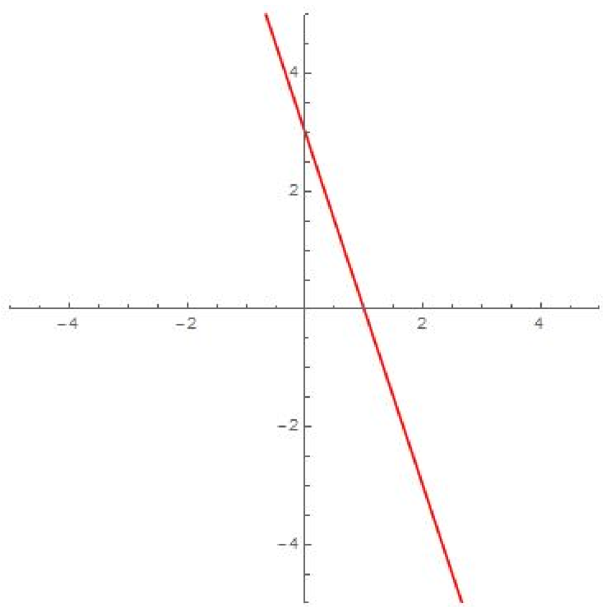
Introduction to Linear Systems

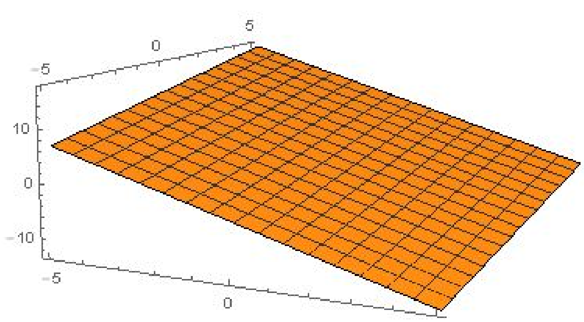
The equations

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

have a common feature: each describes a geometric shape that is "linear". Upon rewriting the first equation $x + \frac{1}{3}y = 1$ as $y = 3 - 3x$, it is clear that this equation in two variables describes the (two-dimensional) line graphed below:



Using similar reasoning, we rewrite the equation $2x - y + z = 2$ as $z = 2 - 2x + y$, and recognize this equation in three variables as the equation of a plane (a sort of three-dimensional analogue of a line), graphed below:



Unfortunately, the equation $3w + x + y + 2z = 0$ is in four variables, thus describes a geometric shape in four-dimensional space. However, this equation is similar enough to the previous two that we can get some feel for its "shape": the equation describes a "linear" structure in four-dimensional space, which we refer to as a *hyperplane*.

Linear equations are intimately tied to both matrices and vector spaces, thus we we will spend several sections studying them. We begin by defining exactly what we mean by a *linear equation*.

Definition. A *linear equation* in the n variables $x_1, x_2, \ldots, x_n$ is an equation which can be written in the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

where b and each of the coefficients $a_1, a_2, \ldots,$ and a_n are real or complex numbers.

If $b = 0$, i.e. the equation has form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0;$$

we say that such a system is *homogeneous*.

There are a few quick things to note about how we use the definition to test an equation for linearity:

1. Each variable is "linear" in the sense that it is raised to the *first* power, and not to any other power.
2. Each variable must have constant coefficients, and *not* variable coefficients.

Example. The equation

$$\sqrt{2}x_1 - i = \tfrac{x_2}{4} + x_3$$

is linear in the variables x_1, x_2, and x_3, since it can be rewritten as

$$\sqrt{2}x_1 - \tfrac{x_2}{4} - x_3 = i;$$

note that each variable has a number as its coefficient, and each variable is raised to the first power.

Example. The equations

$$x_1x_3 + x_2 = 0$$

and

$$x^2 + y^2 = 9$$

each fail to be linear in their variables. The first equation is non-linear in variables x_1, x_2, and x_3 because the coefficient of x_1 is the *variable* x_3. The second equation is non-linear in variables x and y because each variable appears to the second power.

An important point to take away from the definition is that a linear equation in n variables has geometric realization as an n dimensional linear structure.

For example, a linear equation in two variables can be written as $a_1x_1 + a_2x_2 = b$, which is clearly the equation of a line; a linear equation in three variables has form $a_1x_1 + a_2x_2 + a_3x_3 = b$, which is the general form for the equation of a plane in three-dimensional space. Linear equations with more variables simply describe linear structures in higher dimensional space.

Systems of Linear Equations and Their Solutions

Definition. A *system of linear equations* is a collection of equations in the same variables.

