1 Systems of Linear Equations

We're going to start this course with a very concrete, very algebraic problem: solving equations. As the course progresses, we will see how this problem relates to geometric and formal ideas. We will bring in ideas from geometric and formal perspectives to help us approach this problem, and see how we can use our equation-solving techniques to answer questions that arise in geometric and formal settings.

1.1 Basics of Linear Equations

A *linear equation* is an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b \tag{1}$$

where $a_1, \ldots, a_n$, and b are all real numbers, and $x_1, \ldots, x_n$ are *unknowns* or *variables*. (We might write $a_1, \ldots, a_n, b \in \mathbb{R}$; the symbol $\mathbb{R}$ stands for the real numbers, and the symbol $\in$ means "is an element of" or just "in"). We say that this equation has n unknowns.

A *system of linear equations* is a system of the form

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

with the a_{ij} and b_is all real numbers. We say this is a system of m equations in n unknowns.

Importantly, these equations are restricted to be relatively simple. In each equation we multiply each variable by some constant real number, add them together, and set that equal to some constant real number. We aren't allowed to multiply variables together, or do anything else fancy with them. This means the equations can't get too complicated, and are relatively easy to work with.

Example 1.1. A system of two linear equations in two variables is

$$\begin{aligned} 2x + y &= 3 \\ x + 5y &= -3. \end{aligned}$$

A system of two equations in three variables is

$$\begin{aligned} 5x + 2y + z &= 7 \\ 3x + 2y + z &= 6. \end{aligned}$$

A system of three equations in one variable is

$$\begin{aligned} 3x &= 3 \\ 5x &= 5 \\ x &= 2. \end{aligned}$$

We want to find *solutions* to this system of equations. Since there are n variables, a solution much be a list of n real numbers. We write $\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{R}\}$ for the set of ordered lists of n real numbers. (We sometimes call these "ordered n-tuples" or "vectors"). Thus $\mathbb{R}^1 = \mathbb{R}$ is just the set of real numbers; $\mathbb{R}^2$ is the set of ordered pairs that makes up the Cartesian plane.

An element $(x_1, \ldots, x_n) \in \mathbb{R}^n$ is a *solution* to a system of linear equations if all of the equalities hold for that collection of x_i. The *solution set* of a system of linear equations is the set of all solutions, and we say two systems are equivalent if they have the same set of solutions.

Example 1.2. The system

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

has $(2, -1) \in \mathbb{R}^2$ as a solution. We will see later that this is the only solution, and thus the set of solutions is $\{(2, -1)\}$.

The system

$$\begin{aligned} 4x + 2y + 2z &= 8 \\ 3x + 2y + z &= 6 \end{aligned}$$

has $(1, 1, 1)$ as a solution. This is not the only solution; in fact, the set of solutions is $\{(x, 2 - x, 2 - x) : x \in \mathbb{R}\}$. (This means that for each real number x, the ordered triple $(x, 2 - x, 2 - x)$ is a solution to our system). We say this is a *subset* of $\mathbb{R}^3$, since it is a collection of elements of $\mathbb{R}^3$, and write $\{(x, 2 - x, 2 - x) : x \in \mathbb{R}\} \subset \mathbb{R}^3$.

The system

$$\begin{aligned} 3x &= 3 \\ 5x &= 5 \\ x &= 2 \end{aligned}$$

clearly has no solutions, since the first equation implies that $x = 1$ but the third equation implies that $x = 2$. Thus the set of solutions is the *empty set* $\{\} = \varnothing$.

We say that two systems of equations are *equivalent* if they have the same set of solutions. Thus the process of solving a system of equations is mostly the process of converting a system into an equivalent system that is simpler.

There are three basic operations we can perform on a system of equations to get an equivalent system:

1. We can write the equations in a different order.

2. We can multiply any equation by a nonzero scalar.

3. We can add a multiple of one equation to another.

All three of these operations are guaranteed not to change the solution set; proving this is a reasonable exercise. Our goal now is to find an efficient way to use these rules to get a useful solution to our system.

Example 1.3. The system

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

is equivalent to

$$\begin{aligned} 2x + y &= 3 \\ -2x + -10y &= 6 \end{aligned}$$

and then

$$\begin{aligned} 0x + -9y &= 9 \\ -2x + -10y &= 6 \end{aligned}$$

then

$$\begin{aligned} 0x + y &= -1 \\ -2x + -10y &= 6 \end{aligned}$$

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

$$0x + y = 1$$
$$x + 0y = 2$$

which give us our solution of $x = 2, y = 1$ or $(x, y) = (2, 1)$.

This takes up a really awkward amount of space on the page, though, and we'd like to find a better and more systematic way of approaching this process.

