Create a leading 1 in the first row:

$$\begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 2 & 4 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$$

Create zeros under the first leading 1:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix} \xrightarrow{-R_1+R_2} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 4 \\ 3 & 6 & 9 & 12 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 4 \\ 3 & 6 & 9 & 12 \end{bmatrix} \xrightarrow{-3R_1+R_3} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The system is consistent, however, there are only 2 nonzero rows but 3 unknown variables. This means that the solution set will contain $3 - 2 = 1$ **free parameter**. The second row in the augmented matrix is equivalent to the equation:

$$x_3 = 4.$$

The first row is equivalent to the equation:

$$x_1 + 2x_2 + 3x_3 = 4$$

and after substituting $x_3 = 4$ we obtain

$$x_1 + 2x_2 = -8.$$

We now must choose one of the variables x_1 or x_2 to be a parameter, say t, and solve for the remaining variable. If we set $x_2 = t$ then from $x_1 + 2x_2 = -8$ we obtain that

$$x_1 = -8 - 2t.$$

We can therefore write the solution set for the linear system as

$$\begin{aligned} x_1 &= -8 - 2t \\ x_2 &= t \\ x_3 &= 4 \end{aligned} \tag{2.1}$$

where t can be any real number. If we had chosen x_1 to be the parameter, say $x_1 = t$, then the solution set can be written as

$$\begin{aligned} x_1 &= t \\ x_2 &= -4 - \tfrac{1}{2}t \\ x_3 &= 4 \end{aligned} \tag{2.2}$$

Although (2.1) and (2.2) are two different parameterizations, they both give the same solution set. □

In general, if a linear system has n unknown variables and the row reduced augmented matrix has r leading entries, then the number of free parameters d in the solution set is

$$d = n - r.$$

Thus, when performing back substitution, we will have to set d of the unknown variables to arbitrary parameters. In the previous example, there are $n = 3$ unknown variables and the row reduced augmented matrix contained $r = 2$ leading entries. The number of free parameters was therefore

$$d = n - r = 3 - 2 = 1.$$

Because the number of leading entries r in the row reduced coefficient matrix determine the number of free parameters, we will refer to r as the **rank** of the coefficient matrix:

$$r = \operatorname{rank}(\mathbf{A}).$$

Later in the course, we will give a more geometric interpretation to $\operatorname{rank}(\mathbf{A})$.

Example 2.4. Solve the linear system represented by the augmented matrix

$$\begin{bmatrix} 1 & -7 & 2 & -5 & 8 & 10 \\ 0 & 1 & -3 & 3 & 1 & -5 \\ 0 & 0 & 0 & 1 & -1 & 4 \end{bmatrix}$$

Solution. The number of unknowns is $n = 5$ and the augmented matrix has rank $r = 3$ (leading entries). Thus, the solution set is parameterized by $d = 5 - 3 = 2$ free variables, call them t and s. The last equation of the augmented matrix is $x_4 - x_5 = 4$. We choose x_5 to be the first parameter so we set $x_5 = t$. Therefore, $x_4 = 4 + t$. The second equation of the augmented matrix is

$$x_2 - 3x_3 + 3x_4 + x_5 = -5$$

and the unassigned variables are x_2 and x_3. We choose x_3 to be the second parameter, say $x_3 = s$. Then

$$\begin{aligned} x_2 &= -5 + 3x_3 - 3x_4 - x_5 \\ &= -5 + 3s - 3(4 + t) - t \\ &= -17 - 4t + 3s. \end{aligned}$$

We now use the first equation of the augmented matrix to write x_1 in terms of the other variables:

$$\begin{aligned} x_1 &= 10 + 7x_2 - 2x_3 + 5x_4 - 8x_5 \\ &= 10 + 7(-17 - 4t + 3s) - 2s + 5(4 + t) - 8t \\ &= -89 - 31t + 19s \end{aligned}$$

Thus, the solution set is

$$
\begin{aligned}
x_1 &= -89 - 31t + 19s\\
x_2 &= -17 - 4t + 3s\\
x_3 &= s\\
x_4 &= 4 + t\\
x_5 &= t
\end{aligned}
$$

where t and s are arbitrary real numbers.. Choose arbitrary numbers for t and s and substitute the corresponding list $(x_1, x_2, \ldots, x_5)$ into the system of equations to verify that it is a solution. □

2.3 Existence and uniqueness of solutions

The REF or RREF of an augmented matrix leads to three distinct possibilities for the solution set of a linear system.

