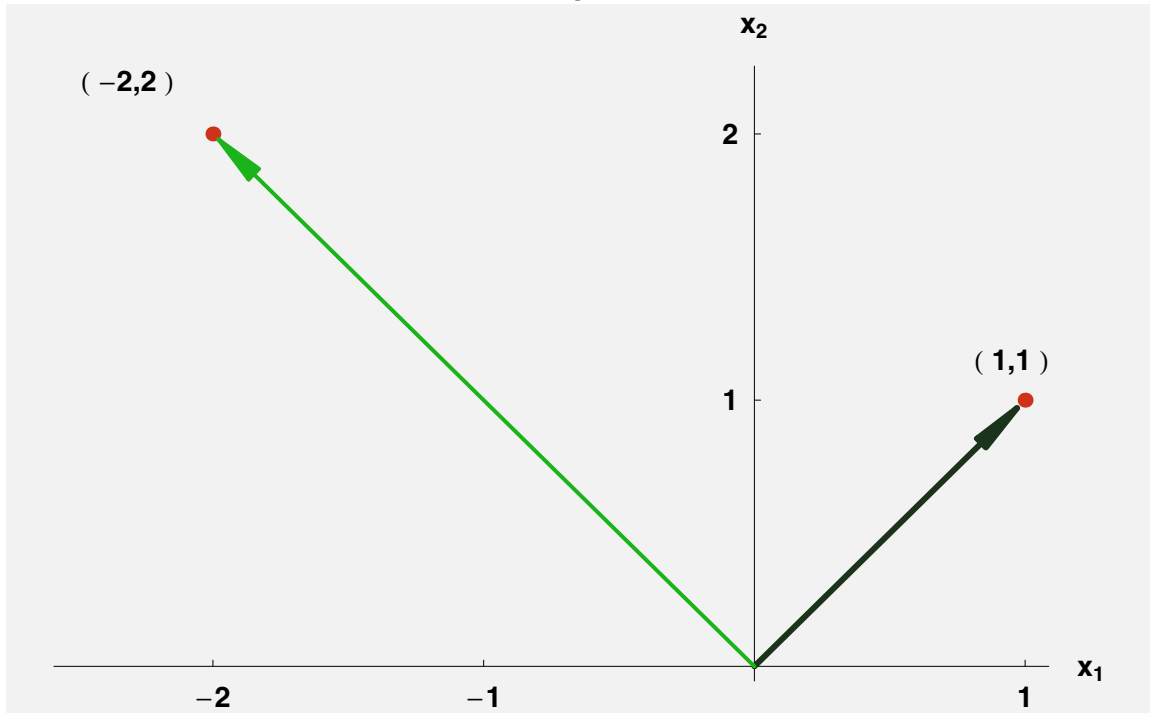
Consider two different vectors that are also orthogonal

$$a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, b = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \quad (52)$$

$$a'b = [1 \ 1] \begin{bmatrix} -2 \\ 2 \end{bmatrix} = (1)(-2) + (1)(2) = 0$$

We can represent this second case graphically as in figure 6

FIGURE 6. Orthogonal Vectors in \mathbb{R}^2



2.1.3. *An example in \mathbb{R}^3 .* Consider the two vectors

$$a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \tag{53}$$

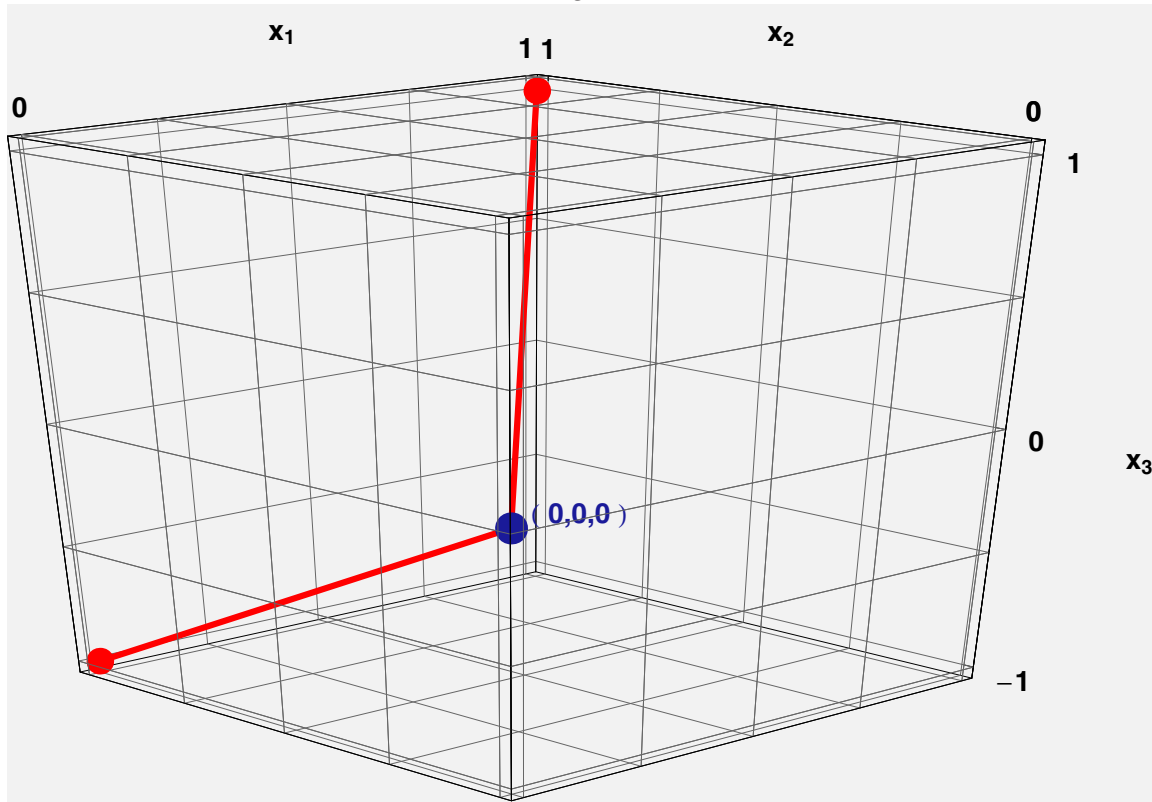
$$a'b = [1 \ 1 \ 1] \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = (1)(0) + (1)(1) + (-1)(1) = 0$$

