Overview

In the previous note, we considered how to represent systems of linear equations. Now, we will develop a systematic algorithm, known as **Gaussian Elimination**, to solve arbitrary systems of linear equations, or determine that no solutions exist.

1.1 Example

We may already be able to solve linear systems in a so-called "ad-hoc" manner - manipulating equations according to our intuition until we get a solution or have convinced ourselves that no solution exists. Let's remind ourselves of how we could do that.

Consider the linear system with unknowns x and y:

$$\begin{aligned} 5x + 6y &= 40 \\ 8x + 9y &= 61. \end{aligned}$$

To solve it, one way would be to use the first equation to express y in terms of x, and then substitute into the second to obtain a single linear equation involving only x. Rearranging the first equation, we find that

$$\begin{aligned} && 5x + 6y &= 40 \\ \Longrightarrow && 6y &= 40 - 5x \\ \Longrightarrow && y &= \frac{20}{3} - \frac{5}{6}x. \end{aligned}$$

Now, substituting into the second equation, we obtain

$$\begin{aligned} && 8x + 9y &= 61 \\ \Longrightarrow && 8x + 9\left(\frac{20}{3} - \frac{5}{6}x\right) &= 61 \\ \Longrightarrow && \frac{1}{2}x + 60 &= 61 \\ \Longrightarrow && \frac{1}{2}x &= 1 \\ \Longrightarrow && x &= 2. \end{aligned}$$

Finally, substituting this value for x into our equation for y in terms of x, we find

$$\begin{aligned} y &= \frac{20}{3} - \frac{5}{6}x \\ &= \frac{20}{3} - \frac{5}{6} \cdot 2 \\ &= 5. \end{aligned}$$

This ad-hoc, substitution-based approach worked well for this small system. However, as we start to consider larger systems, potentially with thousands or even millions of variables, manually solving them quickly becomes infeasible.

We need to develop a systematic approach to solve linear systems that generalizes well as the numbers of equations and variables increases. After developing this approach, we should be able to prove (at least to a certain level of rigor) that our approach is guaranteed to work whenever a solution to a linear system exists.

1.2 Gaussian Elimination

Gaussian elimination is an **algorithm** (a sequence of programmatic steps) that accomplishes this task. Specifically, it can be used to solve any arbitrarily large system of linear equations, or decide that no solution exists. Gaussian elimination isn't the *only* algorithm that does this (for instance, we could try writing an algorithm that formalizes the substitution-based approach that we tried above), but it's pretty good!

Side note: Gaussian elimination was named after Carl Friedrich Gauss, a German mathematician from the 18th century. Despite being its namesake, he did not invent it, though he contributed to its development. As it turns out, Gaussian elimination was initially developed in China over 2000 years ago!

In Europe, Gaussian elimination was refined over the course of 200 years by mathematicians including Newton, Rolle, and Gauss. To quote Newton,

> And you are to know, that by each Æquation one unknown Quantity may be taken away, and consequently, when there are as many Æquations and unknown Quantities, all at length may be reduc'd into one, in which there shall be only one Quantity unknown.

To read about how it evolved, check out `https://www.ams.org/notices/201106/rtx110600782p.pdf`!

1.2.1 Example

Before we look at a precise formulation of Gaussian elimination, let's look at an example of how it works, to build our intuition. Specifically, let's try solving the following example using Gaussian elimination:

$$\begin{aligned} 5x+6y+z &= 43 \quad (1) \\ 8x+9y+2z &= 67 \quad (2) \\ x+y+4z &= 19 \quad (3) \end{aligned}$$

Intuitively, the basic idea behind Gaussian elimination is to use each equation to *eliminate* one variable from all the subsequent equations, until we end up with an equation with just one unknown, which we can directly solve. (we'll make this intuition more rigorous in a moment). Let's try using the first equation to eliminate one of our unknowns - say, x. It would make sense to first **multiply** the first equation by $1/5$, in order to remove the coefficient of 5 in front of x. Thus, we obtain

$$x+\frac{6}{5}y+\frac{1}{5}z=\frac{43}{5}. \quad (4)$$

Now, we will try to eliminate the variable x in (2) and (3). One way would be to write x in terms of y and substitute, like we did earlier. Instead, however, Gaussian elimination requires us to **add multiples** of (4) from (2) and (3), in order to accomplish the same goal (we'll see why we're currently trying to avoid substitution in a moment, when we make this process more rigorous).