For example,

$$\begin{aligned} 3x - y &= 5 \\ 2x + y &= -10 \end{aligned}$$

is a system of two equations in the two unknowns x and y.

The general form for a system of m linear equations in n variables $x_1, x_2, \ldots, x_n$ is

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

Definition. A *solution* to a system of linear equations in n variables is a list of n numbers that, when substituted for the variables in the equations, make the equations true simultaneously.

We can check that $x = -1$, $y = -8$ is a solution to the system

$$\begin{aligned} 3x - y &= 5 \\ 2x + y &= -10 \end{aligned}$$

by substituting $x = -1$ and $y = -8$ in the appropriate positions:

$$3(-1) - (-8) = -3 + 8 = 5$$

and

$$2(-1) + (-8) = -2 - 8 = -10,$$

so clearly $x = -1$, $y = -8$ make the equations true simultaneously.

A slightly more succinct way to write the solution above is as the *ordered pair* $(-1, -8)$. In general, if $s_1, s_2, \ldots, s_n$ are numbers so that

$$x_1 = s_1,\ x_2 = s_2,\ \ldots,\ \text{and } x_n = s_n$$

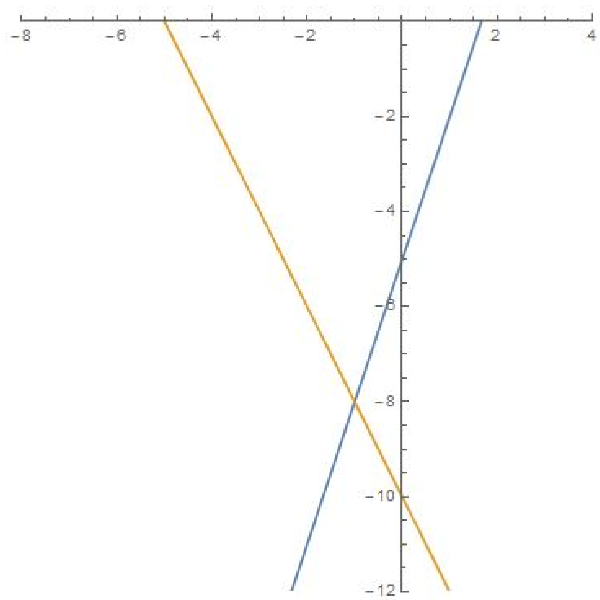
is a solution to the system in (1), then we will write this solution as the *ordered n-tuple*

$$(s_1, s_2, \ldots, s_n),$$

or in “vector” form as

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}.$$

In order to truly understand what a solution to a system of equations means, it is helpful to return to the geometric reasoning we introduced earlier. Recall that the equations $3x - y = 5$ and $2x + y = -10$ each have geometric realization as lines in two-dimensional space. The lines described by the equations are graphed below:



Recall that we called the list $(-1, -8)$ a solution to the system. Geometrically, this means that the point $(-1, -8)$ satisfies both equations; so $(-1, -8)$ must lie on the line $3x - y = 5$, as well as on the line $2x + y = -10$. In other words, the solution is the point of intersection of the two equations, which is clear from the graph above.

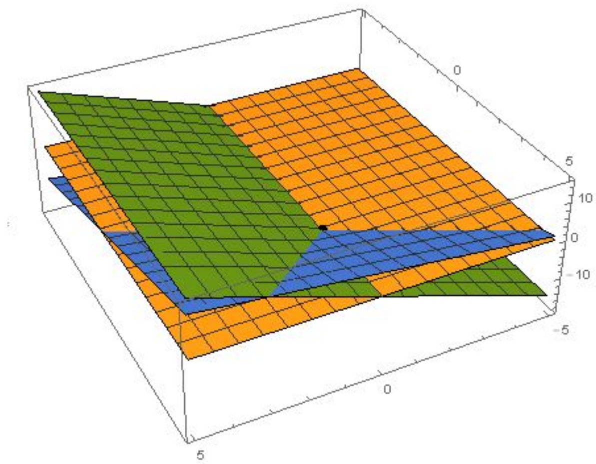
Key Point. A solution to a system of equations is just a description of the points that the graphs of the equations have in common.

Example. The system

$$\begin{aligned} x + y &= 2 \\ x + 2y - z &= 4 \\ 3x - z &= 0 \end{aligned}$$

has solution $x = 0$, $y = 2$, $z = 0$ (or $(0, 2, 0)$).