We can represent this three dimensional case as in figure 7 or alternatively as in figure 8.

2.1.4. *Another example in \mathbb{R}^3 .* Consider the three vectors

$$a = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{2} \\ \frac{3}{2} \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, c = \begin{bmatrix} \frac{4}{9} \\ \frac{-1}{9} \\ \frac{-1}{9} \end{bmatrix} \tag{54}$$

$$c'a = [4/9 \ -1/9 \ -1/9] \begin{bmatrix} \frac{3}{4} \\ \frac{3}{2} \\ \frac{3}{2} \end{bmatrix} = \left(\frac{4}{9}\right) \left(\frac{3}{4}\right) + \left(\frac{-1}{9}\right) \left(\frac{3}{2}\right) + \left(\frac{-1}{9}\right) \left(\frac{3}{2}\right) = 0$$

FIGURE 7. Orthogonal Vectors in \mathbb{R}^3 

Here a and c are orthogonal but a and b are obviously not. We can represent this three dimensional case graphically in figure 9. Note that the line segment d is parallel to c and perpendicular to a , just as c is orthogonal to a .

2.1.5. *Another example in \mathbb{R}^3 .* Consider the four vectors below where p is a scalar multiple of a .

$$a = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, p = \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \\ 0 \end{bmatrix}, e = \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ 1 \end{bmatrix} \quad (55)$$

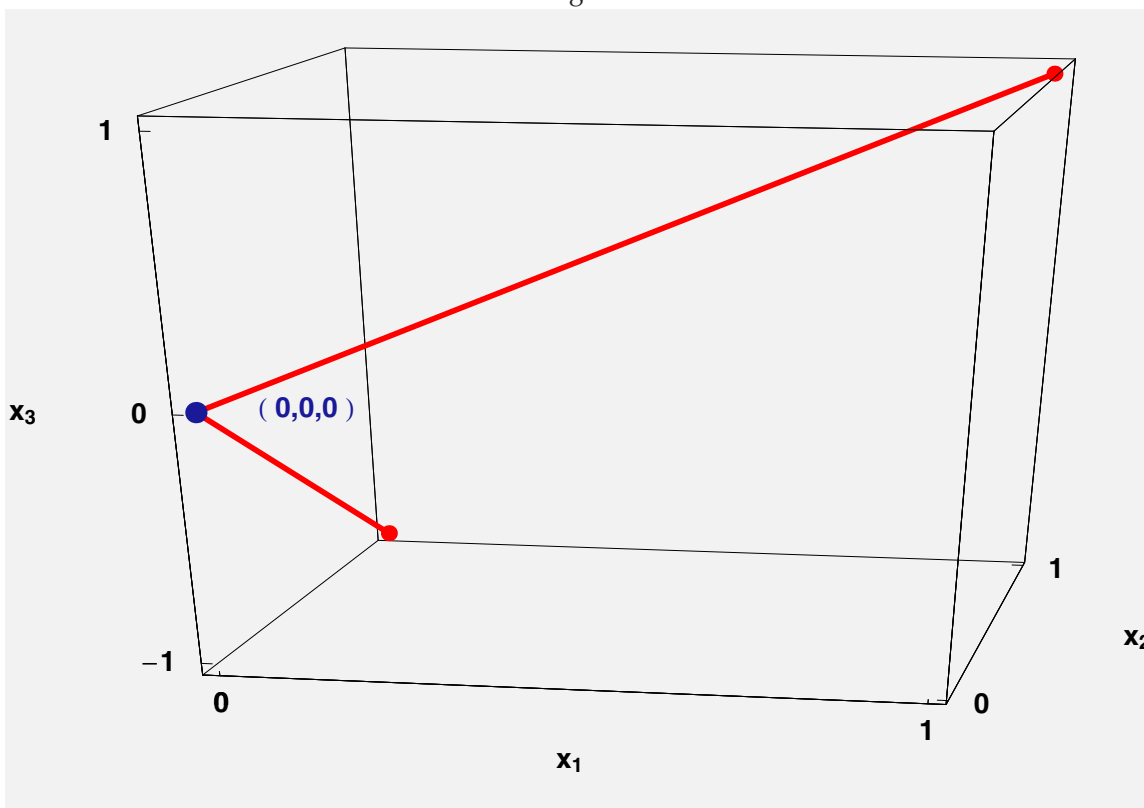
We can see that e is orthogonal to p .