Remark 1.4. There's another possible approach to solving these systems, called the method of substitution. We could observe that if $2x + y = 3$ then $y = 3 - 2x$, and substitute that into our other equation to give

$$\begin{aligned} x + 5(3 - 2x) &= -3 \\ 15 - 9x &= -3 \\ 9x &= 18 \\ x &= 2 \end{aligned}$$

and from here we can see that $y = 3 - 2(2) = -1$.

This is often much simpler to do in your head for small systems. But it scales up really poorly to systems with more than two or three equations and variables, so we'll want to learn something more effective.

1.2 The matrix of a system

Looking at a system of linear equations, we notice that it can be described by an array of real numbers. These numbers are naturally laid out in a rectangular grid, so we wnat to find an efficient way to represent them.

Definition 1.5. A *(real) matrix* is a rectangular array of (real) numbers. A matrix with m rows and n columns is *a* $m \times n$ *matrix*, and we notate the set of all such matrices by $M_{m\times n}$.

A $m \times n$ matrix is *square* if $m = n$, that is, it has the same number of rows as columns. We will sometimes represent the set of $n \times n$ square matrices by M_n.

We will generally describe the elements of a matrix with the notation

$$(a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

We can now take the information from a system of linear equations and encode it in a matrix. Right now, we will just use this as a convenient notational shortcut; we will see later on in the course that this has a number of theoretical and practical advantages.

Definition 1.6. The *coefficient matrix* of a system of linear equations given by

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad & \qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is the matrix

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

and the *augmented coefficient matrix* is

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

Example 1.7. Suppose we have a system

$$\begin{aligned} 4x + 2y + 2z &= 8 \\ 3x + 2y + z &= 6. \end{aligned}$$

Then the coefficient matrix is

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

and the augmented coefficient matrix is

$$\left[\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 3 & 2 & 1 & 6. \end{array}\right]$$

Earlier we listed three operations we can perform on a system of equations without changing the solution set: we can reorder the equations, multiply an equation by a nonzero scalar, or add a multiple of one equation to another. We can do analogous things to the coefficient matrix.

Definition 1.8. The three *elementary row operations* on a matrix are

I Interchange two rows.

II Multiply a row by a nonzero real number.

III Replace a row by its sum with a multiple of another row.

Example 1.9. What can we do with our previous matrix? We can

$$\begin{bmatrix} 4 & 2 & 2 \\ 3 & 2 & 1 \end{bmatrix} \xrightarrow{I} \begin{bmatrix} 3 & 2 & 1 \\ 4 & 2 & 2 \end{bmatrix} \xrightarrow{II} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} \xrightarrow{III} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}.$$

So how do we use this to solve a system of equations? The basic idea is to remove variables from successive equations until we get one equation that contains only one variable—at which point we can substitute for that variable, and then the others. To do that with this matrix, we have

$$\left[\begin{array}{ccc|c} 4 & 2 & 2 & 8 \\ 3 & 2 & 1 & 6 \end{array}\right] \xrightarrow{III} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 6 \end{array}\right] \xrightarrow{III} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 2 & -2 & 0 \end{array}\right] \xrightarrow{II} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 0 \end{array}\right].$$

What does this tell us? That our system of equations is equivalent to the system

$$x + z = 2$$
$$y - z = 0.$$

This gives us the answer I stated earlier: $z = 2 - x$ and $y = z = 2 - x$.

Example 1.10. Solve the system of equations

$$x + 2y + z = 3$$
$$3x - y - 3z = -1$$
$$2x + 3y + z = 4.$$

This system has augmented coefficient matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right] \overset{III}{\to} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 2 & 3 & 1 & 4 \end{array}\right] \overset{III}{\to} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & -1 & -1 & -2 \end{array}\right]$$

$$\overset{II}{\to} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & 1 & 1 & 2 \end{array}\right] \overset{I}{\to} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & -7 & -6 & -10 \end{array}\right] \overset{III}{\to} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

which gives us the system

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

The last equation tells us $z = 4$, which then gives $y = -2$ and $x = 3$. We can check that this solves the system.

1.3 Row Echelon Form

We want to solve systems of linear equations, using these matrix operations. We want to be somewhat more concrete about our goals: what exactly would it look like for a system to be solved?

Definition 1.11. A matrix is in *row echelon form* if

- Every row containing nonzero elements is above every row containing only zeroes; and
- The first (leftmost) nonzero entry of each row is to the right of the first nonzero entry of the above row.

Remark 1.12. Some people require the first nonzero entry in each nonzero row to be 1. This is really a matter of taste and doesn't matter much, but you should do it to be safe; it's an easy extra step to take by simply dividing each row by its leading coefficient.

Example 1.13. The following matrices are all in Row Echelon Form:

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

The following matrices are not in Row Echelon Form:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & 2 & 5 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 5 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Definition 1.14. The process of using elementary row operations to transform a system into row echelon form is *Gaussian elimination.*

A system of equations sometimes has a solution, but does not always. We say a system is *inconsistent* if there is no solution; we say a system is *consistent* if there is at least one solution.