Theorem 2.5: Let $[\mathbf{A}\ \mathbf{b}]$ be the augmented matrix of a linear system. One of the following distinct possibilities will occur:

1. The augmented matrix will contain an inconsistent row.
2. All the rows of the augmented matrix are consistent and there are no free parameters.
3. All the rows of the augmented matrix are consistent and there are $d \geq 1$ variables that must be set to arbitrary parameters

In Case 1., the linear system is inconsistent and thus has no solution. In Case 2., the linear system is consistent and has only one (and thus **unique**) solution. This case occurs when $r = \text{rank}(\mathbf{A}) = n$ since then the number of free parameters is $d = n - r = 0$. In Case 3., the linear system is consistent and has infinitely many solutions. This case occurs when $r < n$ and thus $d = n - r > 0$ is the number of free parameters.

After this lecture you should know the following:

- what the REF is and how to compute it
- what the RREF is and how to compute it
- how to solve linear systems using row reduction (**Practice!!!**)
- how to identify when a linear system is inconsistent
- how to identify when a linear system is consistent
- what is the rank of a matrix
- how to compute the number of free parameters in a solution set
- what are the three possible cases for the solution set of a linear system (Theorem 2.5)

Lecture 3

Vector Equations

In this lecture, we introduce vectors and vector equations. Specifically, we introduce the linear combination problem which simply asks whether it is possible to express one vector in terms of other vectors; we will be more precise in what follows. As we will see, solving the linear combination problem reduces to solving a linear system of equations.

3.1 Vectors in $\mathbb{R}^n$

Recall that a **column vector** in $\mathbb{R}^n$ is a $n \times 1$ matrix. From now on, we will drop the "column" descriptor and simply use the word **vectors**. It is important to emphasize that a vector in $\mathbb{R}^n$ is simply a list of n numbers; you are safe (and highly encouraged!) to forget the idea that a vector is an object with an arrow. Here is a vector in $\mathbb{R}^2$:

$$\mathbf{v} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Here is a vector in $\mathbb{R}^3$:

$$\mathbf{v} = \begin{bmatrix} -3 \\ 0 \\ 11 \end{bmatrix}.$$

Here is a vector in $\mathbb{R}^6$:

$$\mathbf{v} = \begin{bmatrix} 9 \\ 0 \\ -3 \\ 6 \\ 0 \\ 3 \end{bmatrix}.$$

To indicate that $\mathbf{v}$ is a vector in $\mathbb{R}^n$, we will use the notation $\mathbf{v} \in \mathbb{R}^n$. The mathematical symbol $\in$ means "is an element of". When we write vectors within a paragraph, we will write them using list notation instead of column notation, e.g., $\mathbf{v} = (-1, 4)$ instead of $\mathbf{v} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$.

We can add/subtract vectors, and multiply vectors by numbers or **scalars**. For example, here is the addition of two vectors:

$$\begin{bmatrix} 0 \\ -5 \\ 9 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -8 \\ 9 \\ 3 \end{bmatrix}.$$

And the multiplication of a scalar with a vector:

$$3 \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \\ 15 \end{bmatrix}.$$

And here are both operations combined:

$$-2 \begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} -2 \\ 9 \\ 4 \end{bmatrix} = \begin{bmatrix} -8 \\ 16 \\ -6 \end{bmatrix} + \begin{bmatrix} -6 \\ 27 \\ 12 \end{bmatrix} = \begin{bmatrix} -14 \\ 43 \\ 6 \end{bmatrix}.$$

These operations constitute "the algebra" of vectors. As the following example illustrates, vectors can be used in a natural way to represent the solution of a linear system.