$$e'p = \begin{bmatrix} \frac{2}{5} & -\frac{1}{5} & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \\ 0 \end{bmatrix} = \left(\frac{2}{5}\right)\left(\frac{3}{5}\right) + \left(-\frac{1}{5}\right)\left(\frac{6}{5}\right) + (1)(0) = 0 \quad (56)$$

$$e'a = \begin{bmatrix} \frac{2}{5} & -\frac{1}{5} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \left(\frac{2}{5}\right)(1) + \left(-\frac{1}{5}\right)(2) + (1)(2) = 0$$

and to a . This example is represented in figure 10.

FIGURE 8. Orthogonal Vectors in \mathbb{R}^3

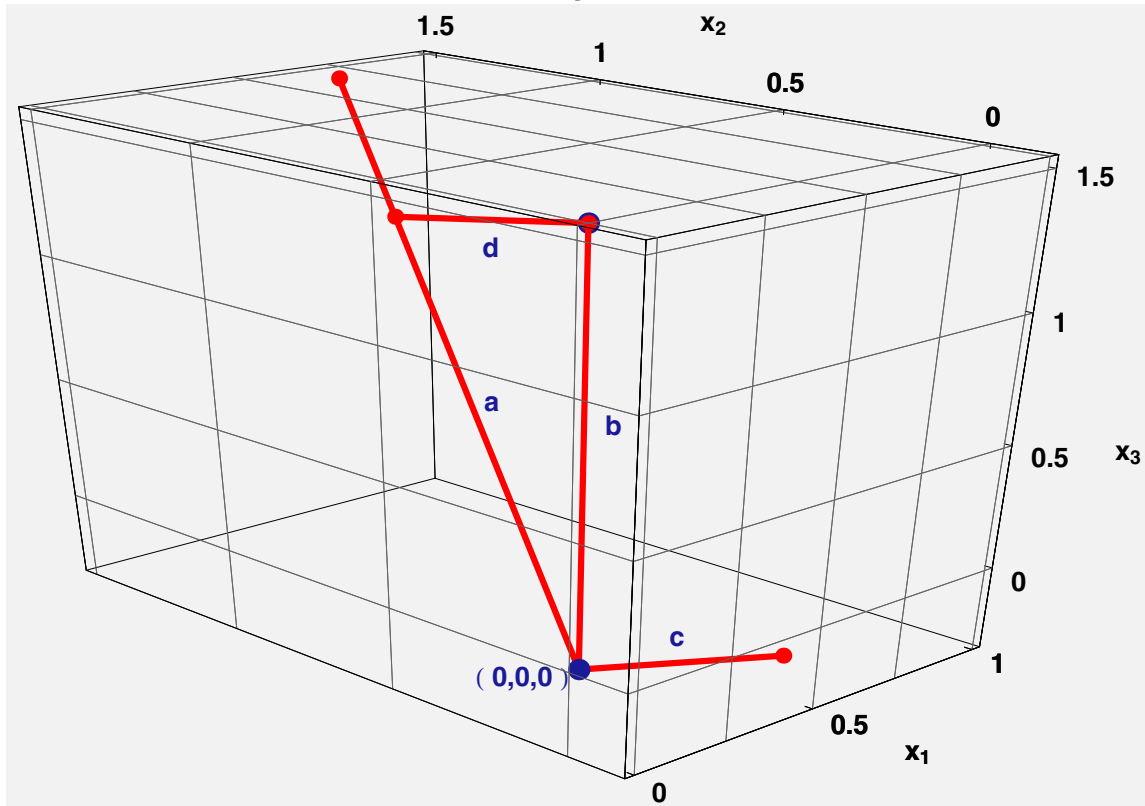


2.2. Projection onto a line.

2.2.1. *Idea and definition of a projection.* Consider a vector $b' = [b_1, b_2, \dots, b_n]$ in \mathbb{R}^n . Also consider a line through the origin in \mathbb{R}^n that passes through the point $a = [a_1, a_2, \dots, a_n]$. To project the vector onto the line a , we find the point on the line through the point a that is closest to b . We find this point on the line a (call it p) by finding the line connecting b and the line through a that is perpendicular to the line through a . This point p (which is on the line through a) will be some multiple of a , i.e. $p = ca$ where c is a scalar. Consider figure 11. The idea is to find the point along the line a that is closest to the tip of the vector b . In other words, find the vector $b-p$ that is perpendicular to a . Given that p can be written as scalar multiple of a , this also means that $b-ca$ is perpendicular or orthogonal to a . This implies that $a \cdot (b-ca) = 0$, or $a \cdot b - ca \cdot a = 0$. Given that $a \cdot b$ and $a \cdot a$ are scalars we can solve this for the scalar c as in equation 57.

$$\begin{aligned} a \cdot (b - ca) &= 0 \\ \Rightarrow a \cdot b - ca \cdot a &= 0 \\ \Rightarrow c &= \frac{a \cdot b}{a \cdot a} = \frac{a'b}{a'a} \end{aligned} \tag{57}$$

Definition 1. The projection of the vector b onto the line through a is the vector $p = ca = \frac{a'b}{a'a} a$.