Example 1.15. Consider the system of equations given by

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ -1x_1 + -1x_2 + x_5 &= -1 \\ -2x_1 + -2x_2 + 3x_5 &= 1 \\ x_3 + x_4 + 3x_5 &= -1 \\ x_1 + x_2 + 2x_3 + 2x_4 + 4x_5 &= 1. \end{aligned}$$

This translates into the augmented matrix

$$\left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 1 & 1 & 2 & 2 & 4 & 1 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right]$$
$$\to \left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right].$$

We see that the final two equations are now $0 = -4$ and $0 = -3$,s othe system is inconsistent.

Example 1.16. Let's look at another system that is almost the same.

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ -1x_1 + -1x_2 + x_5 &= -1 \\ -2x_1 + -2x_2 + 3x_5 &= 1 \\ x_3 + x_4 + 3x_5 &= 3 \\ x_1 + x_2 + 2x_3 + 2x_4 + 4x_5 &= 4. \end{aligned}$$

This translates into the augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 0 & 0 & 1 & 1 & 3 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

We see this system is now consistent. Our three equations are

$$x_1 + x_2 + x_3 + x_4 + x_5 = 1 \qquad x_3 + x_4 + 2x_5 = 0 \qquad x_5 = 3.$$

Via back-substitution we see that we have

$$x_5 = 3 \qquad x_3 + x_4 = -6 \qquad x_1 + x_2 = 4.$$

Thus we could say the set of solutions is $\{(\alpha, 4-\alpha, \beta, -6-\beta, 3)\} \subseteq \mathbb{R}^5$.

What we were just doing definitely worked, but even after we finished transforming the matrix we still needed to do some more work. So we'd like to reduce the matrix even further until we can just read the answer off from it.

Definition 1.17. A matrix is in *reduced row echelon form* if it is in row echelon form, and the first nonzero entry in each row is the only entry in its column.

This means that we will have some number of columns that each have a bunch of zeroes and one 1. Other than that we may or may not have more columns, which can contain basically anything; we've used up all our degrees of freedom to fix those columns that contain the leading term of some row.

Note that the columns we have fixed are not necessarily the first columns, as the next example shows.

Example 1.18. The following matrices are all in reduced Row Echelon Form:

$$\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 17 & 0 & 2 & 8 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

The following matrices are not in reduced Row Echelon Form:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 2 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 15 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Example 1.19. Let's solve the following system by putting the matrix in reduced row echelon form.

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 &= 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 &= 2 \end{aligned}$$

We have

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

$$\to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

From this we can read off the solution $x_1 + x_2 + x_3 = 1, x_4 = 2, x_5 = -1$. Thus the set of solutions is $\{(1 - \alpha - \beta, \alpha, \beta, 2, -1)\}$.

We say some systems of equations are "overdetermined", which means that there are more equations than varaibles. Overdetermined equations are "usually" inconsistent, but not always—they can be consistent when some of the equations are redundant.

Example 1.20. The system

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

gives the matrix

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

$$\to \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3/2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 3/5 & 1 \\ 0 & 1 & 1/5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3/2 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1/10 \\ 0 & 1 & 0 & -3/10 \\ 0 & 0 & 1 & 3/2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This gives us the solution $x_1 = 1/10, x_2 = -3/10, x_3 = 3/2$, which you can go back and check solves the original system.

This overdetermined system does have a solution, but only because two of the equations were redundant, as we could see in the second matrix where two lines are identical. In fact we can go back to the original set of equations, and see that if we add two times the first equation to the second equation, we get the third—which is the redundancy.

Other systems of equations are "underdetermined", which means there are more variables than equations. These systems are usually but not always consistent.

Example 1.21. Let's consider the system

$$\begin{aligned} -x_1 + x_2 - x_3 + 3x_4 &= 0 \\ 3x_1 + x_2 - x_3 - x_4 &= 0 \\ 2x_1 + x_2 - 2x_3 - x_4 &= 0. \end{aligned}$$

This gives us the matrix

$$\left[\begin{array}{cccc|c} -1 & 1 & -1 & 3 & 0 \\ 3 & 1 & -1 & -1 & 0 \\ 2 & 1 & -2 & -1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & -1 & 1 & -3 & 0 \\ 0 & 4 & -4 & 8 & 0 \\ 0 & 3 & -4 & 5 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & -1 & 1 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 3 & -4 & 5 & 0 \end{array}\right]$$
$$\to \left[\begin{array}{cccc|c} 1 & -1 & 1 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & -1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & -1 & 1 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right]$$
$$\to \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

We see that we can't "simplify" the fourth column in any way; we don't have any degrees of freedom after we fix the first three columns. This means that we can pick x_4 to be anything we want, and the other variables are given by $x_1 - x_4 = 0, x_2 - 3x_4 = 0, x_3 + x_4 = 0$. Thus the set of solutions is $\{(\alpha, 3\alpha, -\alpha, \alpha)\}$.

