CHAPTER 1

How to Solve Linear Equations (and How!)

1. Blah Blah

Linear equations are everywhere. That sounded ominous, sorry. I just mean they are everywhere like air, not everywhere like our enemies. Linear equations express simple relationships, and simple relationships are ubiquitous. Anytime you add something to something else, or you multiply something by a number, you’re taking advantage of the two building blocks of linear equations, and indeed, of linear algebra: addition and scalar multiplication. But wait! There’s more! Virtually every linear-algebraic question which we will come across can be paraphrased as a question about linear equations.

This chapter introduces an algorithmic method called row reduction for solving any system of linear equations. Algorithms are a big deal because they’re guaranteed to work. Think about it: how often can you absolutely guarantee that following your directions will solve the problem at hand? Not just the problem at hand, but every problem like it that has ever existed and ever will exist? Every system of linear equation in the Library of Babel can be solved using row reduction. That’s huge!

The bad news—of course, there has to be some bad news to keep the narrative tension interesting—is that the computations can be a bit tedious. But the (additional, free of charge, bonus) good news is two fold: the tedious calculations are straightforward, and the point of learning them is not so that we can do rote work, but so that we can better understand the mechanics of linear equations, their structure if you will, their sole¹.

2. You Already Know How to Solve Systems of Linear Equations

To start, let’s confuse our enemies, who are plentiful, by calling systems of linear equations soles; this acronym is exclusive to our cult. Plato says our soles preexist us, and so we are born with their complete knowledge. My job is easy, then. I don’t have to give you new information, but rather, to organize the information you already have in your head.

2.1. A simple example to illustrate a point. I want to convince you that you already know how to solve soles.

Example 1. Here’s a nice system of two equations in two variables, or a 2×2 sole.

\[
\begin{align*}
  x + y &= -1 \\
  -x + 2y &= 2
\end{align*}
\]

I know what you’re thinking: “I really want to use that first \(x\) to cancel the second \(x\)!” Since you know what you’re doing, let’s follow your lead. But out of an abundance of caution, or OCD, I will change the second equation while keeping the first intact. I will also record the fact that, by using the first \(x\) to cancel the second \(-x\), we’ve added (1 times) equation #1 to equation #2:

\[
\begin{align*}
  x + y &= -1 \\
  -x + 2y &= 2 \\
  1\cdot x + 1\cdot y &= 1 - 1 \\
  3y &= -1
\end{align*}
\]

Again, I know what you’re thinking: “If I divide that second equation by 3, I’ll have solved for \(y\)!” And again, your instinct is correct. However, in keeping with the general linear algebraic theme of adding stuff to other stuff and multiplying stuff by numbers, let’s multiply the second equation by \(\frac{1}{3}\) instead of dividing by 3. Of course, it’s the same thing, but cultists can be particular about ritual.

\[
\begin{align*}
  x + y &= -1 \\
  -x + 2y &= 2 \\
  1\cdot x + 1\cdot y &= 1 - 1 \\
  \frac{1}{3}\cdot 3y &= -1 \\
  y &= \frac{-1}{3}
\end{align*}
\]

We’re almost there, but your instinct may be about to fail you. (Plato says that soles need guidance.) The reasonable thing to do here would be to plug \(y = \frac{-1}{3}\) into the first equation and solve for \(x\). But “plug in”

¹Think you’ve found a typo? I mend typos. Email to let me know. Thanks!
1. LINEAR EQUATIONS

turns out not to be a generalizable direction or a universal form, and sometimes it’s better to immitate instead of act reasonably. Let’s take our hint from the very first step in this example where we used one \( x \) to cancel another, and let’s now use one \( y \) to cancel the other. In particular, let’s use the \( y \) of the second equation to cancel the \( y \) of the first. We’ll do so by adding \(-1\) times equation \#2 to equation \#1. This is exactly the same process as plugging in, but neater.

\[
\begin{align*}
  x &+ y = -1 \\
-2x + 2y &= 2 \\
3y &= 1 \\
-2x + 2y &= 1
\end{align*}
\]

We’re done! We have solved the \( 2 \times 2 \) system using preexisting knowledge. The solution is \( x = -\frac{4}{3}, y = \frac{1}{3} \), but that’s not important. Well, it is important, but beside the point. More on point, we are in a great position to easily check if our answer is correct\(^2\) by simply plugging in: \(-\frac{4}{3} + \frac{1}{3} = -1\sqrt{\cdot} \) and \(-\left(-\frac{4}{3}\right) + 2\left(\frac{1}{3}\right) = 2\sqrt{\cdot} \).

What is important is the method we used to get to the solution. Our method was made up of two main row operations:\(^3\).

1. Add a multiple of one row to another.
2. Multiply a row by a nonzero number scalar.\(^4\)

We applied operation (1) in the first and third steps (\( \frac{-1}{1} \frac{0}{1} \frac{1}{1} \) and \( \frac{-1}{1} \frac{2}{1} \frac{0}{1} \)), and operation (2) in the second step (\( \frac{-1}{1} \frac{2}{1} \)). Note a subtle difference between these two operations. When we add a multiple of one row to another, the one row stays the same and the other changes; when we multiply a row by a scalar, that row changes. This is an important one for future cult promotions, so keep it in the back of your mind. Also note the nonzero restriction in operation (2). If we multiplied an equation by zero, we’d annihilate it, losing the information therein, whereas the point of these row operations is to keep the information intact while simplifying the sole’s form.

Speaking of promotions, we are now ready to apply these operations to larger soles.

2.2. A more complicated example to illustrate the same point. Let’s see how far our row operations can take us with a larger sole.

**Example 2.** We encounter the following system of four equations in five variables, or \( 4 \times 5 \) sole.

\[
\begin{align*}
x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
-2x_1 + 4x_2 - 4x_3 + 5x_4 &= 1 \\
3x_1 - 4x_2 + 8x_3 - 5x_4 + x_5 &= 1 \\
x_1 + 4x_3 + x_5 &= 2
\end{align*}
\]

In our simple example above, we used the \( x \) in the top left to get rid of the \(-x\) below it. Here, let’s to the same, using the \( x_1 \) in the top left to get rid of all the \( x_1 \)s below it. We will use operation (1), add a multiple of one row to another, three times:

- add 2 times row \#1 to row \#2,
- add \(-3\) times row \#1 to row \#3, and
- add \(-1\) times row \#1 to row \#4,

As with the simple example above, we’ll keep track of all equations simultaneously, and we’ll denote the operations we did for posterity.

\[
\begin{align*}
\frac{2}{1} \frac{-1}{1} \frac{0}{1} \\
\frac{-3}{1} \frac{0}{1} \frac{1}{1} \\
\frac{-1}{1} \frac{0}{1} \frac{1}{1}
\end{align*}
\]

\[
\begin{align*}
x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
- x_4 + 2x_5 &= 1 \\
2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4 \\
2x_2 + 2x_3 + 3x_4 &= 3
\end{align*}
\]

---

\(^2\)This is a rarity in life that should be taken advantage of.

\(^3\)Spoiler alert: we will see one more row operation in a bit.

\(^4\)Remember, in a cult, we replace words we know and love with new words that mean exactly the same thing.
Notice how, if we ignore the first row, we have in our possession a new $3 \times 4$ sole. We’ve effectively removed one equation and one unknown from consideration—

\[
\begin{align*}
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
2x_2 & + & 2x_3 & + 3x_4 = 3
\end{array}
\end{align*}
\]

—and so we can apply our operation to this new $3 \times 4$ sole. In other words, all we have to do is use the top left variable in our new sole to eliminate all the variables below it in its column. Alas, there is no $x_2$ in the top left position of our new $3 \times 4$ sole! Oh no!

\[
\begin{align*}
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
?? & - & x_4 & + 2x_5 = 1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
2x_2 & + & 2x_3 & + 3x_4 = 3
\end{array}
\end{align*}
\]

I kid. It’s not a big deal. We can switch two rows. That is the third row operation we need. Since the equations’ order is arbitrary and independent of the information they contain, reordering them will not impact their solutions. Let’s switch equation #2 with equation #3.