Let's try eliminating x from (2) first. What multiple of (4) would be best to use? Well, since (4) has an x term, and we'd like it to cancel out with the $8x$ term from (2), it makes sense to multiply (3) by -8 and add it from (2) to produce (5):

$$\begin{aligned} & (8x+9y+2z)-8\cdot\left(x+\frac{6}{5}y+\frac{1}{5}z\right)=67-8\cdot\frac{43}{5} \\ \implies & \frac{-3}{5}y+\frac{2}{5}z=-\frac{9}{5} \quad (5) \end{aligned}$$

Similarly, it would make sense to simply subtract (4) itself from (3) (equivalently, to multiply (4) by -1 and add it to (3)), since both (3) and (4) have an x term with a coefficient of unity. This produces

$$\begin{aligned} & (x+y+4z)-\left(x+\frac{6}{5}y+\frac{1}{5}z\right)=19-\frac{43}{5} \\ \implies & -\frac{1}{5}y+\frac{19}{5}z=\frac{52}{5}. \quad (6) \end{aligned}$$

Now, observe that (5) and (6) together form a linear system with just two unknowns: y and z, since x has been eliminated from the latter two equations. Now, let's see if we can repeat this process, using (5) to eliminate y from (6). First, as before, we should simplify things by scaling (5) to remove the coefficient of y. Multiplying (5) by $-5/3$, we obtain

$$y-\frac{2}{3}z=3. \quad (7)$$

Again, we'd like to add some scalar multiple of (7) from (6), in order to eliminate y from (6). Since (6) has a $(-1/5)y$ term, it makes sense to multiply (7) by $1/5$ and add it to (6), which produces

$$\begin{aligned} -\frac{1}{5}y+\frac{19}{5}z+\frac{1}{5}\left(y-\frac{2}{3}z\right) &= \frac{52}{5}+\frac{1}{5}\cdot 3 \\ \implies \qquad \frac{11}{3}z &= 11. \qquad (8) \end{aligned}$$

Now, (8) is a simple linear equation in only one variable, which we know how to immediately solve. Awesome! Rearranging a little, we see that

$$z = 3.$$

When we did our initial example above, after solving for x, we substituted it into an equation for y in terms of x. Can we do something similar here? Specifically, do we have an equation for x or y in terms of just z?

Notice that the equation we used to eliminate y (Eq. 7) had no variables "before" y, since they were eliminated, and doesn't even have a coefficient for y. Pulling all terms except for y onto the right-hand-side of the equality, we obtain

$$y = 3+\frac{2}{3}z.$$

Perfect! Substituting in $z = 3$, we find that

$$y = 3+\frac{2}{3}\cdot 3 = 5.$$

Can we do the same thing again? Well, when we eliminated x using (4), we again needed an equation with a unit coefficient for x. Pulling all the terms in (4) except for x to the right-hand-side, we obtain

$$x = \frac{43}{5}-\frac{6}{5}y-\frac{1}{5}z.$$

Again, this is looking pretty good. Substituting in our known values for y and z, we find that

$$x = \frac{43}{5}-\frac{6}{5}\cdot 5-\frac{1}{5}z = 2,$$

so we've solved for all of our unknowns using Gaussian elimination. Awesome!

1.2.2 Steps of Gaussian Elimination

Let's take a moment to reflect on the approach we just used.

- First, we selected an equation involving x (possibly with some coefficient) and scaled it to make the x coefficient unity.
- Then, we added multiples of this equation from all the other equations to eliminate x, producing a system with one fewer unknown and one fewer equation.
- We then repeated the first two steps until we arrived at an equation with exactly one unknown, which we could solve directly.