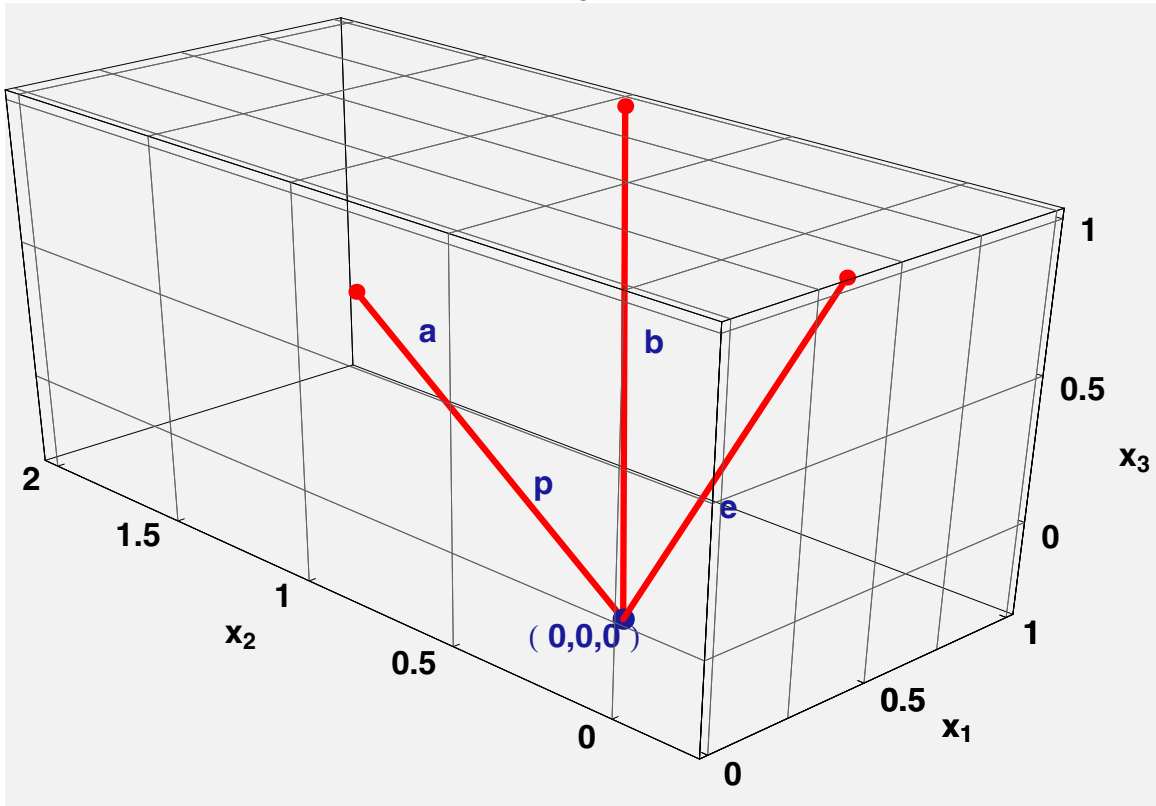
FIGURE 9. Orthogonal Vectors in \mathbb{R}^3 

2.2.2. *An Example.* Project the vector $b' = [1,1,1]$ onto the line through $a' = [1,2,0]$. First find the scalar needed to construct the projection p . For this case

$$\begin{aligned}
 a &= \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
 a'b &= [1, 2, 0] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3 \\
 a'a &= [1, 2, 0] \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = 5 \\
 c &= \frac{a'b}{a'a} = \frac{3}{5}
 \end{aligned} \tag{58}$$

The projection p is then given by ca or

FIGURE 10. Orthogonal Vectors in \mathbb{R}^3



$$p = ca = \frac{3}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \\ 0 \end{bmatrix} \tag{59}$$