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
2x_2 & + & 2x_3 & + 3x_4 = 3
\end{array}
\]

We pause to summarize that we have three row operations.

1. Add a multiple of one row to another.
2. Multiply a row by a nonzero scalar.
3. Switch two rows.

Now we can continue with our row reduction as before, and add $-1$ times the new equation #2 to equation #4. Note that the new equation #3 already has no $x_2$, so we needn’t add a multiple of equation #2 to it.

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1
\end{array}
\]

Look at that! By eliminating the $x_2$ in equation #4, we got rid of the $x_3$ as well. So if we ignore the first two equations, we’re now left with a $2 \times 2$ sole.

\[
\begin{align*}
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1
\end{array}
\end{align*}
\]

As before, we use the top left variable in our new sole, the $-x_4$ in equation #3, to eliminate the $-x_4$ below it.

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1
\end{array}
\]

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1
\end{array}
\]

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1
\end{array}
\]

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1 \\
0 & = & 0
\end{array}
\]

\[
\begin{array}{cccc}
x_1 & - & 2x_2 & + 2x_3 - 3x_4 + x_5 = -1 \\
2x_2 & + & 2x_3 & + 4x_4 - 2x_5 = 4 \\
- & x_4 & + & 2x_5 = 1 \\
0 & = & 0
\end{array}
\]
Since equation #3 completely annihilated equation #4, all we have left is a 0 = 0 to mark the spot where that equation once stood. This is a tradition in our cult: the reduced sole should have the same number of equations as the original sole, even if it means writing 0 = 0. Let’s reexamine where we are in full color:

\[
\begin{align*}
  x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
  2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4 \\
  -x_4 + 2x_5 &= -1 \\
  0 &= 0
\end{align*}
\]

At this stage, our first trick has been exhausted. We can no longer use these top-left, seemingly-important variables (\(x_1\) in equation #1, \(2x_2\) in equation #2, and \(-x_4\) in equation #3) to eliminate the variables underneath them. This seems like a reasonable stage to give a name to: we say the sole is in row echelon form, or echelon form for short, or EF for super short. Notice how, in EF, the first nonzero entry in each row is to the right of the first nonzero entry of the row above it.

We’ve accomplished a lot together so far, and yet, we have not solved our system. To find inspiration on how to proceed, let’s recall our previous, simpler example—in particular, its second step:

Remember that time when

\[
\begin{align*}
  x + y &= -1 \\
  3y &= 1 \\
  y &= \frac{1}{3}
\end{align*}
\]

All that step did was to change the coefficient of the seemingly-important variable \(y\) in equation #2 from 3 to 1. So we follow suit with our more complicated example, multiplying rows by scalars as necessary to change the coefficients of our seemingly-important variables to 1. By the way, the coefficients of the seemingly-important variables are called pivots. Above, the pivots were 1, 2, and \(-1\); next, the pivots will all be 1. We call the number of pivots the rank of the sole. In our example, the rank equals 3 since there are three pivots.

OK, let’s multiply equation #2 by \(\frac{1}{2}\) and equation #3 by \(-1\) to make the pivots equal 1.

\[
\begin{align*}
  x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
  2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4 \\
  -x_4 + 2x_5 &= -1 \\
  0 &= 0
\end{align*}
\]

\[
\begin{array}{c|c}
  \frac{1}{2} \cdot \circ & \frac{1}{3} \\
  \bar{\circ} & \circ
\end{array}
\]

\[
\begin{align*}
  x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
  x_2 + x_3 + 2x_4 - x_5 &= 2 \\
  x_4 - 2x_5 &= 1 \\
  0 &= 0
\end{align*}
\]

We now return to the last step of the simpler example for additional inspiration.

Remember that other time when

\[
\begin{align*}
  x + y &= -1 \\
  y &= \frac{1}{3} \rightarrow x &= \frac{-2}{3} \\
  y &= \frac{-2}{3}
\end{align*}
\]

What we did there was to use the lower right variable \(y\) to eliminate the \(y\) above it. We can try something similar here: use the same seemingly-important variables that eliminated the space underneath in order to eliminate the space above. We can start with \(x_4\) in equation #3 to knock out \(2x_4\) in equation #2 and \(-3x_4\) in equation #1. To that end, we add \(-2\) times equation #3 to equation #2, and 3 times equation #3 to equation #1.

\[
\begin{array}{c|c}
  -2 \cdot \circ & \circ \\
  \bar{\circ} & 3 \cdot \circ
\end{array}
\]

5Most echelon forms that you’ll encounter in nature will be of the row echelon form variety. We will not see column echelon forms explicitly on our journey, so I’ll just drop the ‘row’ designation for brevity (and to confuse our enemies).

6The word traces its origins from scala (Latin for ladder) to eschelon (Old French for ladder) to échelon (French for rung) to echelon (English for rank in an organization, like the military).

7Some other cults call them leading nonzeros. We are at war with those cults.
\[\begin{align*}
\quad x_1 &\quad -2x_2 \quad +2x_3 \quad -5x_5 \quad = \quad 2 \\
\quad x_2 &\quad +x_3 \quad +3x_5 \quad \phantom{=} \quad = \quad 0 \\
\quad x_4 &\quad -2x_5 \quad \phantom{=} \quad = \quad 1 \\
\quad 0 &\quad \phantom{=} \quad \phantom{=} \quad \phantom{=} \quad = \quad 0 \\
\end{align*}\]

And finally, we use the seemingly-important variable \(x_2\) in equation \#2 to get rid of the \(-2x_2\) in equation \#1 by adding 2 times equation \#2 to equation \#1.

\[\begin{align*}
\quad x_1 &\quad +4x_3 \quad +x_5 \quad \phantom{=} \quad = \quad 2 \\
\quad x_2 &\quad +x_3 \quad +3x_5 \quad \phantom{=} \quad = \quad 0 \\
\quad x_4 &\quad -2x_5 \quad \phantom{=} \quad \phantom{=} \quad = \quad 1 \\
\quad 0 &\quad \phantom{=} \quad \phantom{=} \quad \phantom{=} \quad = \quad 0 \\
\end{align*}\]

It may not look like we’re done, but we’re done! Take a moment to celebrate—we’ve gone as far as row reduction allows. Our sole is now in \((row)\) reduced echelon form, or \(\text{REF}\) for short. \(\text{REF}\) is the goal of row reduction.

But what are we celebrating? Where is the solution? Well, it depends. If \(x_3 = 0\) and \(x_5 = 0\), then \(x_1 = 2\), \(x_2 = 0\), and \(x_4 = 1\). If \(x_3 = -1\) and \(x_5 = 3\), then \(x_1 = 3\), \(x_2 = -8\), and \(x_4 = 7\). In fact, if \(x_3 = s\) and \(x_5 = t\), then \(x_1 = 2 - 4s - t\), \(x_2 = -s - 3t\), and \(x_4 = 1 + 2t\). Written in a neater way more appropriate for cult consumption, we may say the \textit{general solution} to our sole is:

\[\begin{align*}
\quad x_1 &\quad = \quad 2 - 4s - t \\
\quad x_2 &\quad = \quad -s - 3t \\
\quad x_3 &\quad = \quad s \\
\quad x_4 &\quad = \quad 1 + 2t \\
\quad x_5 &\quad = \quad t \\
\end{align*}\]

But how is that a solution? I’m glad you asked. Whatever values the variables \(x_3\) and \(x_5\) (or \(s\) and \(t\)) take, we can use those to determine what \(x_1\), \(x_2\), and \(x_4\) must equal. This is the master-slave inversion that Friedrich Nietzsche loved to talk about: the seemingly-important variables \((x_1, x_2,\) and \(x_4)\) that determined the flow of row reduction are now completely dependent on the seemingly-unimportant variables \((x_3\) and \(x_5)\), whereas the seemingly-unimportant variables can equal anything they like. As such, we will call \(x_3\) and \(x_5\) \textit{free variables}, and we will call \(x_1, x_2,\) and \(x_4\) \textit{basic variables}.