Example 3.1. Write the general solution in vector form of the linear system represented by the augmented matrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & -7 & 2 & -5 & 8 & 10 \\ 0 & 1 & -3 & 3 & 1 & -5 \\ 0 & 0 & 0 & 1 & -1 & 4 \end{bmatrix}$$

Solution. The number of unknowns is $n = 5$ and the associated coefficient matrix $\mathbf{A}$ has rank $r = 3$. Thus, the solution set is parametrized by $d = n - r = 2$ parameters. This system was considered in Example 2.4 and the general solution was found to be

$$\begin{aligned} x_1 &= -89 - 31t_1 + 19t_2 \\ x_2 &= -17 - 4t_1 + 3t_2 \\ x_3 &= t_2 \\ x_4 &= 4 + t_1 \\ x_5 &= t_1 \end{aligned}$$

where t_1 and t_2 are arbitrary real numbers. The solution in vector form therefore takes the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -89 - 31t_1 + 19t_2 \\ -17 - 4t_1 + 3t_2 \\ t_2 \\ 4 + t_1 \\ t_1 \end{bmatrix} = \begin{bmatrix} -89 \\ -17 \\ 0 \\ 4 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -31 \\ -4 \\ 0 \\ 1 \\ 1 \end{bmatrix} + t_2 \begin{bmatrix} 19 \\ 3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

□

A fundamental problem in **linear algebra** is solving vector equations for an unknown vector. As an example, suppose that you are given the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -2 \\ 9 \\ 4 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} -14 \\ 43 \\ 6 \end{bmatrix},$$

and asked to find numbers x_1 and x_2 such that $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \mathbf{b}$, that is,

$$x_1 \begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 9 \\ 4 \end{bmatrix} = \begin{bmatrix} -14 \\ 43 \\ 6 \end{bmatrix}.$$

Here the unknowns are the scalars x_1 and x_2. After some guess and check, we find that $x_1 = -2$ and $x_2 = 3$ is a solution to the problem since

$$-2 \begin{bmatrix} 4 \\ -8 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} -2 \\ 9 \\ 4 \end{bmatrix} = \begin{bmatrix} -14 \\ 43 \\ 6 \end{bmatrix}.$$

In some sense, the vector $\mathbf{b}$ is a combination of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. This motivates the following definition.

Definition 3.2: Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ be vectors in $\mathbb{R}^n$. A vector $\mathbf{b}$ is said to be a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ if there exists scalars $x_1, x_2, \ldots, x_p$ such that $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$.

The scalars in a linear combination are called the **coefficients** of the linear combination. As an example, given the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -1 \\ 5 \\ 6 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} -3 \\ 0 \\ -27 \end{bmatrix}$$

you can verify (and you should!) that

$$3\mathbf{v}_1 + 4\mathbf{v}_2 - 2\mathbf{v}_3 = \mathbf{b}.$$

Therefore, we can say that $\mathbf{b}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ with coefficients $x_1 = 3$, $x_2 = 4$, and $x_3 = -2$.

3.2 The linear combination problem

The linear combination problem is the following:

Problem: Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and $\mathbf{b}$, is $\mathbf{b}$ a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$?

For example, say you are given the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \; \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

and also

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

Does there exist scalars x_1, x_2, x_3 such that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}? \tag{3.1}$$

For obvious reasons, equation (3.1) is called a **vector equation** and the unknowns are x_1, x_2, and x_3. To gain some intuition with the linear combination problem, let's do an example by inspection.

Example 3.3. Let $\mathbf{v}_1 = (1, 0, 0)$, let $\mathbf{v}_2 = (0, 0, 1)$, let $\mathbf{b}_1 = (0, 2, 0)$, and let $\mathbf{b}_2 = (-3, 0, 7)$. Are $\mathbf{b}_1$ and $\mathbf{b}_2$ linear combinations of $\mathbf{v}_1, \mathbf{v}_2$?