- Finally, we substituted the known value of the final unknown into a previous equation to recover the last two unknowns, and continued substituting until we recovered all of our unknowns!

These are the key steps of Gaussian elimination. The first three steps are known as **row reduction**, and the final step is known as **back-substitution**.

Now that we know how to perform the steps of Gaussian elimination for systems where a solution is known to exist, it is important to ask ourselves *why* these steps work. In particular, even if we have some intuition for why they work in cases when a system of equations has a unique solution, we'd like to show that these steps remain valid even when working with a system with zero or infinitely many solutions.

1.2.3 Operations

The key idea behind Gaussian elimination is that of "invertible" operations. As we manipulate our equations, we want to preserve their set of solutions. In particular, we neither want to *introduce* new solutions in the process, nor to *remove* potential solutions. To do so, we use three operations that we are certain will *never* change the solution set of a system, and then apply these operations repeatedly in order to solve a linear system. By only applying these operations, we can be confident that our approach will never yield a wrong answer, since the solution set of our system is preserved throughout.

These operations, which we have just seen, are as follows:

1. **Multiplying an equation by a *nonzero* scalar constant.** For instance, if we have the equation

   $$2 \times a + 3 \times b = 4,$$

   we can multiply it by the nonzero scalar -2 to obtain

   $$-4 \times a + (-6) \times b = -8.$$

   Expressed as a single operation, we can write

   $$\begin{aligned} & 2 \times a + 3 \times b = 4 \\ \implies \quad & -4 \times a + (-6) \times b = -8. \end{aligned}$$

   Why does this operation preserve all the solutions to a system? Well, consider any particular solution that satisfies the first equation. Clearly, it still satisfies the second equation, so this operation has not removed any potential solutions.

   But does it introduce a new solution? Consider any particular solution to the second equation. Notice that we can multiply the second equation by the *reciprocal* of our original nonzero scalar multiplier, to obtain the first equation. Thus, this particular solution will also satisfy the first equation. In other words, no solution exists that satisfies the second equation, but not the first. Consequently, the second equation is not only *implied by*, but also *implies* the first equation.

   When each of two equations imply the other, we say that they are *equivalent*, since replacing one with the other does not change their solution set. Notice that, to obtain equivalence, we had to restrict our multiplication to one by a *nonzero* scalar, not an arbitrary scalar. Otherwise, we would not be able to obtain the first equation from the second (since the reciprocal of zero is undefined), so we would only obtain a one-way implication, not two-way equivalence.

2. **Adding a scalar constant multiple of one equation to another.** For instance, if we have the equations

$$\begin{aligned} 5 \times a + 6 \times b &= 7 \qquad (1) \\ 8 \times a + 9 \times b &= 10, \quad (2) \end{aligned}$$

we can multiply the second equation by the scalar 3 and add it to the first, to obtain the new system

$$\begin{aligned} 29 \times a + 33 \times b &= 37 \qquad (3) \\ 8 \times a + 9 \times b &= 10, \quad (2) \end{aligned}$$

Clearly, observe that any solution to the first system will also be a solution to the second, since the first system of equations implies the second. But is the reverse true? Well, observe that equation (1) can be recovered by taking equation (3) and subtracting our scalar (in this case, 3) multiplied by equation (2). In other words, our second system is, not only implied by, but also implies the first system, so it does not introduce any new solutions. Thus, replacing the first system with the second does not change the solution set of our linear system, so this operation is valid.

3. **Swapping two equations.** We have not yet seen when we need to swap two equations (though we will in Example 1.3), but it is clear that the solution set of a linear system of equations does not depend on the order of equations! Therefore, this final operation is clearly valid.