As with the \(2 \times 2\) case, we can check if our answer is correct by simply plugging it back into the original sole. For instance, we could plug the specific solution \(x_1 = 3, x_2 = -8, x_3 = -1, x_4 = 7, x_5 = 3\) into the first equation: \(3 - 2(-8) + 2(-1) - 3(7) + 3 = -1\sqrt{\text{.}}\). A more industrious cultist might plug the general solution into all four equations.

\textbf{2.3. How could our example have gone wrong?} The Easter bunny is delicious, and soles don’t always have solutions.

\textbf{Example 3.} Suppose that we put on our new glasses and notice that Example 2 above had had a \(3\) instead of a \(2\) as the right-hand side of equation \#4.

\[\begin{align*}
\quad x_1 &\quad -2x_2 \quad +2x_3 \quad -3x_4 \quad +x_5 \quad \phantom{=} \quad = \quad -1 \\
\quad -2x_1 &\quad +4x_2 \quad -4x_3 \quad +5x_4 \quad = \quad 1 \\
\quad 3x_1 &\quad -1x_2 \quad +8x_3 \quad -5x_4 \quad +x_5 \quad \phantom{=} \quad = \quad 1 \\
\quad x_1 &\quad +4x_3 \quad +x_5 \quad \phantom{=} \quad = \quad 2 \quad 3 \leftarrow \text{this is different!} \\
\end{align*}\]

Then our process of row reduction would have unfolded quite similarly—indeed, row reduction depends only on the coefficients of the variables on the left-hand side of a sole—up to a point:

\[\begin{align*}
\quad x_1 &\quad -2x_2 \quad +2x_3 \quad -3x_4 \quad +x_5 \quad \phantom{=} \quad = \quad -1 \\
\quad -4x_1 &\quad +8x_2 \quad -8x_3 \quad +10x_4 \quad +2x_5 \quad = \quad 4 \\
\quad 2x_2 &\quad +2x_3 \quad +4x_4 \quad -2x_5 \quad = \quad 4 \\
\quad 2x_2 &\quad +2x_3 \quad +3x_4 \quad = \quad 3 \\
\end{align*}\]
1. LINEAR EQUATIONS

Our sole is now in EF, almost identical to the previous example, with one minor but devastating difference: equation #4 has become 0 = 1 instead of 0 = 0. There are no possible values of \(x_1, x_2, x_3, x_4,\) and \(x_5\) that could possible make equation #4 true! Zero cannot ever equal one. That is, this sole cannot have any solutions. We will call a sole that has solutions consistent, and a sole that doesn’t have solutions inconsistent. We have witnessed the only way a sole can be inconsistent—when its EF has a row of the form 0 = nonzero. Otherwise, row reduction always leads to solutions.

2.4. How could our example have gone right? The Easter bunny is delicious, and some soles always have solutions.

Example 4. Suppose that in a fit of anger, we change the right-hand sides of Example 2 to all zeros. Such a sole is called homogeneous.

\[
\begin{align*}
x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
-2x_1 + 4x_2 - 4x_3 + 5x_4 &= -1 \\
3x_1 - 4x_2 + 8x_3 - 5x_4 + x_5 &= 0 \\
x_1 + 4x_3 + x_5 &= 2
\end{align*}
\]

As before, the row reduction steps would have been then same because they depend only on the left-hand side coefficients. But I don’t care (IDC) about the left-hand side. Instead, notice that if the right-hand sides all equaled zero, then none of the row operations could change them to nonzero. So at every stage of row reduction, our sole will look like this:

\[
\begin{align*}
\text{IDC} &= 0 \\
&= 0 \\
&= 0 \\
&= 0
\end{align*}
\]

Therefore, there will never be a row of the form 0 = nonzero. In other words, homogeneous soles are always consistent. Can you find at least one solution to any homogeneous sole?

You now know all there is to know about solving systems of linear equations. Meditate on that power while we review the salient features of row reduction.

---

8Except in rare circumstances when there are no numbers in the universe other than 0 or 1. True story.
3. Some Commentary on the Previous Examples

If the unexamined life is not worth living, then the unexamined example is a waste of time. In this section, we revisit the previous example with an eye towards extracting wisdom, generalizing patterns, and making the most of time. We begin to see how, following the simple and elegant algorithm of row reduction, the underlying structure of linear algebra reveals its patterns like worm molds under a rock in the sun-baked mud. Or something. Read these.

DEFINITION 5. A linear equation is an equation of the form \( a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b \). The numbers \( a_1, a_2, \cdots, a_n \) are called coefficients, \( b \) is the right-hand side constant term, and \( x_1, x_2, \cdots, x_n \) are variables, or unknowns. We must be more careful with the notation when looking at a system of \( m \) equations in \( n \) variables:

\[
\begin{align*}
& a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = b_1 \\
& a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = b_2 \\
& \quad \vdots \\
& a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = b_m
\end{align*}
\]

Notice how we write the row subscript first and the column subscript second; this is a convenient convention. We usually put the constant terms on the right hand side of the equations and align the variables on top of each other for reasons both aesthetic and functional. For brevity and pun potential, we refer to systems of linear equations as soles.

The sole in Example 2 was a \( 4 \times 5 \) system with 20 coefficients and 4 right-hand side constants. Following the notation introduced in Definition 5, we have \( a_{13} = 2 \), \( a_{25} = 0 \), and \( b_4 = 2 \), whereas \( a_{52} \) does not exist.

\[
\begin{align*}
& a_{13} \quad \quad \quad x_1 - 2x_2 + \frac{2}{3} x_3 - 3x_4 + x_5 = -1 \\
& a_{25} \quad \quad \quad -2x_1 + 4x_2 - 4x_3 + 5x_4 + b_4 = 1 \\
& \quad \quad \quad 3x_1 - 4x_2 + 8x_3 - 5x_4 + x_5 = 1 \\
& a_{52} \quad \quad \quad x_1 + 4x_3 + x_5 = \frac{2}{3}
\end{align*}
\]

While most linear equations might look familiar to you, the same equation can be linear in one context but not in the other. For instance, \( 4 \sin(x) + 7 \sin(y) = 9 \) is not linear in \( x \) and \( y \), but it is linear in \( \sin(x) \) and \( \sin(y) \). Similarly, the equation \( x/(y+z) = 5 \) can be written in the linear form \( x - 5y - 5z = 0 \), but we should note that the first form does not admit the solution \( x = y = z = 0 \), while the second (linear) form does.

DEFINITION 6. There are three row operations that we use to solve soles.

1. Add a multiple of one row to another.
2. Multiply a row by a nonzero scalar.
3. Switch two rows.

The process of using row operations to find solutions to a sole is called row reduction. Repeated uses of row operation (1) in the same step should be done with care. It is safe to do so if we’re adding multiples of the same row to different rows; indeed, we did plenty of that in the above examples. But mistakes can easily arise if we instead add multiples of different rows to other rows in the same step. For instance, consider the following row operations impatiently inflicted on the sole in EF. Can you spot the...
ensuing mistake?

\[
\begin{align*}
8x + 3y - 2z &= 1 \\
y + z &= -1 \\
z &= 2
\end{align*}
\]

So in applying the first row operation—add a multiple of one row to another—remember to use one one row at a time.

**Definition 7.** A sole is in echelon form (EF) if the first nonzero entry in each row is to the right of the first nonzero entry in the row above it.

Soles in EF are easy to spot in nature because of their distinctive stair-like pattern where each step is one equation high. Take another look at the sole from Example 2 at two different stages.

\[
\begin{align*}
x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4
\end{align*}
\]
not EF
\[
\begin{align*}
x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4 \\
-x_4 + 2x_5 &= -1
\end{align*}
\]
EF! Hooray!