Solution. For any scalars x_1 and x_2

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

and thus no, $\mathbf{b}_1$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. On the other hand, by inspection we have that

$$-3\mathbf{v}_1 + 7\mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 7 \end{bmatrix} = \mathbf{b}_2$$

and thus yes, $\mathbf{b}_2$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. These examples, of low dimension, were more-or-less obvious. Going forward, we are going to need a systematic way to solve the linear combination problem that does not rely on pure inspection. □

We now describe how the linear combination problem is connected to the problem of solving a system of linear equations. Consider again the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \; \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \; \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

Does there exist scalars x_1, x_2, x_3 such that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}? \tag{3.2}$$

First, let's expand the left-hand side of equation (3.2):

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \begin{bmatrix} x_1 \\ 2x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2x_3 \\ x_3 \\ 2x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 2x_3 \\ 2x_1 + x_2 + x_3 \\ x_1 + 2x_3 \end{bmatrix}.$$

We want equation (3.2) to hold so let's equate the expansion $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$ with $\mathbf{b}$. In other words, set

$$\begin{bmatrix} x_1 + x_2 + 2x_3 \\ 2x_1 + x_2 + x_3 \\ x_1 + 2x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

Comparing component-by-component in the above relationship, we seek scalars x_1, x_2, x_3 satisfying the equations

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 1 \\ x_1 + 2x_3 &= -2. \end{aligned} \tag{3.3}$$

This is just a linear system consisting of $m = 3$ equations and $n = 3$ unknowns! Thus, the linear combination problem can be solved by solving a system of linear equations for the unknown scalars x_1, x_2, x_3. We know how to do this. In this case, the augmented matrix of the linear system (3.3) is

$$[\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & 0 & 2 & -2 \end{bmatrix}$$

Notice that the 1st column of $\mathbf{A}$ is just $\mathbf{v}_1$, the second column is $\mathbf{v}_2$, and the third column is $\mathbf{v}_3$, in other words, the augment matrix is

$$[\mathbf{A}\ \mathbf{b}] = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b} \end{bmatrix}$$

Applying the row reduction algorithm, the solution is

$$x_1 = 0,\ x_2 = 2,\ x_3 = -1$$

and thus these coefficients solve the linear combination problem. In other words,

$$0\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{b}$$

In this case, there is only one solution to the linear system, so $\mathbf{b}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in only one (or unique) way. You should verify these computations.

We summarize the previous discussion with the following:

> The problem of determining if a given vector $\mathbf{b}$ is a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ is equivalent to solving the linear system of equations with augmented matrix
>
> $$\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_p & \mathbf{b} \end{bmatrix}.$$

Applying the existence and uniqueness Theorem 2.5, the only three possibilities to the linear combination problem are:

1. If the linear system is inconsistent then $\mathbf{b}$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, i.e., there does not exist scalars $x_1, x_2, \ldots, x_p$ such that $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$.

2. If the linear system is consistent and the solution is unique then $\mathbf{b}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in only one way.

3. If the the linear system is consistent and the solution set has free parameters, then $\mathbf{b}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in infinitely many ways.

Example 3.4. Is the vector $\mathbf{b} = (7, 4, -3)$ a linear combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}?$$

Solution. Form the augmented matrix:

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix}$$

The RREF of the augmented matrix is

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

and therefore the solution is $x_1 = 3$ and $x_2 = 2$. Therefore, yes, $\mathbf{b}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2$:

$$3\mathbf{v}_1 + 2\mathbf{v}_2 = 3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} = \mathbf{b}$$

Notice that the solution set does not contain any free parameters because $n = 2$ (unknowns) and $r = 2$ (rank) and so $d = 0$. Therefore, the above linear combination is the only way to write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. $\square$

Example 3.5. Is the vector $\mathbf{b} = (1, 0, 1)$ a linear combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}?$$

Solution. The augmented matrix of the corresponding linear system is

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

After row reducing we obtain that

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

The last row is inconsistent, and therefore the linear system does not have a solution. Therefore, no, $\mathbf{b}$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. □

Example 3.6. Is the vector $\mathbf{b} = (8, 8, 12)$ a linear combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 4 \\ 9 \end{bmatrix}?$$