Now we have developed these three operations, we can repeatedly use them in a structured manner to solve arbitrary systems of linear equations, as will be illustrated in the following examples:

Example 1.1 (System of 2 equations): Consider the following system of two equations with two variables:

$$\begin{array}{rcrcl} x & - & 2y & = & 1 \qquad (1) \\ 2x & + & y & = & 7 \qquad (2) \end{array}$$

We would like to find an explicit formula for x and y, but the presence of both x and y in each of the equations prevents this. If we can eliminate a variable from one of the equations, we can get an explicit formula for the remaining variable. To eliminate x from (Eq. 2), we can subtract 2 times (Eq. 1) from (Eq. 2) to obtain a new equation, (Eq. $2'$):

$$\begin{array}{rcrcr} 2x & + & y & = & 7 \\ -2 \times (x & - & 2y & = & 1) \\ & & \Downarrow & & \\ 2x & + & y & = & 7 \\ -2x & + & 4y & = & -2 \\ & & \Downarrow & & \\ & & 5y & = & 5 \qquad (2') \end{array}$$

Scaling (Eq. 1) by the amount that x is scaled in (Eq. 2) allows us to cancel the x term. As a result, we can replace (Eq. 2) with (Eq. $2'$) to rewrite our system of equations as:

$$\begin{array}{rrcrcl} & x & - & 2y & = & 1 \qquad (1) \\ \text{(Eq. 2)} - 2 \times \text{(Eq. 1):} & & & 5y & = & 5 \qquad (2') \end{array}$$

From here, we can divide both sides of (Eq. $2'$) by 5 to see that $y = 1$. We will call this (Eq. $2''$). Next, we would like to solve for x. It would be natural to proceed by substituting $y = 1$ into (Eq. 1) to solve directly for x, and doing this will certainly give you the correct result. However, our goal is to find an *algorithm* for solving systems of equations. This means that we would like to be able to repeat the same sequence of operations over and over again to come to the solution. Recall that to cancel x in (Eq. 2), we:

1. **Scaled** (Eq. 1) by a factor of 2.
2. **Subtracted** (Eq. 1) from (Eq. 2).

To solve for x, we would like to eliminate y from (Eq. 1) using a similar process of scaling and subtracting. Because y is scaled by a factor of -2 in (Eq. 1), we can scale (Eq. $2''$) by -2 and subtract it from (Eq. 1) to cancel the y term. Doing so gives:

$$\begin{array}{rrcrcrl} \text{(Eq. 1)} + 2 \times \text{(Eq. } 2''\text{):} & x & & & = & 3 & (1') \\ \text{(Eq. } 2'\text{)} \nabla\cdot 5\text{:} & & & y & = & 1 & (2'') \end{array}$$

Soon we will generalize this technique so that it can be extended to any number of equations. Right now we will use an example with 3 equations to help build intuition.

Example 1.2 (System of 3 equations): Suppose we would like to solve the following system of 3 equations:

$$\begin{array}{rcrcrcrl} x & - & y & + & 2z & = & 1 & (1) \\ 2x & + & y & + & z & = & 8 & (2) \\ -4x & + & 5y & & & = & 7 & (3) \end{array}$$

As in the 2 equation case, our first step is to eliminate x from all but one equation by adding or subtracting scaled versions of the first equation from the remaining equations. Because x is scaled by 2 and -4 in (Eq. 2) and (Eq. 3) (respectively), we can multiply (Eq. 1) by these factors and subtract it from the corresponding equations:

$$\begin{array}{rrcrcrcrl} & x & - & y & + & 2z & = & 1 & (1) \\ \text{(Eq. 2)} - 2\times \text{(Eq. 1):} & & & 3y & - & 3z & = & 6 & (2') \\ \text{(Eq. 3)} + 4\times \text{(Eq. 1):} & & & y & + & 8z & = & 11 & (3') \end{array}$$

Next, we would like to eliminate y from (Eq. $3'$). First, we can divide (Eq. $2'$) by 3 such that y is scaled by 1:

$$\begin{array}{rrcrcrcrl} & x & - & y & + & 2z & = & 1 & (1) \\ \text{(Eq. } 2'\text{)}\nabla\cdot 3\text{:} & & & y & - & z & = & 2 & (2'') \\ & & & y & + & 8z & = & 11 & (3') \end{array}$$