**Definition 8.** The **pivots** of a sole in EF are the first nonzero coefficients in each row.

We can speak of pivots only once a sole is in EF. Just as with EF itself, pivots are easy to spot in nature.

**Definition 9.** A variable is called **basic** if it corresponds to a column with a pivot. Otherwise, it is called **free**.

**Definition 10.** The **rank** of a sole equals the number of pivots—that is, the number of basic variables.

Once we know where the pivots are in the EF of a sole, we can identify variables as basic or free and determine the sole’s rank.

\[
\begin{align*}
1x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 &= -1 \\
2x_2 + 2x_3 + 4x_4 - 2x_5 &= 4 \\
-x_4 + 2x_5 &= -1 \\
0 &= 0
\end{align*}
\]

**Observation 11.** In a sole that’s in EF, there can be at most one pivot per row and per column.

**Proof.** Since a pivot is by definition the first nonzero coefficient of a row, it’s not possible to have more than one first nonzero coefficient per row. And two pivots in the same column would violate the definition of EF, where the first nonzero entry in each row must to the right of the first nonzero entry in the row above it.

**Observation 12.** The rank of an \( m \times n \) sole must be no larger than \( m \) and no larger than \( n \).
3. SOME COMMENTARY ON THE PREVIOUS EXAMPLES

Proof. This follows from Observation 11: the rank of a sole is the number of pivots, and there can be at most one pivot per row and one pivot per column. If the rank of an \( m \times n \) sole equals \( r \), then we have \( r \leq \min(m, n) \).

**Definition 13.** A sole is in reduced echelon form (REF) if it is in EF, each of the pivots equals 1, and each entry above (as well as below) the pivots equals zero.

![Zeroes above pivots]

I have to get some fine print about notation out of the way in order to satisfy our legal team. If you are not litigious, you can skip this\(^{11}\) footnote.

**Observation 14.** If a sole is in REF, then each basic variable appears in exactly one row.

Proof. Recall that a basic variable is one associated to a pivot. By definition, in REF, all entries above and below the pivot equal zero. As such, if a variable is associated to a pivot in a certain row of a sole in EF, then that row will contain the only occurrence of that variable.

**Definition 15.** A sole is consistent if it has at least one solution. Otherwise, it is inconsistent.

**Observation 16.** If a sole is consistent, then its general solution can be found by writing each of the basic variables in terms of the free variables.

Proof. According to Observation 14, once a sole is in REF, each basic variable appears in only one row. Therefore, in that row, all the other variables are free. By solving for the basic variable in that row, we can write it in terms of only the free variables. It is that fact that makes row reduction work in the sense that it leads to a solution every time one exists. Moreover, any specific solution to our sole can be recovered from the general one by assigning the specific values to the free variables. As such, all solutions can be found in this manner.

**Observation 17.** A sole is consistent if and only if\(^2\) it does not have a row of the form \( 0 = \text{nonzero} \) in EF.

Proof. It’s clear that a sole with a row of the form \( 0 = \text{nonzero} \) at any stage, not just in echelon form, cannot be consistent, since no value of the variables can satisfy the equation \( 0 = \text{nonzero} \). On the other hand, if a sole does not have a row of the form \( 0 = \text{nonzero} \) in echelon form, then the process of row reduction can continue to REF, at which point the general solution can be found as described in the Observation 16.

**Definition 18.** A sole whose right-hand sides all equal zero is called homogeneous.

**Observation 19.** A homogeneous sole is always consistent.

Proof. If our sole is homogeneous, (that is, if its right-hand sides all equal zeroes), then none of the row operations can introduce a nonzero to the right-hand side. In particular, there will never be a row of the form \( 0 = \text{nonzero} \). So by Observation 17, our sole will be consistent. Alternatively, we can note that setting all the variables to equal zero will give us a solution.

---

\(^{11}\)Fine print: As long as you’re in the safe confines of this book, EF will mean echelon form and REF will mean reduced echelon form. However, out there in the confusing and inefficient world, REF might mean row echelon form, in which case, RREF would be used for reduced row echelon form. That extra \( R \) (for row) is injected there to differentiate this from column echelon form (CEF) and reduced column echelon form (RCEF). We will skip that extra level of notation for four reasons: first, we will not use column reduction here, so there’s no ambiguity; second, few enough people use column reduction in the world today that we’re not going to waste an extra letter just for them; third, column reduction may have uses for matrices (coming up in Chapter 2), but it doesn’t make sense to add a multiple of one column to another when we’re dealing with soles because that changes the solutions; and fourth, in a pinch, we’ll see that matrix transposes (ditto) allow use to perform column reduction by swapping rows and columns. All of this is to say that EF stands for (row) echelon form and REF stands for reduced (row) echelon form. By using this textbook, you consent to sharing your personal data with Mom & Pop Productions Inc\(^\circledR\).

\(^{12}\)“Statement A if and only if statement B” means that statement A implies statement B, and statement B implies statement A. When proving such an if-and-only-if claim, we have to prove both directions.
Definition 20. The solution to a homogeneous sole where all the variables equal zero is called the *trivial solution*.

Observation 21. *A sole cannot have 5 solutions or 47 solutions. The number of solutions is either none, one, or infinity.*

**Proof.** By Observation 16, the general solution of any sole can be found by writing each of the basic variables in terms of only the free ones, where the free ones can equal anything they want. As such, either our sole has at least one free variable, in which case it has infinitely many solutions, or else all the variables are basic, in which case there is only one solution. □

Definition 22. In case the rank of a sole equals \( \min(m, n) \), we say our sole has *full rank*.

‘Full rank’ just means that the rank is as large as it can be under the restrictions of Observation 12. For instance, a \( 3 \times 4 \) sole can have, a priori, ranks of 0, 1, 2, or 3; if such a sole claims it has rank 4, it must be lying. So a full rank \( 3 \times 4 \) system must have rank 3. What can you conclude if a \( 3 \times 4 \) sole has full rank? What about a \( 4 \times 3 \) full rank sole?

### 3.2. Unsolicited advice about learning: memorize all the words.

Like much of mathematics, linear algebra is a language. And like any language, there’s an immersive learning process that takes you from vocabulary to grammar to syntax, and from words to sentences to ideas. The progression may not always be that clean, even if it’s useful to tell such a story in your head. But what is incontrovertible is that the ideas cannot exist without the words. It is essential to store the words in an easily accessible part of your memory that, with time, becomes hard-wired second nature. There are two reasons for this.

First, the definitions are always the same, and they are precise. Unlike pretty much all other languages, mathematics relies on definitions that are fixed. While the ambiguity of other languages makes human relationships interesting, the precision of mathematical language makes communication possible. Definitions do not change over time, or based on cultural context (though cult differences do have an impact), or depending on the listener’s childhood traumas. Always and anywhere, a sole is in echelon form if the first nonzero entry in each row is to the right of the first nonzero entry in the row above it. In fact, the phrase “the first nonzero entry in each row is to the right of the first nonzero entry in the row above it” should with time become one word in your brain.

Second, it is nearly impossible to think in ideas without having a solid grasp of the words. Maybe you’re the kind of person who generally develops a big-picture intuition and fills in the details later on. And maybe that has even worked well so far in building an understanding of the world around you. But chances are very good that that will not work with mathematical understanding; even worse, it will create a nefarious illusion of understanding. I am willing to bet you a small but nontrivial amount of money that if you say “I feel like I have a solid understand of the material, but I have no idea what this question is asking,” then you are suffering from an intuition-before-definition syndrome.

Third (and before you object, this is not really third, but rather, first plus second), there is such a thing as mathematical intuition, and it grows organically out of habit and constant exposure. This should strike you as entirely bizarre—since how could there be such a thing as intuition, that most human of emotions, about abstract mathematical structures?—or as entirely self evident—since what is intuition but habit taken over by the unconscious mind? What this should not strike you as is anything in between.