Solution. The augmented matrix is

$$\begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 2 & 4 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix} \xrightarrow{\text{REF}} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The system is consistent and therefore $\mathbf{b}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. In this case, the solution set contains $d = 1$ free parameters and therefore, it is possible to write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in infinitely many ways. In terms of the parameter t, the solution set is

$$\begin{aligned} x_1 &= -8 - 2t \\ x_2 &= t \\ x_3 &= 4 \end{aligned}$$

Choosing any t gives scalars that can be used to write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. For example, choosing $t = 1$ we obtain $x_1 = -10$, $x_2 = 1$, and $x_3 = 4$, and you can verify that

$$-10\mathbf{v}_1 + \mathbf{v}_2 + 4\mathbf{v}_3 = -10 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} + 4 \begin{bmatrix} 6 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \\ 12 \end{bmatrix} = \mathbf{b}$$

Or, choosing $t = -2$ we obtain $x_1 = -4$, $x_2 = -2$, and $x_3 = 4$, and you can verify that

$$-4\mathbf{v}_1 - 2\mathbf{v}_2 + 4\mathbf{v}_3 = -4 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} + 4 \begin{bmatrix} 6 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \\ 12 \end{bmatrix} = \mathbf{b}$$

□

We make a few important observations on linear combinations of vectors. Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, there are certain vectors $\mathbf{b}$ that can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in an obvious way. The zero vector $\mathbf{b} = \mathbf{0}$ can always be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$:

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p.$$

Each $\mathbf{v}_i$ itself can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, for example,

$$\mathbf{v}_2 = 0\mathbf{v}_1 + (1)\mathbf{v}_2 + 0\mathbf{v}_3 + \cdots + 0\mathbf{v}_p.$$

More generally, any scalar multiple of $\mathbf{v}_i$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, for example,

$$x\mathbf{v}_2 = 0\mathbf{v}_1 + x\mathbf{v}_2 + 0\mathbf{v}_3 + \cdots + 0\mathbf{v}_p.$$

By varying the coefficients $x_1, x_2, \ldots, x_p$, we see that there are infinitely many vectors $\mathbf{b}$ that can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$. The "space" of all the possible linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ has a name, which we introduce next.

3.3 The span of a set of vectors

Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$, we have been considering the problem of whether or not a given vector $\mathbf{b}$ is a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. We now take another point of view and instead consider the idea of **generating** all vectors that are a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. So how do we generate a vector that is guaranteed to be a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$? For example, if $\mathbf{v}_1 = (2, 1, 3)$, $\mathbf{v}_2 = (4, 2, 6)$ and $\mathbf{v}_3 = (6, 4, 9)$ then

$$-10\mathbf{v}_1 + \mathbf{v}_2 + 4\mathbf{v}_3 = -10\begin{bmatrix}2\\1\\3\end{bmatrix} + \begin{bmatrix}4\\2\\6\end{bmatrix} + 4\begin{bmatrix}6\\4\\9\end{bmatrix} = \begin{bmatrix}8\\8\\12\end{bmatrix}.$$

Thus, by construction, the vector $\mathbf{b} = (8, 8, 12)$ is a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. This discussion leads us to the following definition.

Definition 3.7: Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ be vectors. The set of all vectors that are a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ is called the **span** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, and we denote it by

$$S = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}.$$

By definition, the span of a set of vectors is a collection of vectors, or a **set** of vectors. If $\mathbf{b}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ then $\mathbf{b}$ is an **element** of the set $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$, and we write this as

$$\mathbf{b} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}.$$

By definition, writing that $\mathbf{b} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ implies that there exists scalars $x_1, x_2, \ldots, x_p$ such that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}.$$

Even though $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is an infinite set of vectors, it is not necessarily true that it is the whole space $\mathbb{R}^n$.

The set $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is just a collection of infinitely many vectors but it has some geometric structure. In $\mathbb{R}^2$ and $\mathbb{R}^3$ we can visualize $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. In $\mathbb{R}^2$, the span of a single nonzero vector, say $\mathbf{v} \in \mathbb{R}^2$, is a line through the origin in the direction of $\mathbf{v}$, see Figure 3.1.