Remark 1.22. A system of any size can be either consistent or inconsistent. $0 = 1$ is an inconsistent system with one equation, and

$$\begin{aligned} x_1 + \cdots + x_{100} &= 0 \\ x_1 + \cdots + x_{100} &= 1 \end{aligned}$$

is an inconsistent system with a hundred variables and only two equations. In contrast,

$$\begin{aligned} x_1 &= 1 \\ x_1 &= 1 \\ &\vdots \\ x_1 &= 1 \end{aligned}$$

has only one variable, and many equations, and is still consistent.

1.4 Matrix Algebra

So far we've treated matrices as just being a convenient way to write down a bunch of numbers. But matrices are interesting mathematical objects in their own right, and we can do a lot of useful calculations with them.

1.4.1 Simple Operations

We want to start with a couple of simple operations. Neither of these operations really depend on the structure of the matrix; they treat the matrix as a list of numbers.

Definition 1.23. If $A = (a_{ij})$ is an $m \times n$ matrix, and $r \in \mathbb{R}$ is a real number, then we can multiply each entry of the matrix A by the real number R. This is called *scalar multiplication* and we say that r is a *scalar.*

$$rA = (ra_{ij}) = \begin{bmatrix} ra_{11} & ra_{12} & \dots & ra_{1n} \\ ra_{21} & ra_{22} & \dots & ra_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \dots & ra_{mn} \end{bmatrix}.$$

Definition 1.24. If $A = (a_{ij})$ and $B = (b_{ij})$ are two $m \times n$ matrices, we can add the two matrices by adding each individual pair of coordinates together.

$$A + B = (a_{ij} + b_{ij}) = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

Example 1.25.

$$3\begin{bmatrix} 2 & 5 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 6 & 15 \\ -3 & 12 \end{bmatrix} \qquad \begin{bmatrix} 4 & 1 & 3 \\ -2 & 5 & -1 \end{bmatrix} + \begin{bmatrix} -2 & 7 & 5 \\ 1 & -6 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 8 & 8 \\ -1 & -1 & 3 \end{bmatrix}$$

1.4.2 Matrix Multiplication

Definition 1.26. If $A \in M_{\ell \times m}$ and $B \in M_{m \times n}$, then there is a matrix $AB \in M_{\ell \times n}$ whose ij element is

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}.$$

If you're familiar with the dot product, you can think that the ij element of AB is the dot product of the ith row of A with the jth column of b.

Note that A and B don't have to have the same dimension! Instead, A has the same number of columns that B has rows. The new matrix will have the same number of rows as A and the same number of columns as B.

Example 1.27.

$$\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix}\begin{bmatrix}5 & -1\\ 3 & 2\end{bmatrix} = \begin{bmatrix}1\cdot 5 + 3\cdot 3 & 1\cdot(-1) + 3\cdot 2\\ 2\cdot 5 + 4\cdot 3 & 2\cdot(-1) + 4\cdot 2\end{bmatrix} = \begin{bmatrix}14 & 5\\ 22 & 6\end{bmatrix}$$

$$\begin{bmatrix}4 & 6\\ 2 & 1\end{bmatrix}\begin{bmatrix}3 & 1 & 5\\ 4 & 1 & 6\end{bmatrix} = \begin{bmatrix}4\cdot 3 + 6\cdot 4 & 4\cdot 1 + 6\cdot 1 & 4\cdot 5 + 6\cdot 6\\ 2\cdot 3 + 1\cdot 4 & 2\cdot 1 + 1\cdot 1 & 2\cdot 5 + 1\cdot 6\end{bmatrix} = \begin{bmatrix}36 & 10 & 56\\ 10 & 3 & 16\end{bmatrix}.$$

Matrix multiplication is *associative*, by which we mean that $(AB)C = A(BC)$.

Matrix multiplication is not commutative: in general, it's not even the case that AB and BA both make sense. If $A \in M_{3\times 4}$ and $B \in M_{4\times 2}$ then AB is a 3×2 matrix, but BA isn't a thing we can compute. But even if AB and BA are both well-defined, they are not equal.