Ultimately, the mathematical endorphin rush and the giddy thrills of aesthetic discovery flow from meditation on the beautiful and intricate connections between all these objects. Memorize the objects’ definitions and their properties, and you will profit.

### 4. More Examples (as an Excuse to Introduce New Notation)

If you think about it, variables are just placeholders. The equations \( x - 2y = 47 \) and \( \diamondsuit - 2\spadesuit = 47 \) contain the same information—I say \( x \) and you say \( \diamondsuit \). So instead of writing the variables over and over again, we can just agree on their location and then write \((1 \quad -2 \quad | \quad 47)\) to mean \( x - 2y = 47 \).
**Definition 23.** We will use the augmented matrix\(^{13}\) notation for soles as follows.

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
    \vdots &
    \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]

is the same as

\[
\begin{pmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} & | & b_1 \\
    a_{21} & a_{22} & \cdots & a_{2n} & | & b_2 \\
    \vdots & \vdots & & \vdots & | & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn} & | & b_m
\end{pmatrix}
\]

The austerity of the augmented matrix notation allows row operations to rightly focus on the coefficients. For instance, adding a multiple of one row to another is the same as adding one multiple of that row’s coefficients to the other row’s coefficients. As long as we remember what the notation means, we can apply row operations much more efficiently, and in a way, we will be more attuned the heart of the matter\(^{14}\).

Additionally, now that we’re old hands at row reduction, we’re going to stop annotating our arrows with the row operations used; those should be evident from context, and indeed, it may be useful cult-practice to annotate them yourself.

**Example 24.** This is how Example 1 looked in the classic notation:

\[
\begin{align*}
    x + y &= -1 \\
    -x + 2y &= 2 \\
    3y &= 1 \\
    x &= -\frac{1}{3} \quad y = \frac{4}{3}
\end{align*}
\]

And this is how it looks in the new & improved notation:

\[
\begin{pmatrix}
    1 & 1 & -1 \\
    -1 & 2 & 2
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
    1 & 1 & -1 \\
    0 & 3 & 1
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
    1 & 1 & -1 \\
    0 & 1 & \frac{1}{3}
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
    1 & 0 & -\frac{4}{3} \\
    0 & 1 & \frac{4}{3}
\end{pmatrix}
\]

We used to omit zero coefficients in the old notation (3y = 1 instead of 0x + 3y = 1) since there was no ambiguity in such an expression; however, we do write the zero coefficients down in the augmented matrix for added clarity and symmetry. Notice how easy it is to read off the solution \(x = -\frac{4}{3}, y = \frac{4}{3}\) from the augmented matrix REF.

**Example 25.** For completeness, we also revisit Example 2, noting only the initial form, EF, and REF of the sole in both classical and augmented matrix notations.

<table>
<thead>
<tr>
<th>Classical Notation</th>
<th>Augmented Matrix Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 = -1)</td>
<td>(\begin{pmatrix} 1 &amp; -2 &amp; 2 &amp; -3 &amp; 1 &amp; -1 \end{pmatrix})</td>
</tr>
<tr>
<td>(-2x_1 + 4x_2 - 4x_3 + 5x_4 = 1)</td>
<td>(\begin{pmatrix} -2 &amp; 4 &amp; -4 &amp; 5 &amp; 0 &amp; 1 \end{pmatrix})</td>
</tr>
<tr>
<td>(3x_1 - 4x_2 + 8x_3 - 5x_4 + x_5 = 1)</td>
<td>(\begin{pmatrix} 3 &amp; -4 &amp; 8 &amp; -5 &amp; 1 &amp; 1 \end{pmatrix})</td>
</tr>
<tr>
<td>(x_1 + 4x_3 + x_5 = 2)</td>
<td>(\begin{pmatrix} 1 &amp; 0 &amp; 4 &amp; 0 &amp; 1 &amp; 2 \end{pmatrix})</td>
</tr>
<tr>
<td>(x_1 - 2x_2 + 2x_3 - 3x_4 + x_5 = -1)</td>
<td>(\begin{pmatrix} 1 &amp; -2 &amp; 2 &amp; -3 &amp; 1 &amp; -1 \end{pmatrix})</td>
</tr>
<tr>
<td>(2x_2 + 2x_3 + 4x_4 - 2x_5 = 4)</td>
<td>(\begin{pmatrix} 0 &amp; 2 &amp; 2 &amp; 4 &amp; -2 &amp; 4 \end{pmatrix})</td>
</tr>
<tr>
<td>(-x_4 + 2x_5 = -1)</td>
<td>(\begin{pmatrix} 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 \end{pmatrix})</td>
</tr>
<tr>
<td>(0 = 0)</td>
<td>(\begin{pmatrix} 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{pmatrix})</td>
</tr>
<tr>
<td>(x_1 + 4x_3 + x_5 = 2)</td>
<td>(\begin{pmatrix} 1 &amp; 0 &amp; 4 &amp; 0 &amp; 1 &amp; 2 \end{pmatrix})</td>
</tr>
<tr>
<td>(x_2 + x_3 + 3x_5 = 0)</td>
<td>(\begin{pmatrix} 0 &amp; 1 &amp; 1 &amp; 0 &amp; 3 &amp; 0 \end{pmatrix})</td>
</tr>
<tr>
<td>(x_4 - 2x_5 = 1)</td>
<td>(\begin{pmatrix} 0 &amp; 0 &amp; 0 &amp; 1 &amp; -2 &amp; 1 \end{pmatrix})</td>
</tr>
<tr>
<td>(0 = 0)</td>
<td>(\begin{pmatrix} 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \end{pmatrix})</td>
</tr>
</tbody>
</table>

**4.1. Examples with morals.** Mathematics is a human discipline because of its reliance on learning from experience. This is therefore a fundamentally human section for all the generalities we will construct from our specific anecdotes.

**Example 26.** A \(3 \times 2\) sole walks into the room. We decide to solve it using augmented matrix notation.

\[
\begin{align*}
    u + 2v &= 1 \\
    2u + 3v &= 0 \\
    -u - v &= -2
\end{align*}
\]

\[
\begin{pmatrix}
    1 & 2 & 1 \\
    2 & 3 & 0 \\
    -1 & -1 & -2
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
    1 & 2 & 1 \\
    0 & -1 & -2 \\
    0 & 1 & -1
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
    1 & 2 & 1 \\
    0 & -1 & -2 \\
    0 & 0 & -3
\end{pmatrix}
\]
Our sole is in EF. From the third line (0 = −3) we conclude that the system is inconsistent—there are no solutions. We can therefore stop the row reduction process and move on to bigger & better things. But first, let’s learn something from this example so that our effort is not wasted.

In a 3 × 2 sole, we knew all along from Observation 11 that we weren’t going to get more than two pivots. Said differently, we knew all along that (at least) one of our three rows in EF was not going to have a pivot. Such a row always runs the risk of being of the form 0 = nonzero. So Observation 11 already warned us that our sole may be inconsistent.

**Example 27.** I adjusted the 3 × 2 sole from Example 26 to make it consistent.

\[
\begin{bmatrix}
1 & 2 & 1 \\
0 & 1 & 2 \\
0 & 0 & 0
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 & -3 \\
0 & 1 & 2 \\
0 & 0 & 0
\end{bmatrix} \Rightarrow 15 \ u = -3. \ v = 2.
\]

Can you adjust this example so that the sole remains consistent, but there are instead infinitely many solutions?