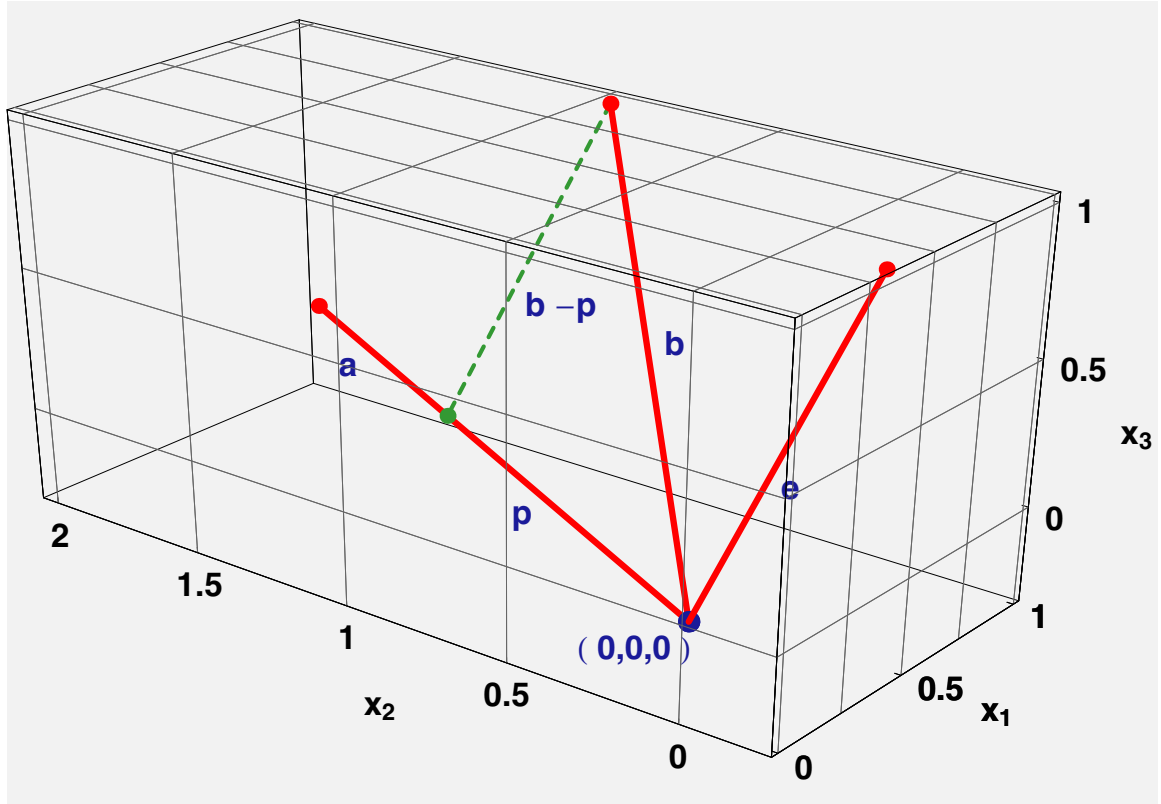
The difference between b and p (the error vector, if you will) is $b-p$. This is computed as follows.

$$e = b - p = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ 1 \end{bmatrix} \tag{60}$$

We can think of the vector b as being split into two parts, the part along the line through a , that is p , and the part perpendicular to the line through a , that is e . We can then write $b = p+e$.

2.2.3. *The projection matrix.* We can write the equation for the projection p in another useful fashion.

FIGURE 11. Projection on to a Line



$$\begin{aligned}
 p &= a c \\
 &= a \frac{a'b}{a'a} \\
 &= \frac{a a'}{a'a} b \\
 &= P b \text{ where } P = \frac{a a'}{a'a}
 \end{aligned} \tag{61}$$

We call the matrix P the projection matrix.

2.2.4. *Example projection matrix.* Consider the line through the point a where $a = [1, 2, 2]$. We find the projection matrix for this line as follows. First find $a'a$.

$$a = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad a'a = [1, 2, 2] \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = 9 \tag{62}$$

Then find aa' .

$$aa' = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} [1, 2, 2] = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix} \tag{63}$$

Then construct P as follows

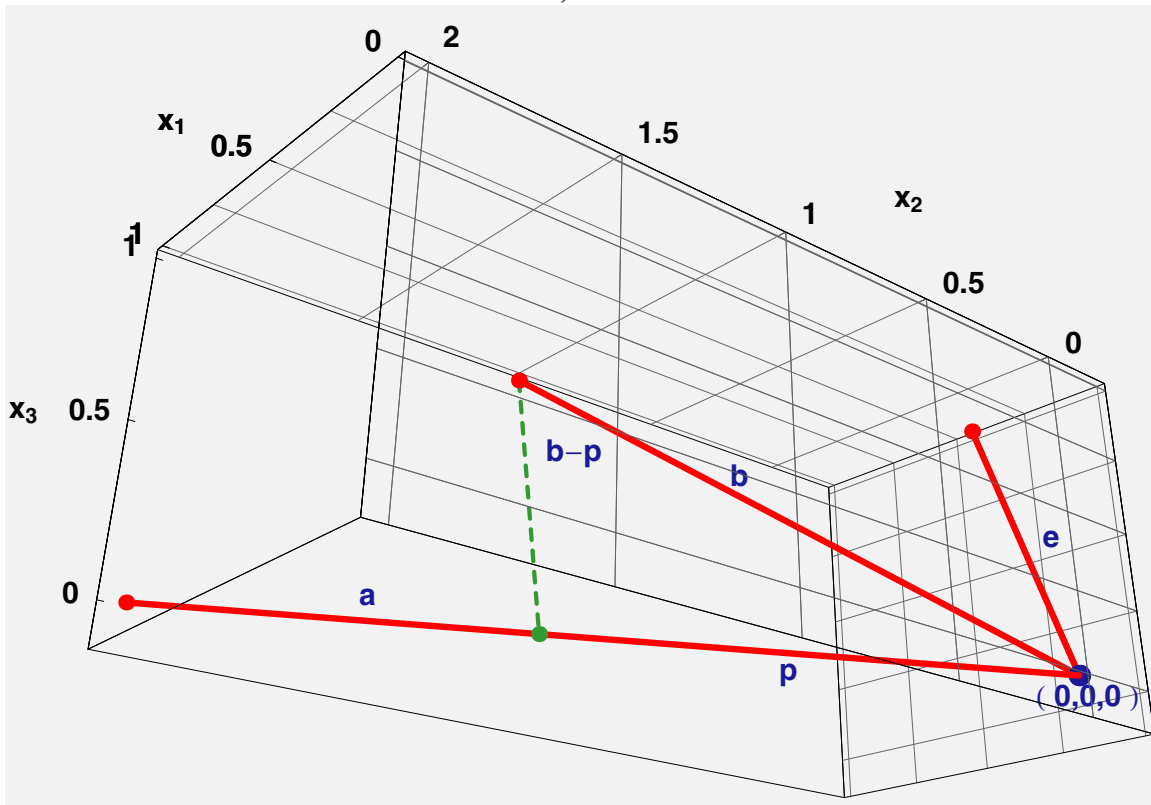
$$P = \begin{bmatrix} \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \end{bmatrix} \tag{64}$$