Example 1.28.

$$\begin{bmatrix}3 & 5 & 1\\ -2 & 0 & 2\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & 3\\ 4 & 1\end{bmatrix} = \begin{bmatrix}3\cdot 2 + 5\cdot 1 + 1\cdot 4 & 3\cdot 1 + 5\cdot 3 + 1\cdot 1\\ -2\cdot 2 + 0\cdot 1 + 2\cdot 4 & -2\cdot 1 + 0\cdot 3 + 2\cdot 1\end{bmatrix} = \begin{bmatrix}15 & 19\\ 4 & 0\end{bmatrix}$$

$$\begin{bmatrix}2 & 1\\ 1 & 3\\ 4 & 1\end{bmatrix}\begin{bmatrix}3 & 5 & 1\\ -2 & 0 & 2\end{bmatrix} = \begin{bmatrix}2\cdot 3 + 1\cdot(-2) & 2\cdot 5 + 1\cdot 0 & 2\cdot 1 + 1\cdot 2\\ 1\cdot 3 + 3\cdot(-2) & 1\cdot 5 + 3\cdot 0 & 1\cdot 1 + 3\cdot 2\\ 4\cdot 3 + 1\cdot(-2) & 4\cdot 5 + 1\cdot 0 & 4\cdot 1 + 1\cdot 2\end{bmatrix} = \begin{bmatrix}4 & 10 & 4\\ -3 & 5 & 7\\ 10 & 20 & 6\end{bmatrix}.$$

Particularly nice things happen when our matrices are square. Any time we have two $n \times n$ matrices we can multiply them by each other in either order (though we will still get different things each way!).

Example 1.29.

$$\begin{bmatrix}4 & 1\\ -3 & 5\end{bmatrix}\begin{bmatrix}-1 & 1\\ 1 & -2\end{bmatrix} = \begin{bmatrix}-3 & 2\\ 8 & -13\end{bmatrix}$$

$$\begin{bmatrix}-1 & 1\\ 1 & -2\end{bmatrix}\begin{bmatrix}4 & 1\\ -3 & 5\end{bmatrix} = \begin{bmatrix}-7 & 4\\ 10 & -9\end{bmatrix}.$$

However, matrix multiplication does satisfy the *distributive* and *associative* properties.

Fact 1.30. *If $A \in M_{\ell\times m}$ and $B, C \in M_{m\times n}$ then $A(B + C) = AB + AC$.*

If $A \in M_{\ell\times m}, B \in M_{m\times n}, C \in M_{n\times p}$ then $(AB)C = A(BC)$.

Example 1.31. Let

$$A = \begin{bmatrix}4 & 1\\ -3 & 5\end{bmatrix} \qquad B = \begin{bmatrix}-1 & 1\\ 1 & -2\end{bmatrix} \qquad C = \begin{bmatrix}3 & 2\\ 1 & -5\end{bmatrix}.$$

Then we have

$$AB = \begin{bmatrix} -3 & 2 \\ 8 & -13 \end{bmatrix} \qquad AC = \begin{bmatrix} 13 & 3 \\ -4 & -31 \end{bmatrix} \qquad AB + AC = \begin{bmatrix} 10 & 5 \\ 4 & -44 \end{bmatrix}$$
$$B + C = \begin{bmatrix} 2 & 3 \\ 2 & -7 \end{bmatrix} \qquad A(B+C) = \begin{bmatrix} 10 & 5 \\ 4 & -44 \end{bmatrix}.$$

Thus we see $AB + AC = A(B + C)$.

We can similarly compute

$$AB = \begin{bmatrix} -3 & 2 \\ 8 & -13 \end{bmatrix} \qquad (AB)C = \begin{bmatrix} -3 & 2 \\ 8 & -13 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & -5 \end{bmatrix} = \begin{bmatrix} -7 & -16 \\ 11 & 81 \end{bmatrix}$$
$$BC = \begin{bmatrix} -2 & -7 \\ 1 & 12 \end{bmatrix} \qquad A(BC) = \begin{bmatrix} 4 & 1 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} -2 & -7 \\ 1 & 12 \end{bmatrix} = \begin{bmatrix} -7 & -16 \\ 11 & 81 \end{bmatrix}$$

1.4.3 Transposes

Definition 1.32. If A is a $m \times n$ matrix, then we can form a $n \times m$ matrix B by flipping A across its diagonal, so that $b_{ij} = a_{ji}$. We say that B is the *transpose* of A, and write $B = A^T$.

If $A = A^T$ we say that A is *symmetric*. (Symmetric matrices must always be square).

Example 1.33.

$$\text{If } A = \begin{bmatrix} 1 & 3 & 5 \\ -1 & 4 & 2 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 1 & -1 \\ 3 & 4 \\ 5 & 2 \end{bmatrix}.$$
$$\text{If } B = \begin{bmatrix} 5 & 3 \\ 3 & -2 \end{bmatrix} \text{ then } B^T = \begin{bmatrix} 5 & 3 \\ 3 & -2 \end{bmatrix}$$

and thus B is symmetric.

Fact 1.34.

- $(A^T)^T = A$.
- $(A + B)^T = A^T + B^T$.
- $(rA)^T = rA^T$.
- *If* $A \in M_{\ell \times m}$ *and* $B \in M_{m \times n}$ *then* $(AB)^T = B^T A^T$.