**Example 28.** Yet another sole comes to visit, this time with 3 equations in 3 unknowns. We call such a sole **square** because it has the same number of equations as variables.

\[
\begin{bmatrix}
x - y + 3z = 1 \\
-3x + z = -6 \\
x + 2y = 4
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & -1 & 3 & 1 \\
0 & -3 & 10 & -3 \\
0 & 3 & -3 & 3
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & -1 & 3 & 1 \\
0 & -3 & 10 & -3 \\
0 & 0 & 7 & 0
\end{bmatrix}
\]

The sole is now in EF, and so we can already draw some conclusions. Note that the sole is consistent. You may be inclined to object that there’s a row of the form 0 = nonzero; but upon further reflection, you’ll realize that the row is of the form nonzero = 0. According to the definition of augmented matrix notation, that last row (0 0 7 | 0) actually means 7z = 0, which is perfectly legal. Compare this to the inconsistency of Example 26, where (0 0 | −3) meant 0 = −3. Also note that here all three variables are basic, so we can expect there to be only one solution. Onward to REF!

\[
\begin{bmatrix}
1 & -1 & 3 & 1 \\
0 & 1 & -10 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & -1 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 & 0 & 2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0
\end{bmatrix}
\]

As expected, there is only one solution, and that solution is easily read from the REF of our sole:

\[
x = 2 \\
y = 1 \\
z = 0
\]

Square soles are special; here’s one reason why. If, as in this example, we have as many pivots as rows, then necessarily, we have as many pivots as columns. A pivot in every row means we avoid the danger of inconsistency, and a pivot in every column means we avoid infinitely many solutions. While those two facts are generally independent of each other, in square soles they are one and the same fact.

---

15We’ve been using the single arrow, \(\rightarrow\), to indicate a transition in the form of our sole; later on, we’ll use it to mean something else in the context of linear transformations. The double arrow, on the other hand, has a more specific and universal sense: \(\implies\) always means this implies. Regardless of how you might have used it in your former, more troubled lives, \(\Rightarrow\) never means this equals. (More succinctly, \(\implies\neq\).)
Example 29. Dig if you will the following $4 \times 5$ sole.

$$\begin{pmatrix} x_1 + 2x_2 + 3x_4 - x_5 & = & 2 \\ 3x_1 + 6x_2 + 12x_4 + 4x_5 & = & 7 \\ x_1 + 2x_2 - x_3 + 5x_4 - 3x_5 & = & 3 \\ 2x_1 + 4x_2 + 2x_3 - x_4 + 3x_5 & = & 0 \end{pmatrix}$$

Can you guess in advance whether it’s going to be consistent or not? Let’s say I spoiled the ending for you and said this sole is consistent, can you guess whether there will be one solution or infinitely many solutions? Make your bets.

$$\begin{pmatrix} 1 & 2 & 0 & 3 & -1 \\ 3 & 6 & 0 & 12 & 4 \\ 1 & 2 & -1 & 5 & -3 \\ 2 & 4 & 2 & -1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 3 & -1 \\ 0 & 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 & 7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 3 & -1 \\ 0 & 0 & -1 & 2 & -2 \\ 0 & 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix}$$

In EF, we pause to take stock. The pivots are 1, −1, 3, and 8, the basic variables are $x_1, x_3, x_4,$ and $x_5$, and the rank equals 4. The sole is consistent because there are no rows of the form 0 = nonzero in echelon form; in fact, regardless of what the right-hand side equaled, this sole never ran the danger of being inconsistent. Also, the presence of a free variable implies that the sole will have infinitely many solutions. While we couldn’t know in advance that our sole was going to be consistent, we could have deduced from the beginning that the sole couldn’t have exactly one solution. How? Hint: Observation 11. Now we continue to REF.

$$\rightarrow \begin{pmatrix} 1 & 0 & 3 & -1 & 0 \\ 0 & 0 & 1 & 7/3 & 1/3 \\ 0 & 0 & 0 & 1 & -1/8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 3 & -1 & 15/8 \\ 0 & 1 & -2 & 0 & -3/4 \\ 0 & 0 & 0 & 0 & 1/8 \end{pmatrix}$$

$$\begin{align*}
\text{EF: } & x_1 = -2s \\
& x_2 = s \\
& x_3 = 1/2 \\
& x_4 = 5/8 \\
& x_5 = -1/8
\end{align*}$$

Both this sole and the one in Example 2 have infinitely many solutions. However, the fact that there is one free variable here and two free variables in Example 2 tells us that, in some sense, Example 2 has more solutions than this example. We will make that sense more precise in Chapter 3 when we learn all about dimension.

Example 30. Let’s end our bountiful example run-in with a homogeneous $3 \times 3$ sole.

$$\begin{align*}
2s - u & = 0 \\
-s + t + u & = 0 \\
3s + t - u & = 0
\end{align*}$$

$$\begin{pmatrix} 2 & 0 & -1 \\ -1 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 1 & 1/2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{pmatrix}$$

In EF we find no rows of the form 0 = nonzero, so we conclude that our sole is consistent. This is no big surprise, since all homogeneous soles are consistent. Our pivots are 1 and 2, our basic variables are $s$ and $t$, and our rank equals 2.

$$\begin{pmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} s = a/2 \\ t = -a/2 \\ u = a \end{pmatrix}$$

Note that if $u = 0$, then $s = 0$ and $t = 0$. This is the specific, trivial solution that we always expect to be part of the general solution of a homogenous system.

We reaffirm the wisdom we learned from Example 28. A row of zero coefficients is the same as saying a row coefficients, or a row without a pivot; and since the sole is square, a row without a pivot implies a column without a pivot, which in turn implies a free variable. This is an exclusive bonus fact about square soles that is not true in general.
4.2. **A brief summary of facts we learned from examples in this section.** These are by no means all the morals we can learn from the above examples, to say nothing about all the morals we can learn from linear equations in general. However, instead of despairing at the seeming infinity of what can be known vis-à-vis the finite time we have, we instead revel in the thrill of discovery that each new insight affords us.

- A $3 \times 2$ sole always runs the risk of inconsistency because, in EF, that sole will have at least one row of all zero coefficients. Such a row will run the risk of being of the form $0 = \text{nonzero}$ in case its right-hand side doesn’t equal zero.
- A $3 \times 3$ sole has a curious connection between its rows and its columns. If in EF there is a row whose coefficients all equal zero, then one of the variables must be free. In general, there is no connection between rows without pivots and columns without pivots, but square soles are special.
- A $4 \times 5$ sole cannot have exactly one solution. That’s because a $4 \times 5$ sole cannot have more than four pivots, so if it is consistent, then one of its five variables must be free. A $4 \times 5$ sole will have a free variable regardless of whether it has a row of zeroes in EF.
- A consistent sole with one free variable has fewer solutions than a consistent sole with two free variables, even though both soles have infinitely many solutions.

All the above wisdom relates to the interplay between rank, number of equations, and number of unknowns. While the number of specific instances of the interplay can be endless, the principles of the interplay are few and simple\textsuperscript{16}.

5. **The Geometry of Linear Equations**

You may be excited to read this section because you are a visual learner. Our cult welcomes all learners and celebrates their differences. However, we also caution that “I am a [favorite adjective] type of learner” is a double-edged sword. Self knowledge is helpful; it can also lead to complacency and psychological blockage. The cult leadership council strongly advises you that [less favorite adjective] type of learning is not an excuse to avoid certain perspectives, it is an opportunity to expand minds and horizons. What can happen to you at the micro level has been characteristic of the most fundamental leaps in mathematical understanding at the macro level. Linear algebra is as much geometric as it is algebraic. Reconcile both aspects and you will achieve enlightenment.

Having said that... In what follows, we’ll talk about the ‘dimension’ of a space. For now, this is an undefined idea that you will understand in reference to your Platonic geometric foreknowledge. Well will make the term more precise, and more linear algebraic, in future chapters.