The graph of each of the equations above is a plane in three-dimensional space, as indicated below:

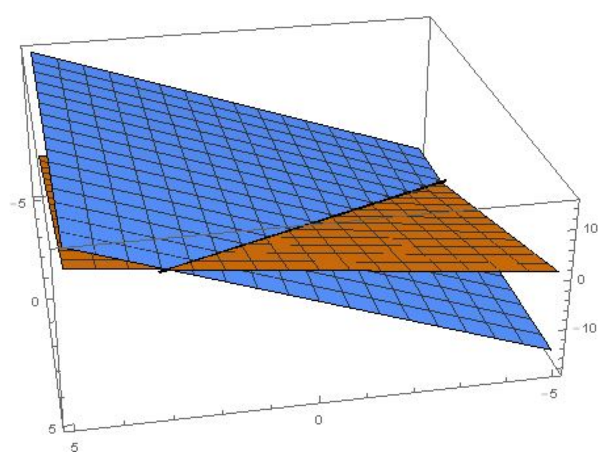


Notice that the solution $(0, 2, 0)$ is the *unique point* which lies *on each of the three planes.*

Example. Of course, a solution to a system doesn't have to be a single point. Consider the system we get by deleting the first equation in the system from the previous example:

$$\begin{aligned} x + 2y - z &= 4 \\ 3x - z &= 0 \end{aligned}$$

Let's attempt to describe the form of a solution to this system using geometric reasoning. The graphs of the equations above are below:

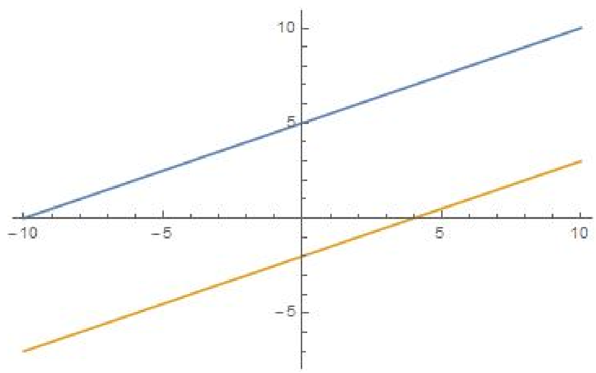


Notice that, in this case, the graphs intersect in *infinitely many points*, in particular on the line graphed in black above. Indeed, the solution to the system can be written as $(t, 2+t, 3t)$, a parametric description of the line.

Example. There is one more interesting type of behavior that we may observe from a system, as illustrated by the following example. The graphs of the equations in the system

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

are both lines, which clearly have the same slope. In other words, the lines are parallel; it is clear from their graphs below that the lines do not intersect:



Thus it is clear that the system has *no solution*.

In order to differentiate between systems that can be solved, and systems that *can not* be solved, we record the following definition:

Definition. A system is called *inconsistent* if it has no solution; otherwise, the system is *consistent*.

Key Point. So far, we have seen three distinct possibilities for the solution to a system of equations:

1. A unique solution (consistent)

2. Infinitely many solutions (also consistent)

3. No solution (inconsistent).

It turns out that these are the *only* possibilities for the solution set to a system of linear equations:

Theorem. A system of m linear equations in n unknowns, $0 < m, n < \infty$, has either no solution, a unique solution, or infinitely many solutions.

Proof. Consider the system

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

Without loss of generality, we may assume that, for each i, there is at least one j so that $a_{ij} \neq 0$, and for each j, there is at least one i so that $a_{ij} \neq 0$. Thus each equation has at least one nonzero term, and each variable appears at least once with nonzero coefficient.

Suppose that the system has at least two solutions, say

$$(s_1,\ s_2,\ \ldots,\ s_n) \text{ and } (s'_1,\ s'_2,\ \ldots,\ s'_n).$$

In other words,

$$\begin{aligned} b_i &= a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n \\ &= a_{i1}s'_1 + a_{i2}s'_2 + \ldots + a_{in}s'_n \end{aligned}$$

for all i, $1 \leq i \leq m$.

Thus we have

$$\begin{aligned} 0 &= a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n - (a_{i1}s'_1 + a_{i2}s'_2 + \ldots + a_{in}s'_n) \\ &= (s_1 - s'_1)a_{i1} + (s_2 - s'_2)a_{i2} + \ldots + (s_n - s'_n)a_{in} \end{aligned}$$

for all i; and since the original solutions are different, there is at least one j so that $s_j - s'_j \neq 0$.