1.4.4 Matrices and Systems of Equations

We will do a lot with matrices in the future (a linear algebra class that doesn't cover general vector spaces is often called a matrix algebra class). In the current context we mostly want it to make it easier to talk about systems of equations.

Let

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

be a system of linear equations. Then $A = (a_{ij}) \in M_{m\times n}$ is its coefficient matrix, and $\mathbf{b} = (b_1, \ldots, b_m)$ is an element of $\mathbb{R}^m$, but we can also think of it as a $m \times 1$ matrix $b = [b_1, \ldots, b_m]^T$. If we take $\mathbf{x} = [x_1, \ldots, x_n]^T$ to be a $n \times 1$ matrix, we can rewrite our linear system as the equation

$$A\mathbf{x} = \mathbf{b},$$

which is certainly much easier to write down.

Example 1.35. If $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ and $\mathbf{b} = [4, 6]^T$, then the equation $A\mathbf{x} = \mathbf{b}$ is

$$\begin{aligned} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 4 \\ 6 \end{bmatrix} \\ \begin{bmatrix} x + 3y \\ 2x + 4y \end{bmatrix} &= \begin{bmatrix} 4 \\ 6 \end{bmatrix} \\ x + 3y &= 4 \\ 2x + 4y &= 6 \end{aligned}$$

1.5 The identity matrix and matrix inverses

We just saw that any system of linear equations can be written $A\mathbf{x} = \mathbf{b}$, which reminds us of the single-variable linear equation $ax = b$. In the single-variable case we can just divide both sides of the equation by a, as long as $a \neq 0$; it would be nice if we can do the same thing for any system of linear equations.

But what does it mean to divide by a matrix? When we define division, we often start by understanding reciprocals $\frac{1}{a}$. So we start by asking what matrix is the equivalent of the number 1.

Definition 1.36. For any n we define the *identity matrix* to be $I_n \in M_n$ to have a 1 on every diagonal entry, and a zero everywhere else. For example,

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

If $A \in M_n$ then $I_n A = A = A I_n$. Thus it is a *multiplicative identity* in the ring of $n \times n$ matrices.

The identity matrix is symmetric (that is, $I_n^T = I_n$).

Now we want to define multiplicative inverses, the equivalent of reciprocals. The definition is not difficult to invent:

Definition 1.37. Let A and B be $n \times n$ matrices, such that $AB = I_n = BA$. Then we say that B is the *inverse* (or *multiplicative inverse*) of A, and write $B = A^{-1}$.

If such a matrix exists, we say that A is *invertible* or *nonsingular*. If no such matrix exists, we say that A is *singular*.

Example 1.38. The identity matrix I_n is its own inverse, and thus invertible.

The matrices

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1/10 & 2/5 \\ 3/10 & -1/5 \end{bmatrix}$$

are inverses to each other, as you can check.

Example 1.39. The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ has no inverse, since

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$$

won't be the identity for any a, b, c, d. Thus this matrix is singular.

Remark 1.40. If $AB = I_n$ then $BA = I_n$. This isn't really trivial but we won't prove it.

As the last example shows, finding the inverse to a matrix is a matter of solving a big pile of linear equations at the same time (one for each coefficient of the inverse matrix). Fortunately, we just got good at solving linear equations. Even more fortunately, there's an easy way to organize the work for these problems.

Proposition 1.41. *Let A be a $n \times n$ matrix. Then if we form the augmented matrix $\begin{bmatrix} A & I_n \end{bmatrix}$, then A is invertible if and only if the reduced row echelon form of this augmented matrix is $\begin{bmatrix} I_n & B \end{bmatrix}$ for some matrix B, and furthermore $B = A^{-1}$.*

Proof. Let X be a $n \times n$ matrix of unknowns, and set up the system of equations implied by $AX = I_n$. This will be the same set of equations we are solving with this row reduction, and thus a matrix X exists if and only if this system has a solution, which happens if and only if the reduced row echelon form of $\begin{bmatrix} A & I_n \end{bmatrix}$ has no all-zero rows on the A side. □

Example 1.42. Let's find an inverse for $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$.

We form and reduce the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & -5 & 1 & -2 & 0 \\ 0 & 1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 5 \\ 0 & 1 & 0 & 0 & 1 & -4 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

Thus $A^{-1} = \begin{bmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix}$. We can check this by multiplying the matrices back together.

Example 1.43. Find the inverse of $B = \begin{bmatrix} 1 & 0 & 4 \\ 1 & 1 & 6 \\ -3 & 0 & -10 \end{bmatrix}$.