5.1. **Two variables.** As you may remember from kindergarten, a linear equation in two variables can be represented geometrically by a line in the plane\textsuperscript{17}. For instance, the equation $x + 2y = 4$ corresponds to a line in $\mathbb{R}^2$ with with slope $-\frac{1}{2}$ and $y$-intercept 2. If we were to row reduce this $1 \times 2$ sole to REF and find the general solution, we’d get

$$\begin{pmatrix} 1 & 2 \\ 4 \end{pmatrix} \rightarrow \text{already in REF!} \Rightarrow x = 4, \quad y = \frac{-2t}{t}$$

Each point $(x, y)$ on the line is a specific solution to our sole, and the line itself is the general solution.

---

\textsuperscript{16}Complexity of behavior does not necessarily stem from structural complexity. Some of the more pleasingly complex systems, like linear algebra, say, arise from the interactions between a few simple rules.

\textsuperscript{17}Not a plane, but the plane, $\mathbb{R}^2$, the canonical, real, two-dimensional Cartesian $xy$-plane that you know and love
Similarly, an \( m \times 2 \) sole corresponds to \( m \) straight lines in \( \mathbb{R}^2 \), and the solution of such a sole corresponds to all points \((x, y)\) that are on all \( m \) lines simultaneously. If the sole is inconsistent, then there is no point through which all the lines pass.

If the sole is consistent, then there is either one solution (all the lines have exactly one point in common, though some of the lines might coincide) or there are infinitely many solutions (all \( m \) equations actually define the same line).

5.2. Three variables. Analogously, an equation in 3 variables defines a plane\(^{18}\) in three dimensions\(^{19}\). This might not be as familiar an object to you as lines are, but the geometry could still make sense on an intuitive\(^{20}\) level.

Solving an \( m \times 3 \) sole is the same as finding all points \((x, y, z)\) in \( \mathbb{R}^3 \) where all \( m \) planes intersect simultaneously. If the system is inconsistent, then there is no point through which all the planes pass. If the system is consistent, then there is either one solution (all the planes have exactly one point in common) or there are infinitely many solutions (either all the planes coincide, or they intersect in a straight line).

One way to guarantee that lines (and planes) share a point in common with each other is to insist that they pass through the origin. From a 2-dimensional (or 3-dimensional) dimensional point of view, we can think of Observation 19 as follows. An \( m \times 2 \) sole (or \( m \times 3 \) sole) can be represented as a system of \( m \) straight lines.

---

\(^{18}\)A plane, not the plane. This sounds like me just having fun, and it is. But it foreshadows vector spaces, coming in Chapter 3, and linear transformations, coming in Chapter 4.

\(^{19}\)Not all kindergarten sections cover this topic. But Calculus 3 does. You should take Calculus 3.

\(^{20}\)Beware of intuition!
lines (or planes) all of which pass through the origin. Hence they all intersect (at least) at the origin, namely at the point \((0, 0)\) (or \((0, 0, 0)\)).

5.3. Even more variables. Higher-dimensional analogues exist and follow a very similar line of thinking, though they have fancier sounding names.

**Definition** A hyperplane in \(\mathbb{R}^n\) consists of all points \((x_1, x_2, \cdots, x_n)\) that satisfy the equation \(a_1x_1 + \cdots + a_nx_n = b\), where \(a_1, \cdots, a_n, b\) are real numbers, and at least one of the \(a_i\) is nonzero.

A hyperplane is best understood by analogy\(^{21}\). Since \(a_1x + a_2y = b\) can be represented by a line in \(\mathbb{R}^2\) and \(a_1x + a_2y + a_3z = b\) can be represented by a plane in \(\mathbb{R}^3\), then lines in \(\mathbb{R}^2\) and planes in \(\mathbb{R}^3\) are hyperplanes. More generally, a hyperplane is a linear object (like a line or plane) that has dimension one less than the dimension of the space that it belongs to. Higher-dimensional hyperplanes are hard to visualize. Not impossible, though. Just close your eyes and think in analogy...

**Exercises**

(1) Let’s start out with some drills. For each of the following systems of linear equations, row-reduce to echelon form, identify the basic and free variables, and determine the rank. If the system is consistent, row-reduce to reduced echelon form and find all solutions. Keep track of each step by annotating the row operations that you applied. You may use the abbreviated notation if you’re careful.

\[
\begin{align*}
(a) \quad x - 2y &= -3 \\
\quad x - &y = -1 \\
(b) \quad 2x - 3y &= -1 \\
\quad -6x + &9y = 4 \\
(c) \quad x + 2y + 3z &= -1 \\
\quad y + 2z &= 4 \\
\quad 2x + y &= -5 \\
(d) \quad -u + &v = 3 \\
\quad 2u + &2v = 2 \\
\quad 3u - &v = -5 \\
(h) \quad -v + 4x &= y \quad = 1 \\
\quad 2u - &v + 5y = 4z \quad = 2 \\
\quad -u + &2x - 3y + 2z = 3 \\
(i) \quad x_1 + &2x_2 - 3x_3 - 2x_4 + 4x_5 = 1 \\
\quad 2x_1 + &5x_2 - 8x_3 - x_4 + 6x_5 = 4 \\
\quad x_1 + &4x_2 - 7x_3 + 5x_4 + 2x_5 = 8 \\
(j) \quad x_1 + &2x_2 - 3x_3 - 2x_4 + 4x_5 = 0 \\
\quad 2x_1 + &5x_2 - 8x_3 - x_4 + 6x_5 = 0 \\
\quad x_1 + &4x_2 - 7x_3 + 5x_4 + 2x_5 = 0 \\
\end{align*}
\]

(2) (a) By picking specific values for the free variables, find two different solutions to part (1i)—call them \(x_1 = a_1, \cdots, x_5 = a_5\) and \(x_1 = b_1, \cdots, x_5 = b_5\). Compute their difference: \(c_1 = a_1 - b_1, \cdots, c_5 = a_5 - b_5\). Is \(x_1 = c_1, \cdots, x_5 = c_5\) a solution to part (1i)? How about to part (1j)?

(b) Show that what you saw in part (2a) was no accident: the difference of any two specific solutions to part (1i) will always be a solution to part (1j).

---

\(^{21}\)There’s a wonderful book about dimensional analogy. It is called *Flatland: A Romance of Many Dimensions*. It was published in 1884. Read it now.
(c) Is that true in general: Will the difference between two solutions to a nonhomogeneous sole 
be a solution to the corresponding homogeneous sole (that is, the sole with the same left-hand 
sides and a homogeneous right-hand side)?

(3) (a) Sketch the three straight lines

\[
\begin{align*}
2x + y &= 2 \\
x - y &= 0 \\
y &= 3
\end{align*}
\]

(b) Can these three equations be solved simultaneously?
(c) What happens to the figure if all right hand sides are zero?
(d) Is there any nonzero choice of right hand sides which allows the three lines to intersect at the 
same point?

(4) Consider the following system of 2 equations in 3 variables:

\[
\begin{align*}
x + 2y - z &= 3 \\
2x - y + z &= -1
\end{align*}
\]

(a) Find a third equation to include in the system so that there is a unique solution to all three 
equations.
(b) Find a third equation which has no coefficients in common with the first two (try not to use 
the same equations as the first two with only modifications to the right-hand-side; make it 
interesting!) such that there are no solutions.
(c) Find a third equation which has no coefficients in common with the first two (i.e. make it 
interesting again!) such that there are infinitely many solutions.

(5) Find all values of \(c\) such that the following system of equations:

(a) does not have any solutions.
(b) has only one solution.
(c) has infinitely many solutions.

\[
\begin{align*}
x + y &= 2 \\
x - y + cz &= -1 \\
2x + cy + 8z &= 0
\end{align*}
\]

(6) Find all values of \(a\) and \(b\) such that the following system of equations:

(a) does not have any solutions.
(b) has only one solution.
(c) has infinitely many solutions.

\[
\begin{align*}
2x + 3y + z &= 0 \\
x + 4y - 2z &= 2 \\
x + y + az &= b
\end{align*}
\]

(7) In this exercise, the polynomial coefficients become the unknowns!
(a) Find a straight line which passes through the points \((1, 1)\) and \((3, -2)\).
(b) Find a parabola which passes through the points \((1, 1)\) and \((3, -2)\).
(c) Find a parabola which passes through the origin and points \((1, 1)\) and \((3, -2)\).
(d) Find a polynomial \(p(x)\) of degree 3 such that \(p(1) = 1, p(2) = 2, p(3) = 3\) and \(p(4) = 4\).
(e) Exploration. Formulate a conjecture which generalizes these computations.