Now for all i, $1 \leq i \leq m$, and any parameter t, we see that

$$\begin{aligned} 0 + \mathrm{b_i} &= t\Big((s_1 - s'_1)a_{i1} + (s_2 - s'_2)a_{i2} + \ldots + (s_j - s'_j)a_{ij} \ldots + (s_n - s'_n)a_{in}\Big) \\ &\quad + \Big(a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n\Big) \\ &= \big(t(s_1 - s'_1) + s_1\big)a_{i1} + \big(t(s_2 - s'_2) + s_2\big)a_{i2} + \ldots + \big(t(s_n - s'_n) + s_n\big)a_{in}. \end{aligned}$$

Thus the list

$$\big(t(s_1 - s'_1) + s_1,\ t(s_2 - s'_2) + s_2,\ \ldots,\ t(s_n - s'_n) + s_n\big)$$

is a solution to the system for all t; since $s_j - s'_j \neq 0$, the jth coordinate of the solution is a nonconstant linear function. Thus each solution is unique, so that the system has infinitely many solutions.

Solving Systems of Linear Equations

Now that we know what types of solutions we might encounter when faced with a particular system of equations, we should discuss methods for actually finding these solutions. There are ad-hoc techniques for solving systems, but the solution method that will work most reliably stems from an encoding of the system in matrix form. Thus we introduce the idea of the *matrix of a linear system.*

The Matrix of a Linear System

Let us briefly recall the definition of am $m \times n$ matrix:

Definition. An $m \times n$ matrix is an array of mn numbers arranged in m (horizontal) rows and n (vertical) columns; such a matrix has form

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

where each a_{ij} is a number in the ith row and jth column.

As mentioned above, we will be able to use matrices to solve linear systems; the first step in doing so is to understand how we can use a linear system to build an appropriate matrix:

Definition. The linear system

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

has *augmented matrix*

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

For example, the system

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

has augmented matrix

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

The augmented matrix we created above has encoded the data from the linear system; let's investigate this concept in more detail. First of all, we should notice that the rows of the augmented matrix correspond to equations in the system. For example, row 1 corresponds to equation 1:

$$\begin{pmatrix} 1 & 1 & 0 & | & 2 \\ 1 & 2 & -1 & | & 4 \\ 3 & 0 & -1 & | & 0 \end{pmatrix} \to 1x_1 + 1x_2 + 0x_3 = 2.$$

Second, it is important to note that *columns* of the augmented matrix correspond to *coefficients* of a particular variable from the system. For example, the second column of the augmented matrix encodes the coefficients of x_2 from the original system:

$$\begin{pmatrix} 1 & 1 & 0 & | & 2 \\ 1 & 2 & -1 & | & 4 \\ 3 & 0 & -1 & | & 0 \end{pmatrix} \to \begin{matrix} x_1 + 1x_2 + 0x_3 & = & 2 \\ x_1 + 2x_2 - x_3 & = & 4 \\ 3x_1 + 0x_2 - x_3 & = & 0. \end{matrix}$$

If we are told that a matrix is the augmented matrix of a system, it is quite simple to recover the system. For example, suppose that we know that the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \tag{2}$$

is the augmented matrix for a system. In order to recover the system, we begin by noting that the system must consist of 3 equations, one corresponding to each row.

Consider the first row of the augmented matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 4 \end{pmatrix}.$$

The first three entries of the row correspond to coefficients of the system's variables, while the last entry is the constant on the right-hand side of the equation's equality. Thus equation corresponding to the first row of the matrix is

$$x_1 + 0x_2 + 0x_3 = 4.$$

Similarly, the second row

$$\begin{pmatrix} 0 & 1 & 0 & 3 \end{pmatrix}$$

of the augmented matrix must correspond to the equation

$$0x_1 + x_2 + 0x_3 = 3;$$

the last row

$$\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

corresponds to the equation

$$0x_1 + 0x_2 + x_3 = \tfrac{1}{2}.$$

So the system of equations corresponding to the augmented matrix in (2) is

$$\begin{aligned} x_1 + 0x_2 + 0x_3 &= 4 \\ 0x_1 + x_2 + 0x_3 &= 3 \\ 0x_1 + 0x_2 + x_3 &= \tfrac{1}{2}. \end{aligned} \tag{3}$$

As mentioned above, the matrix in (2) is really just a convenient way to encode the data from the system in (3).

Incidentally, this particular system is already solved: clearly

$$x_1 = 4,\ x_2 = 3,\ \text{and } x_3 = \tfrac{1}{2};$$

indeed, we could have simply looked at the system's augmented matrix to determine this solution!

Elementary Row Operations

In order to use the matrix of a linear system to solve the system, we will need one more tool, known as the *elementary row operations*.