Consider the projection of the point $b = [1, 1, 1]$ onto the line through a . We obtain

$$\begin{aligned} p &= Pb \\ &= \begin{bmatrix} \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{4}{9} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{9} \\ \frac{10}{9} \\ \frac{10}{9} \end{bmatrix} \end{aligned} \tag{65}$$

This is represented in figure 12.

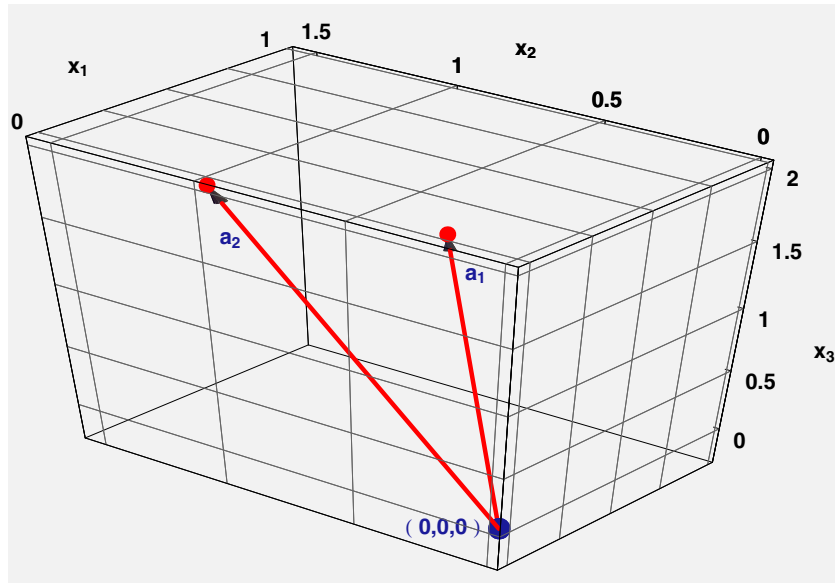
FIGURE 12. Projection on to a Line



2.3. Projection onto a more general subspace.

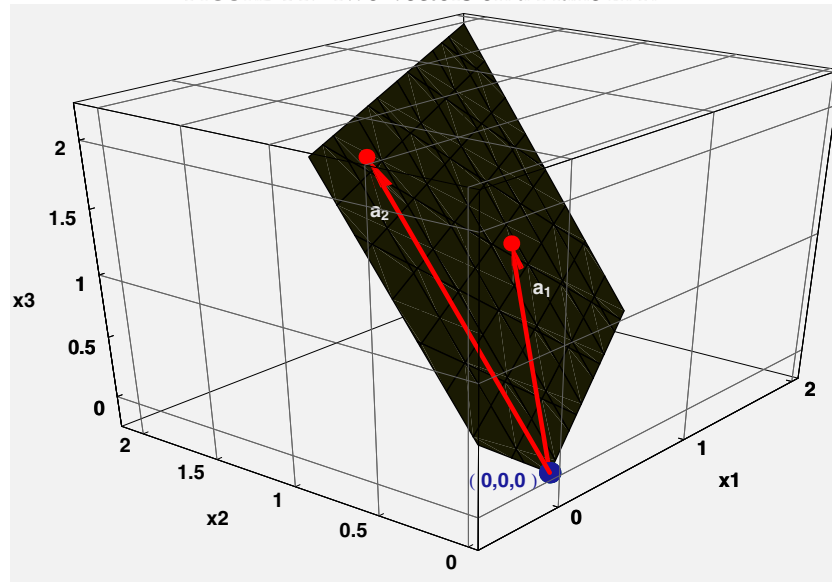
2.3.1. *Idea and definition.* Consider a vector $b' = [b_1, b_2, \dots, b_m]$ in R^m . Then consider a set of n $m \times 1$ vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$. The goal is to find the combination $c_1 \vec{a}_1 + c_2 \vec{a}_2 + \dots + c_n \vec{a}_n$ that is closest to the vector b . This is called the projection of vector b onto the n -dimensional subspace spanned by the \vec{a} 's. If we treat the vectors as columns of a matrix A , then we are looking for the vector in the column space of A that is closest to the vector b , i.e., the point c , such that b and Ac are closest. We find the point in this subspace such that the error vector $b - Ac$ is perpendicular to the subspace. This point is some linear combination of vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$. Consider figure 13 where we plot the vectors $a_1 = [1,1,1]$ and $a_2 = [0,1,2]$.