(8) We continue to expand on the ideas in the previous exercise.
(a) Find two polynomials \(p(x) = ax + b\) and \(q(x) = cx + d\) whose sum equals \(x + 1\) and their 
difference \(p - q\) equals \(47x\).
(b) Consider the function \(f(x) = \alpha \sin(x) + \beta \cos(x)\) (this is called a linear combination of \(\sin(x)\) 
and \(\cos(x)\)). Find numbers \(\alpha\) and \(\beta\) if \(f(\pi/4) = 3\) and \(f(\pi/3) = 4\).
(c) Can you find numbers \(\alpha\) and \(\beta\) such that \(f(x) = x^2 + x + 1\)? If so, we say that \(x^2 + x + 1\) can 
be written as a linear combination of \(\sin(x)\) and \(\cos(x)\).

(9) Using the lingo of the previous exercise, we say that a polynomial \(q(x)\) can be written as a linear 
combination of the polynomials \(p_1(x), p_2(x), \cdots, p_k(x)\) if we can find numbers \(a_1, a_2, \cdots, a_k\) such 
that

\[
a_1 p_1(x) + a_2 p_2(x) + \cdots + a_k p_k(x) = q(x)
\]
For instance, any polynomial of degree 3 or less can be written as a linear combination of $x^3$, $x^2$, $x$ and 1. Determine whether $q(x) = 2x^3 - x + 4$ can be written as a linear combination of each of the following sets of polynomials.

(a) $p_1(x) = x^3$
$p_2(x) = x^3 + x^2$
$p_3(x) = x^3 + x^2 + x$
$p_4(x) = x^3 + x^2 + x + 1$

(b) $p_1(x) = 2x^3 + x^2 - x - 2$
$p_2(x) = -x^3 + x^2 + 1$
$p_3(x) = x^2 + 3x^2 + 1$
$p_4(x) = x^3 + 3x^2 + 2x$

(c) $p_1(x) = 2x^3 + 2x^2 + x + 1$
$p_2(x) = 2x^3 + 4x^2 + 3x - 2$

(10) Equation Island is inhabited by chickens and foxes. Eighty percent of the chickens are hens, and each hen has 2 chicks every year. Half the foxes are female, and each female has one pup every year. Additionally, each of the adult foxes (i.e. you don’t have to take the newborns into account) eats 12 chickens a year. Denote by $f_n$ and $c_n$ the number of foxes and chickens, respectively, alive on the island at the beginning of year $n$.
(a) Find a formula for $f_{n+1}$ in terms of $f_n$ and $c_n$. Do the same for $c_{n+1}$.
(b) Find a formula of $f_{n+2}$ in terms of $f_n$ and $c_n$. Do the same for $c_{n+2}$.
(c) Can you find a general formula for $f_{n+k}$ in terms of $f_n$ and $c_n$? How about for $c_{n+k}$.

(11) Consider the “generic” system of 2 equations in two unknowns.

\[
\begin{align*}
ax + by &= \alpha \\
cx + dy &= \beta 
\end{align*}
\]

(a) Suppose $ad \neq bc$. Show that this system has a unique solution. (Hint: divide your answer into two cases, depending on whether $a$ equals zero or not.)
(b) Suppose $ad = bc$. What can you say about the existence of solutions?

(12) Observe that the two equations $x - 2y = 2$ and $2x + y = -6$ (call them (1) and (2), respectively) can be combined to produce the equation $x - y = 0$ (call it (3)), since since 3/5 times the first equation plus 1/5 times the second equation yields the third: $\frac{3}{5} \cdot (1) + \frac{1}{5} \cdot (2) = (3)$. Show that the first two equations in the following sole can be combined to yield the third equation. Express that description explicitly.

\[
\begin{align*}
x - y + 3z &= 1 \\
-3x + z &= -6 \\
x + 2y + 7z &= 4
\end{align*}
\]

It may help to row reduce the sole to EF.

(13) How do soles respond to a small change? We subtract 0.01 from the first coefficient to find out.
(a) Compare the solutions to the following two soles.

\[
\begin{align*}
x + y &= 1 & 0.99x + y &= 1 \\
x + 2y &= 3 & x + 2y &= 3
\end{align*}
\]

(b) Compare the solutions to the following two soles.

\[
\begin{align*}
x + y &= 1 & 0.99x + y &= 1 \\
x + 1.01y &= 3 & x + 1.01y &= 3
\end{align*}
\]

(c) Write a haiku describing your feelings about this experience.

(14) Consider the sequence 0, 1, 1, 3, 5, 11, 21, 43, 85, · · · . Let $x_n$ denote the $n$th term of this sequence; so $x_0 = 0$, $x_1 = 1$, $x_2 = 1$, $x_3 = 3$, etc.
(a) Write down a linear equation which defines $x_n$ in terms of previous two terms in the sequence.
(b) Can you find a general formula for $x_n$ in terms of $n$? (In general, this is difficult without eigentheory).
(c) Can you guess the limit of ratios of successive terms, $x_{n+1}/x_n$, as $n \to \infty$? Is that useful for finding an approximate formula for $x_n$?
(15) Repeat previous exercise for 0, 1, 1, 2, 3, 5, 8, \ldots.

(16) Most mathematical questions whose answers are “no!” can be reformulated with an eye towards more of a compromise. For instance, if a system of linear equations is inconsistent, perhaps we can find the next best thing to a solution.
(a) Verify that the following system of equations is inconsistent.
\[
\begin{align*}
-x &+ y = -1 \\
-x &- y = 2 \\
x &+ 2y = 3
\end{align*}
\]
(b) Since we cannot find a solution to the above system, let us try to find solutions which make the left-hand-sides as “close” as possible to their respective right-hand-sides. More precisely, find numbers \(x\) and \(y\) such that the expression
\[
(-x + y + 1)^2 + (x - y - 2)^2 + (x + 2y - 3)^2
\]
is as small as possible.
(c) What would happen if we used the same approach on a consistent system of linear equations?

(17) How expensive is row reduction? How many products and sums would you expect to perform when finding the solutions to an \(m \times n\) sole?

(18) Show that if, in a system of \(m\) equations in \(n\) variables, \(m < n\), then the system either is inconsistent or else has infinitely many solutions (that is, it cannot have only one solution). What more can you say if the system is homogeneous?

(19) Prove or Dare. Determine if each of the following statements is true. If it is, provide a proof. If it is false, demonstrate that with a counterexample, and change the statement in a minimal way to make it true.
(a) If \(x = x_1, y = y_1\) and \(x = x_2, y = y_2\) are two solutions to the equation \(ax + by = c\), then so is \(x = x_1 + x_2, y = y_1 + y_2\). In other words, the sum of solutions is also a solution.
(b) If homogeneous sole
\[
\begin{align*}
ax &+ by = 0 \\
cx &+ dy = 0
\end{align*}
\]
has infinitely many solutions, then \(a = c\) and \(b = d\).
(c) A homogeneous sole whose echelon form has a row of zeroes will have infinitely many solutions.
(d) Let \(r\) be the rank of a system of \(m\) equations in \(n\) unknowns.
(i) If \(r = m\), then the system must be consistent.
(ii) If \(r = n\), then the system can have at most one solution.

(20) Starting with Example 2, delete columns three and five (recall these were the columns without pivots) and row reduce the new sole. Repeat by deleting columns 2 and 3. Compare the sequence of row operations as well as the solution to those of Example 2. Can you generalize?