We form and reduce the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 1 & 1 & 6 & 0 & 1 & 0 \\ -3 & 0 & -10 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 2 & 3 & 0 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 3/2 & 0 & 1/2 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & 0 & -2 \\ 0 & 1 & 2 & -4 & 1 & -1 \\ 0 & 0 & 1 & 3/2 & 0 & 1/2 \end{array}\right].$$

Thus $B^{-1} = \begin{bmatrix} -5 & 0 & -2 \\ -4 & 1 & -1 \\ 3/2 & 0 & 1/2 \end{bmatrix}$.

Example 1.44. What happens if we try to find an inverse for $C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$? We start with

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

but then there is no way to make the left-side block of the matrix into the identity I_2. Thus this matrix C is not invertible.

There are many more interesting properties of inverse matrices we'd like to discuss, but we don't have the tools to explain them properly yet. We will be returning to the properties of matrices throughout the course as we develop more techniques and vocabulary.

1.6 Homogeneous systems and subspaces

There's one particular category of systems of linear equations that's especially important to us, and will lead into the main subject matter of the course.

Definition 1.45. The $n \times 1$ matrix $\mathbf{0} = [0, \ldots, 0]^T$ whose entries are all zero is called the *zero vector*.

A system of linear equations $A\mathbf{x} = \mathbf{b}$ is called *homogeneous* if $\mathbf{b} = \mathbf{0}$, that is, if the constant term in each equation is zero. Otherwise, it is *non-homogeneous*.

It's pretty clear that every homogeneous system has at least one solution: the solution where every variable is equal to zero. It may have many more solutions than that.

Definition 1.46. For a given matrix A, the subspace of solutions to the equation $A\mathbf{x} = \mathbf{0}$ is called the *nullspace* $N(A)$ or the *kernel* $\ker(A)$ of the matrix A.

Example 1.47. Find the null space of $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix}$.

We row reduce the matrix

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & -2 & 1 & 0 \end{array}\right]$$

We see that x_3 and x_4 are fixed variables, and x_1, x_2 are determined by x_3 and x_4. (You could of course do this the other way around). Then we have $x_1 = x_3 - x_4$ and $x_2 = x_4 - 2x_3$.

Thus $N(A) = \{(\alpha - \beta, \beta - 2\alpha, \alpha, \beta)\} = \{\alpha(1, -2, 1, 0) + \beta(-1, 1, 0, 1)\}$.

Remark 1.48. It's not too hard to see that a square matrix A is invertible if and only if $N(A) = \{\mathbf{0}\}$. If the matrix is invertible, then row-reducing it gets to be the identity matrix—and so the solution to the associated homogeneous system is just $\mathbf{0}$. Conversely, if the only solution is $\mathbf{0}$ then you must not have any rows of all zeros in the reduced form of your matrix, so it's invertible.

We can see that if we add together two solutions to this system of equations, we will get another. In fact, this must be true of any homogeneous system.

Proposition 1.49 (Homogeneity). *Suppose $A\mathbf{x} = \mathbf{0}$ is a homogeneous sytem of linear equations. Then:*

1. *$\mathbf{0}$ is a solution to the system.*
2. *If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions to this system, then $\mathbf{x}_1 + \mathbf{x}_2$ is a solution.*
3. *If $\mathbf{x}$ is a solution to this system, and r is a real number, then $r\mathbf{x}$ is a solution.*

Remark 1.50. We can rephrase this result: for any matrix A, we have

1. $\mathbf{0} \in N(A)$
2. If $\mathbf{x}_1, \mathbf{x}_2 \in N(A)$ then $\mathbf{x}_1 + \mathbf{x}_1 \in N(A)$
3. If $r \in \mathbb{R}$ and $\mathbf{x} \in N(A)$ then $r\mathbf{x} \in N(A)$.

This says exactly the same thing, but puts the emphasis on the matrix A rather than on the equation $A\mathbf{x} = \mathbf{0}$.

Proof. 1. Calculation confirms that $A\mathbf{0} = \mathbf{0}$.

2. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions, then $A\mathbf{x}_1 = \mathbf{0}$ and $A\mathbf{x}_2 = 0$, so we have

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Thus $\mathbf{x}_1 + \mathbf{x}_2$ is a solution.

3. If $\mathbf{x}$ is a solution and $r \in \mathbb{R}$, then

$$A(r\mathbf{x}) = rA\mathbf{x} = r\mathbf{0} = \mathbf{0}.$$

Thus $r\mathbf{x}$ is a solution.

□

In contrast, the set of solutions to a non-homogeneous system $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b} \neq \mathbf{0}$ never has these nice properties.

1. The zero vector is never a solution, since $A\mathbf{0} = \mathbf{0} \neq \mathbf{b}$.
2. Adding two solutions doesn't give you another solution: $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{b} + \mathbf{b} = 2\mathbf{b} \neq \mathbf{b}$.
3. Multiplying a solution by a scalar doesn't give another solution: $Ar\mathbf{x} = r\mathbf{b} \neq \mathbf{b}$ unless $r = 1$.