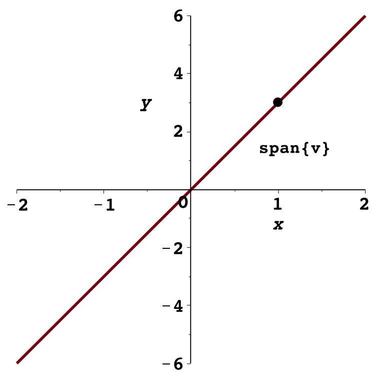

Figure 3.1: The span of a single non-zero vector in $\mathbb{R}^2$.

In $\mathbb{R}^2$, the span of two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$ that are not multiples of each other is all of $\mathbb{R}^2$. That is, $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathbb{R}^2$. For example, with $\mathbf{v}_1 = (1, 0)$ and $\mathbf{v}_2 = (0, 1)$, it is true that $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \mathbb{R}^2$. In $\mathbb{R}^3$, the span of two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$ that are not multiples of each other is a plane through the origin containing $\mathbf{v}_1$ and $\mathbf{v}_2$, see Figure 3.2. In $\mathbb{R}^3$, the

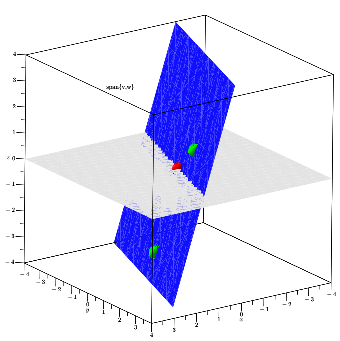

Figure 3.2: The span of two vectors, not multiples of each other, in $\mathbb{R}^3$.

span of a single vector is a line through the origin, and the span of three vectors that do not depend on each other (we will make this precise soon) is all of $\mathbb{R}^3$.

Example 3.8. Is the vector $\mathbf{b} = (7, 4, -3)$ in the span of the vectors $\mathbf{v}_1 = (1, -2, -5), \mathbf{v}_2 = (2, 5, 6)$? In other words, is $\mathbf{b} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$?

Solution. By definition, $\mathbf{b}$ is in the span of $\mathbf{v}_1$ and $\mathbf{v}_2$ if there exists scalars x_1 and x_2 such that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \mathbf{b},$$

that is, if $\mathbf{b}$ can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. From our previous discussion on the linear combination problem, we must consider the augmented matrix $\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{b}\end{bmatrix}$. Using row reduction, the augmented matrix is consistent and there is only one solution (see Example 3.4). Therefore, yes, $\mathbf{b} \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and the linear combination is unique. $\square$

Example 3.9. Is the vector $\mathbf{b} = (1, 0, 1)$ in the span of the vectors $\mathbf{v}_1 = (1, 0, 2), \mathbf{v}_2 = (0, 1, 0), \mathbf{v}_3 = (2, 1, 4)$?

Solution. From Example 3.5, we have that

$$\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b}\end{bmatrix} \xrightarrow{\text{REF}} \begin{bmatrix}1 & 0 & 2 & 1\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}$$

The last row is inconsistent and therefore $\mathbf{b}$ is not in $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. $\square$

Example 3.10. Is the vector $\mathbf{b} = (8, 8, 12)$ in the span of the vectors $\mathbf{v}_1 = (2, 1, 3), \mathbf{v}_2 = (4, 2, 6), \mathbf{v}_3 = (6, 4, 9)$?

Solution. From Example 3.6, we have that

$$\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b}\end{bmatrix} \xrightarrow{\text{REF}} \begin{bmatrix}1 & 2 & 3 & 4\\ 0 & 0 & 1 & 4\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

The system is consistent and therefore $\mathbf{b} \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. In this case, the solution set contains $d = 1$ free parameters and therefore, it is possible to write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in infinitely many ways. $\square$

Example 3.11. Answer the following with True or False, and explain your answer.