FIGURE 13. Two Vectors in R^3



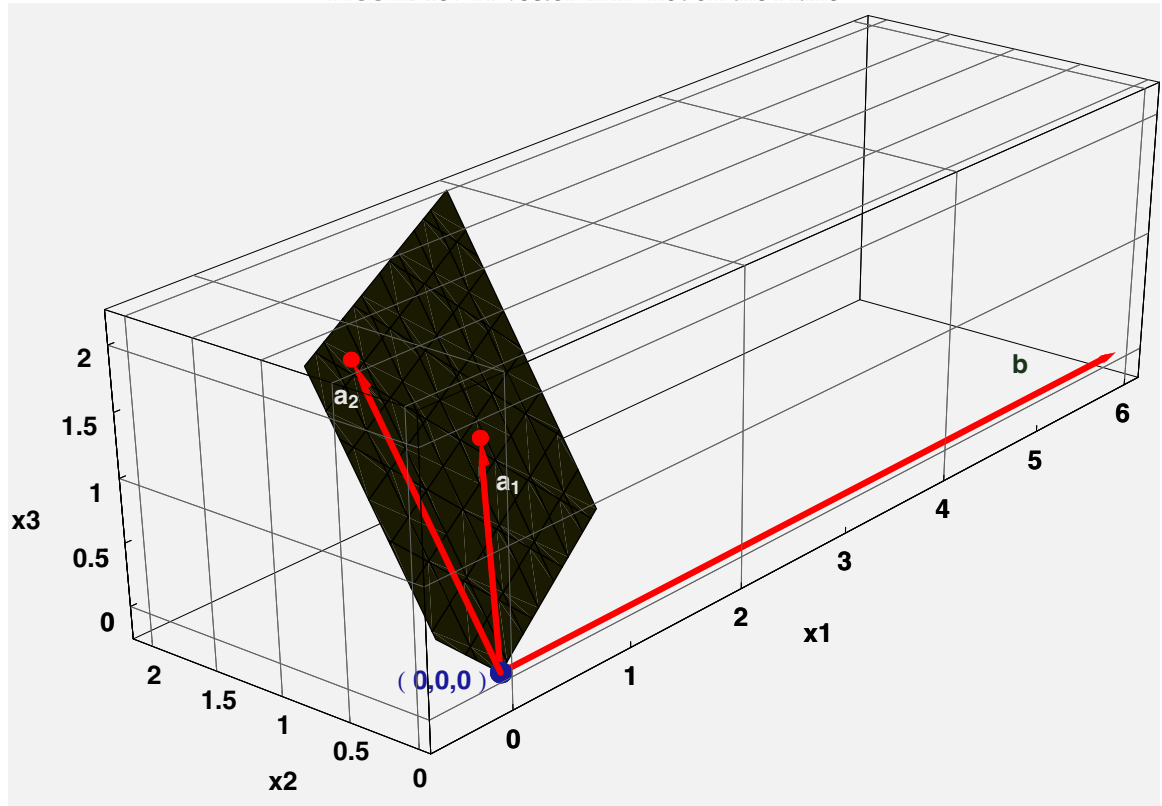
The subspace spanned by these two vectors is a plane in \mathbb{R}^3 . The plane is shown in figure 14.

FIGURE 14. Two Vectors on a Plane in \mathbb{R}^3



Now consider a third vector given by $b = [6,0,0]$. We represent the three vectors in figure 15.

FIGURE 15. A Vector in \mathbb{R}^3 not on the Plane



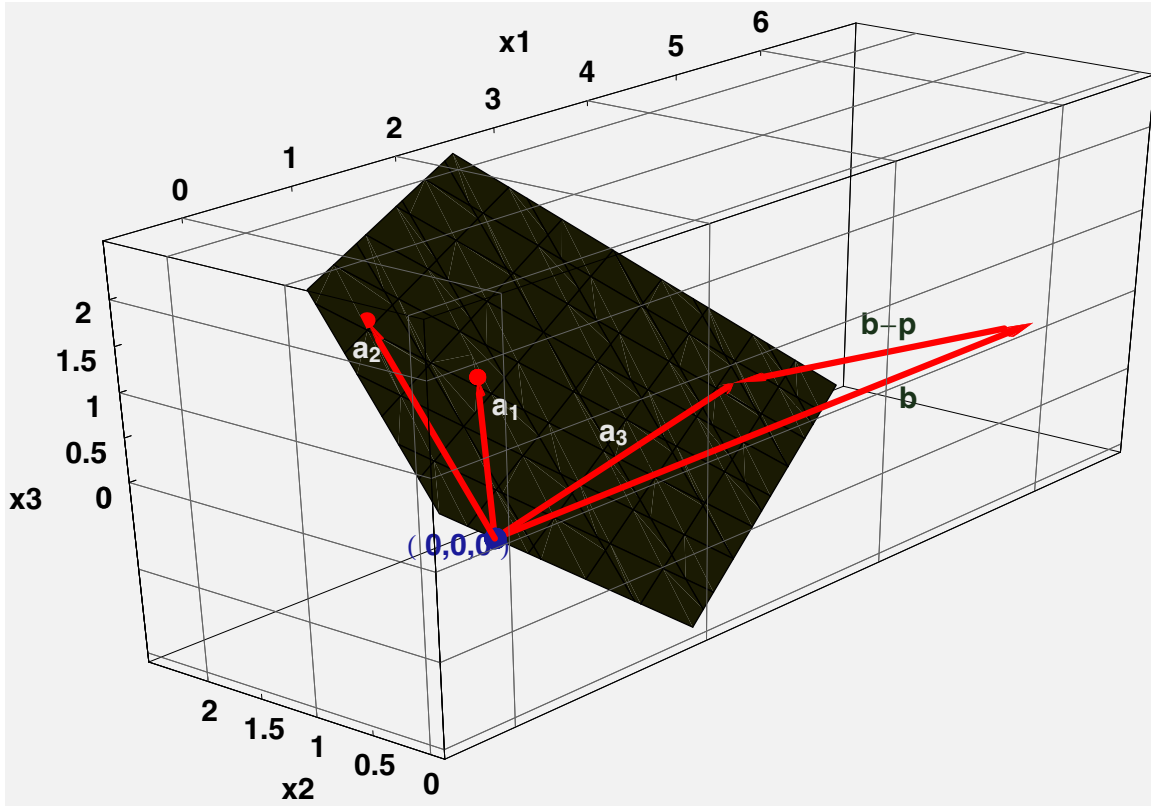
This vector does not lie in the space spanned by a_1 and a_2 . The idea is to find the point in the subspace spanned by a_1 and a_2 that is closest to the point $(6,0,0)$. The error vector $e = b - Ac$ will be perpendicular to the subspace. In figure 16 the vector (labeled a_3) in the space spanned by a_1 and a_2 that is closest to the point b is drawn. The diagram also shows the vector $b - a_1$.