So there's something special about homogeneous systems, which we will discuss in more detail in 2.3.

But even though the set of solutions to a non-homogeneous system doesn't have the nice properties of proposition 1.48, we can still say a lot about what it looks like.

Proposition 1.51. *Suppose $A\mathbf{x} = \mathbf{b}$ is a non-homogeneous linear system.*

If $U = N(A))$ and $\mathbf{x}_0$ is a solution to $A\mathbf{x} = \mathbf{b}$, then the set of solutions to the system $A\mathbf{x} = \mathbf{b}$ is the set

$$N(A) + \mathbf{x}_0 = \{\mathbf{y} + \mathbf{x}_0 : \mathbf{y} \in N(A)\}.$$

Proof. We want to show that two sets are equal, so we show that each is a subset of the other.

First, suppose that $\mathbf{x}_1$ is a solution to $A\mathbf{x}_1 = \mathbf{b}$. Then we have

$$\begin{aligned} b &= A\mathbf{x}_0 \\ b &= A\mathbf{x}_1 \\ b - b &= A\mathbf{x}_1 - A\mathbf{x}_0 = A(\mathbf{x}_1 - \mathbf{x}_0) \\ \mathbf{0} &= A(\mathbf{x}_1 - \mathbf{x}_0). \end{aligned}$$

Thus $\mathbf{y} = \mathbf{x}_1 - \mathbf{x}_0$ is a solution to $A\mathbf{x} = \mathbf{0}$, and then $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}$ for some $\mathbf{y} \in U$.

Conversely, suppose $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{y}$ for some $\mathbf{y} \in U$. Then

$$A\mathbf{x}_1 = A(\mathbf{x}_0 + \mathbf{y}) = A\mathbf{x}_0 + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Thus x_1 is a solution to $A\mathbf{x} = \mathbf{b}$. □

Remark 1.52. Notice this did not depend on the specific matricx, or even really the fact that A is a matrix at all; it only depends on the ability to distribute matrix multiplication across sums of vectors. Operations with this property are called "linear" and we will discuss them in much more detail in section 4.

Definition 1.53. Suppose $A\mathbf{x} = \mathbf{b}$ is a system of linear equations. We call the equation $A\mathbf{x} = \mathbf{0}$ the associated homogeneous system of linear equations. That is, the associated homogeneous system has the same coefficients for all the variables, but the constants are all zero.

Thus proposition 1.50 lets us understand the set of solutions to a non-homogeneous system based on the solutions to the associated homogeneous system.

Example 1.54. Let's find a set of solutions to the system

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 + 2x_2 + 3x_3 &= 6 \\ 2x_1 + 3x_2 + 4x_2 &= 9. \end{aligned}$$

Gaussian elimination gives

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 3 & 6 \\ 2 & 3 & 4 & 9 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Taking $x_3 = \alpha$ as a free variable, our solution set is $\{(\alpha, 3-2\alpha, \alpha)\} = \{(0,3,0)+\alpha(1,-2,1)\}$. Indeed, we see that this set corresponds to elements of the vector space spanned by $\{(1,-2,1)\}$, plus a specific solution $(0,3,0)$.

Alternatively, we could have solved the homogeneous system first, and seen that the solution was $x_1 - x_3 = 0, x_2 + 2x_3 = 0$, telling us that $N(A) = \{\alpha(1,-2,1)\}$. Then we just need to find a solution; to my eyes the obvious solution is $(1,1,1)$. So our theorem tells us that the solution set is $\{(1,1,1) + \alpha(1,-2,1)\}$. This may not *look* like the solution we got before, but it is in fact the same set, since $(1,1,1) = (0,3,0) + (1,-2,1)$.

Example 1.55. Now consider the system

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

It's easy enough to see that this system has no solutions, since the sum of the first two equations should be the third.

This at first might seem concerning, since $N(A)$ is never empty. But our proposition assumed that there was at least one solution to the non-homogeneous system; when there are no solutions, the proposition doesn't actually say anything. But *if* any solution exists, proposition 1.50 tells us that the set of solutions is just the nullspace of A, plus an offset.

Example 1.56. Let's find the set of solutions to

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

We form the matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -2 & 2 & 4 \\ 1 & 2 & -1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & 4 \\ 0 & 1 & -2 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & -3 & 1 & 4 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & -5 & 10 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|c} 1 & 0 & 3 & -2 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

giving us the sole solution $x_1 = 4, x_2 = -2, x_3 = -2$.

If we look at the corresponding homogeneous system, we see that we can reduce the matrix to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and thus the sole solution to the homogeneous system of equations is $x_1 = x_2 = x_3 = 0$. Then every solution to our non-homogeneous system is a solution to our homogeneous system plus some vector in $\{\vec{0}\}$. Since there is only one vector in that set, there is only one solution to our system.