(a) The vector $\mathbf{b} = (1, 2, 3)$ is in the span of the set of vectors

$$\left\{ \begin{bmatrix}-1\\ 3\\ 0\end{bmatrix}, \begin{bmatrix}2\\ -7\\ 0\end{bmatrix}, \begin{bmatrix}4\\ -5\\ 0\end{bmatrix} \right\}.$$

(b) The solution set of the linear system whose augmented matrix is $\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b}\end{bmatrix}$ is the same as the solution set of the vector equation $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$.

(c) Suppose that the augmented matrix $\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{b}\end{bmatrix}$ has an inconsistent row. Then either $\mathbf{b}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ or $\mathbf{b} \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(d) The span of the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ (at least one of which is nonzero) contains only the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and the zero vector $\mathbf{0}$.

After this lecture you should know the following:

- what a vector is
- what a linear combination of vectors is
- what the linear combination problem is
- the relationship between the linear combination problem and the problem of solving linear systems of equations
- how to solve the linear combination problem
- what the span of a set of vectors is
- the relationship between what it means for a vector $\mathbf{b}$ to be in the span of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ and the problem of writing $\mathbf{b}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$
- the geometric interpretation of the span of a set of vectors

Lecture 4

The Matrix Equation $\mathbf{Ax} = \mathbf{b}$

In this lecture, we introduce the operation of matrix-vector multiplication and how it relates to the linear combination problem.

4.1 Matrix-vector multiplication

We begin with the definition of matrix-vector multiplication.

Definition 4.1: Given a matrix $\mathbf{A} \in M_{m\times n}$ and a vector $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

we define the product of $\mathbf{A}$ and $\mathbf{x}$ as the vector $\mathbf{Ax}$ in $\mathbb{R}^m$ given by

$$\mathbf{Ax} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

For the product $\mathbf{Ax}$ to be well-defined, the number of columns of $\mathbf{A}$ must equal the number of components of $\mathbf{x}$. Another way of saying this is that the outer dimension of $\mathbf{A}$ must equal the inner dimension of $\mathbf{x}$:

$$(m \times n) \cdot (n \times 1) \to m \times 1$$

Example 4.2. Compute $\mathbf{Ax}$.

(a)

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ -4 \\ -3 \\ 8 \end{bmatrix}$$

(b)

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

(c)

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

Solution. We compute:

(a)

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} 1 & -1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \\ -3 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} (1)(2) + (-1)(-4) + (3)(-3) + (0)(8) \end{bmatrix} = \begin{bmatrix} -3 \end{bmatrix} \end{aligned}$$

(b)

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} 3 & 3 & -2 \\ 4 & -4 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} (3)(1) + (3)(0) + (-2)(-1) \\ (4)(1) + (-4)(0) + (-1)(-1) \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ 5 \end{bmatrix} \end{aligned}$$

(c)

$$
\begin{aligned}
\mathbf{A}\mathbf{x} &= \begin{bmatrix} -1 & 1 & 0 \\ 4 & 1 & -2 \\ 3 & -3 & 3 \\ 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} (-1)(-1) + (1)(2) + (0)(-2) \\ (4)(-1) + (1)(2) + (-2)(-2) \\ (3)(-1) + (-3)(2) + (3)(-2) \\ (0)(-1) + (-2)(2) + (-3)(-2) \end{bmatrix} \\
&= \begin{bmatrix} 3 \\ 2 \\ -15 \\ 2 \end{bmatrix}
\end{aligned}
$$

□

We now list two important properties of matrix-vector multiplication.