2.3.2. *Formula for the projection p and the projection matrix P .* The error vector $e = b - Ac$ will be orthogonal to each of the vectors in the column space of A , i.e., it will be orthogonal to $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$. We can write this as n separate conditions as follows.

$$\begin{aligned}
 a_1'(b - Ac) &= 0 \\
 a_2'(b - Ac) &= 0 \\
 &\vdots \\
 a_n'(b - Ac) &= 0
 \end{aligned}
 \tag{66}$$

We can also write this as a matrix equation in the following manner.

FIGURE 16. Projection onto a Subspace



$$\begin{aligned}
 A'(b - Ac) &= 0 \\
 \Rightarrow A'b - A'Ac &= 0 \\
 \Rightarrow A'Ac &= A'b
 \end{aligned}
 \tag{67}$$

Equation 67 can then be solved for the coefficients c by finding the inverse of $A'A$. This is given by

$$\begin{aligned}
 A'Ac &= A'b \\
 \Rightarrow c &= (A'A)^{-1}A'b
 \end{aligned}
 \tag{68}$$

The projection $p = Ac$ is then

$$\begin{aligned}
 p &= Ac \\
 &= A(A'A)^{-1}A'b
 \end{aligned}
 \tag{69}$$

The projection matrix that produces $p = Pb$ is

$$P = A(A'A)^{-1}A'
 \tag{70}$$

Given any $m \times n$ matrix A and any vector b in \mathbb{R}^m , we can find the projection of b onto the column space of A by premultiplying the vector b by this projection matrix. The vector b is split into

two components by the projection. The first component is the vector p which lies in the subspace spanned by the columns of the A matrix. The second component is the vector e , which is the difference between p and b . The idea is that we get to b by first getting to p and then moving in a perpendicular direction from p to b as in figure 16.

2.3.3. *Example.* Project the vector $b' = [6,0,0]$ onto the space spanned by the vectors $a_1 = [1,1,1]$ and $a_2 = [0,1,2]$. The matrices are as follows.

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, A' = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}, b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \quad (71)$$

$A'A$ and $A'b$ are as follows

$$A'A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \quad (72)$$

$$A'b = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

To compute c we need to find the inverse of $A'A$. We can do so as follows

$$\begin{aligned} (A'A)^{-1} &= \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}^{-1} = \frac{1}{5} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \\ &= \begin{bmatrix} \frac{5}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{aligned} \quad (73)$$

We can now solve for c

$$\begin{aligned} c &= (A'A)^{-1} A'b = \begin{bmatrix} \frac{5}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 6 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ -3 \end{bmatrix} \end{aligned} \quad (74)$$

The projection p is then given by Ac

$$\begin{aligned} p &= Ac \\ &= \begin{bmatrix} 1 & -0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} \end{aligned} \quad (75)$$

and while $e = b - p$ is

$$e = b - p = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \tag{76}$$

This solves the problem for this particular vector b . For the more general problem is it useful to construct the projection matrix P .

$$\begin{aligned} P &= A(A'A)^{-1}A' \\ &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{5}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{5}{6} & -\frac{1}{2} \\ \frac{1}{3} & 0 \\ -\frac{1}{6} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix} \end{aligned} \tag{77}$$

Notice that e is perpendicular to both a_1 and a_2 .

$$a_1' e = [1, 1, 1] \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 0 \tag{78}$$

$$a_2' e = [0, 1, 2] \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 0$$

We can think of the vector b as being split into two parts, the part in the space spanned by the columns of A , that is p , and the part perpendicular to this subspace, that is e . We can then write $b = p+e$. Also note that $Pb = p$.

$$Pb = p$$

$$\begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} \tag{79}$$

2.4. Projections, linear independence, and inverses. To find the coefficients c , the projection p , and the projection matrix P , we need to be able to invert the matrix, $A'A$.

Theorem 3. *The matrix $A'A$ is invertible if and only if the columns of A are linearly independent.*