Elementary Row Operations:

- Multiply any row by a non-zero constant
- Interchange any pair of rows
- Add or subtract a constant multiple of one row from another.

The following theorem indicates why we should care about elementary row operations when we are solving linear systems:

Theorem. Suppose that the augmented matrix M of a linear system $\mathcal{M}$ is obtained from the augmented matrix L of the linear system $\mathcal{L}$ by a finite sequence of elementary row operations on L. Then $\mathcal{M}$ and $\mathcal{L}$ have the same solution set.

Again, it is important to emphasize that these three operations, while apparently changing the form of a matrix, *do not affect the solutions of the corresponding linear system*, and thus we can use these operations to our advantage when we wish to solve a system of equations.

Example. Given the system

$$\begin{aligned} x_3 &= \tfrac{1}{2} \\ x_1 + x_2 &= 7 \\ x_1 &= 4, \end{aligned}$$

1. Find the augmented matrix for the system.

2. Apply elementary row operations to the augmented matrix to reduce it to the form of the matrix in (2),
$$\begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 0 & 1 & 0 & | & 3 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix}.$$

3. Use the reduction to solve the system.

1. The augmented matrix for the system is quite easy to build–recall that each row of the system corresponds to a row of the matrix, whose entries are the coefficients of the variables. Thus the system
$$\begin{aligned} x_3 &= \tfrac{1}{2} \\ x_1 + x_2 &= 7 \\ x_1 &= 4 \end{aligned}$$
has augmented equation
$$\begin{pmatrix} 0 & 0 & 1 & | & \frac{1}{2} \\ 1 & 1 & 0 & | & 7 \\ 1 & 0 & 0 & | & 4 \end{pmatrix}.$$

2. In order to reduce the matrix above to the form
$$\begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 0 & 1 & 0 & | & 3 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix},$$
let's begin by noting that the last row of
$$\begin{pmatrix} 0 & 0 & 1 & | & \frac{1}{2} \\ 1 & 1 & 0 & | & 7 \\ 1 & 0 & 0 & | & 4 \end{pmatrix}$$
looks just like the desired first row; we can use the second elementary row operation, switching rows, to start the reduction:
$$\begin{pmatrix} 0 & 0 & 1 & | & \frac{1}{2} \\ 1 & 1 & 0 & | & 7 \\ 1 & 0 & 0 & | & 4 \end{pmatrix} \xrightarrow{\text{switch rows 1 and 3}} \begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 1 & 1 & 0 & | & 7 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix}.$$
Now we see that the first and third rows have the right form, but row 2 is off. However, if we use the third elementary row operation and subtract row 1 from row 2, we will move closer to the desired form:

$$\begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 1 & 1 & 0 & | & 7 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix} \xrightarrow{\text{row 2-row 1}} \begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 1-1 & 1-0 & 0-0 & | & 7-4 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 0 & 1 & 0 & | & 3 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix}.$$

Notice that we have now reduced our original matrix to the one desired.

3. Above, I claimed that applying elementary row operations to a matrix does not change the solutions to the corresponding linear system; we've already seen that the matrix

$$\begin{pmatrix} 1 & 0 & 0 & | & 4 \\ 0 & 1 & 0 & | & 3 \\ 0 & 0 & 1 & | & \frac{1}{2} \end{pmatrix}$$

implies solution

$$x_1 = 4, \ x_2 = 3, \text{ and } x_3 = \tfrac{1}{2}.$$

Since we got to this matrix from the original one via elementary row operations, their systems should share solutions:

$$(4, 3, \tfrac{1}{2})$$

must be a solution to the system

$$\begin{aligned} x_3 &= \tfrac{1}{2} \\ x_1 + x_2 &= 7 \\ x_1 &= 4. \end{aligned}$$

Let's verify the solution: clearly $x_1 = 4$ and $x_3 = \frac{1}{2}$ are correct, so we only need to check the second row of the equation, $x_1 + x_2 = 7$. We've claimed that $x_1 = 4$ and $x_2 = 3$, so the second row does indeed check out, and

$$(4, 3, \tfrac{1}{2})$$

is a solution to the system.

The example above should give us some motivation as to how we will go about solving linear systems; the general procedure is as follows:

1. Write the system's augmented equation.
2. Use elementary row operations to rewrite the augmented matrix in a simpler form (i.e., one whose solutions are easy to find).
3. Since elementary row operations *do not alter solutions*, the solutions found above are also solutions to the original system.

The only tricky part about solving a system is the second step–determining a form for the augmented matrix whose solutions are easy to find. We will focus on this step in the next section.