Theorem 4.3: Let $\mathbf{A}$ be an $m \times n$ a matrix.
(a) For any vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ it holds that

$$\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v}.$$

(b) For any vector $\mathbf{u}$ and scalar c it holds that

$$\mathbf{A}(c\mathbf{u}) = c(\mathbf{A}\mathbf{u}).$$

Example 4.4. For the given data, verify that the properties of Theorem 4.3 hold:

$$\mathbf{A} = \begin{bmatrix} 3 & -3 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad c = -2.$$

4.2 Matrix-vector multiplication and linear combinations

Recall the general definition of matrix-vector multiplication $\mathbf{A}\mathbf{x}$ is

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \tag{4.1}$$

There is an important way to decompose matrix-vector multiplication involving a linear combination. To see how, let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ denote the columns of $\mathbf{A}$ and consider the following linear combination:

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \begin{bmatrix} x_1a_{11} \\ x_1a_{21} \\ \vdots \\ x_1a_{m1} \end{bmatrix} + \begin{bmatrix} x_2a_{12} \\ x_2a_{22} \\ \vdots \\ x_2a_{m2} \end{bmatrix} + \cdots + \begin{bmatrix} x_na_{1n} \\ x_na_{2n} \\ \vdots \\ x_na_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}. \tag{4.2}$$

We observe that expressions (4.1) and (4.2) are equal! Therefore, if $\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ then

$$\mathbf{Ax} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n.$$

In summary, the vector $\mathbf{Ax}$ is a linear combination of the columns of $\mathbf{A}$ where the scalar in the linear combination are the components of $\mathbf{x}$! This (important) observation gives an alternative way to compute $\mathbf{Ax}$.

Example 4.5. Given

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

compute $\mathbf{Ax}$ in two ways: (1) using the original Definition 4.1, and (2) as a linear combination of the columns of $\mathbf{A}$.

4.3 The matrix equation problem

As we have seen, with a matrix $\mathbf{A}$ and any vector $\mathbf{x}$, we can produce a new output vector via the multiplication $\mathbf{Ax}$. If $\mathbf{A}$ is a $m \times n$ matrix then we must have $\mathbf{x} \in \mathbb{R}^n$ and the output vector $\mathbf{Ax}$ is in $\mathbb{R}^m$. We now introduce the following problem:

Problem: Given a matrix $\mathbf{A} \in M_{m\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^m$, find, if possible, a vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$\mathbf{Ax} = \mathbf{b}. \tag{$\star$}$$

Equation ($\star$) is a **matrix equation** where the unknown variable is $\mathbf{x}$. If $\mathbf{u}$ is a vector such that $\mathbf{Au} = \mathbf{b}$, then we say that $\mathbf{u}$ is a solution to the equation $\mathbf{Ax} = \mathbf{b}$. For example, suppose that

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} -3 \\ 7 \end{bmatrix}.$$

Does the equation $\mathbf{Ax} = \mathbf{b}$ have a solution? Well, for any $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ we have that

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}$$

and thus any output vector $\mathbf{Ax}$ has equal entries. Since $\mathbf{b}$ does not have equal entries then the equation $\mathbf{Ax} = \mathbf{b}$ has no solution.

We now describe a systematic way to solve matrix equations. As we have seen, the vector $\mathbf{Ax}$ is a linear combination of the columns of $\mathbf{A}$ with the coefficients given by the components of $\mathbf{x}$. Therefore, the matrix equation problem is equivalent to the linear combination problem. In Lecture 2, we showed that the linear combination problem can be solved by solving a system of linear equations. Putting all this together then, if $\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$ and $\mathbf{b} \in \mathbb{R}^m$ then:

To find a vector $\mathbf{x} \in \mathbb{R}^n$ that solves the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

we solve the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n & \mathbf{b} \end{bmatrix}.$$

From now on, a system of linear equations such as

$$\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

will be written in the compact form

$$\mathbf{Ax} = \mathbf{b}$$

where $\mathbf{A}$ is the coefficient matrix of the linear system, $\mathbf{b}$ is the output vector, and $\mathbf{x}$ is the unknown vector to be solved for. We summarize our findings with the following theorem.

Theorem 4.6: Let $\mathbf{A} \in M_{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The following statements are equivalent:
(a) The equation $\mathbf{Ax} = \mathbf{b}$ has a solution.
(b) The vector $\mathbf{b}$ is a linear combination of the columns of $\mathbf{A}$.
(c) The linear system represented by the augmented matrix $\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix}$ is consistent.