Now, since y is also scaled by 1 in (Eq. $3'$), we can subtract (Eq. $2''$) from (Eq. $3'$) to get a formula[1] with only z:

$$\begin{array}{rrcrcrcrl} & x & - & y & + & 2z & = & 1 & (1) \\ & & & y & - & z & = & 2 & (2'') \\ \text{(Eq. } 3'\text{)} - \text{(Eq. } 2''\text{):} & & & & & 9z & = & 9 & (3'') \end{array}$$

[1] At this point we have made a decision in our algorithm to eliminate y from (Eq. $3'$) but *not* (Eq. 1). The motivation for this might not be completely evident now, but approaching it this way can be more computationally efficient for certain systems of linear equations — typically if the system has an infinite number of solutions or no solutions.

Dividing (Eq. $3''$) by 9 gives an explicit formula for z:

$$\begin{array}{rrrrrrrcll} & x & - & y & + & 2z & = & 1 & (1) \\ & & & y & - & z & = & 2 & (2'') \\ (\text{Eq. } 3'')\nabla\cdot 9: & & & & & z & = & 1 & (3''') \end{array}$$

At this point, we can see that our system of equations has a “triangular" structure — all three variables are contained in (Eq. 1), two are in (Eq. $2''$), and only z remains in (Eq. $3'''$). If we look back to the previous example with 2 equations, we obtained a similar result after eliminating x from (Eq. 2):

System of 2 Equations

$$\begin{array}{rrrcl} x & - & 2y & = & 1 \\ & & 5y & = & 5 \end{array}$$

System of 3 Equations

$$\begin{array}{rrrrrcl} x & - & y & + & 2z & = & 1 \\ & & y & - & z & = & 2 \\ & & & & z & = & 1 \end{array}$$

This similarity is not coincidental, but a direct result of the way in which we successively eliminate variables in our algorithm moving left to right. To understand why it is useful to have our system of equations in this format, we will now proceed to solve for the remaining variables in this 3-equation example. First, we would like to eliminate z from (Eq. 1) and (Eq. $2''$). As usual, we can accomplish this by scaling (Eq. $3'''$) by the amount z is scaled in (Eq. 1) and (Eq. $2''$) and subtracting this from these equations:

$$\begin{array}{rrrrrcrl} (\text{Eq. } 1) - 2\times (\text{Eq. } 3'''): & x & - & y & & = & -1 & (1') \\ (\text{Eq. } 2'') + (\text{Eq. } 3'''): & & & y & & = & 3 & (2''') \\ & & & & z & = & 1 & (3''') \end{array}$$

Finally, by adding (Eq. $2'''$) to (Eq. $1'$), we can find the solution:

$$\begin{array}{rrrrcrl} (\text{Eq. } 1') + (\text{Eq. } 2'''): & x & & & = & 2 & (1'') \\ & & y & & = & 3 & (2''') \\ & & & z & = & 1 & (3''') \end{array}$$

After obtaining an explicit equation for z using a repetitive process of scaling and subtraction, we were able to obtain an explicit equation for y, and then x, using this same process — this time propagating equations upwards instead of downwards.

So far, the two operations we’ve previously encountered seem to be sufficient to solve every system of equations we’ve encountered. Are there any other operations we might need to perform in addition to scaling and adding/subtracting equations?

Example 1.3 (System of 3 equations): Suppose we would like to solve the following system of 3 equations:

$$\begin{array}{rrrrrcrl} & & 2y & + & z & = & 1 & (1) \\ 2x & + & 6y & + & 4z & = & 10 & (2) \\ x & - & 3y & + & 3z & = & 14 & (3) \end{array}$$

As in the 2 equation case, our first step is to eliminate x from all but one equation. Since x is the first variable to be eliminated, we want the equation containing it to be at the top. However, the first equation does not contain x. To solve this problem, we **swap** the first two equations. Clearly, swapping two equations does not change the system's solution set, so we will obtain the equivalent linear system:

$$\begin{array}{rrcrcrcrl}
\text{(Eq. 2):} & 2x & + & 6y & + & 4z & = & 10 & (1') \\
\text{(Eq. 1):} & & & 2y & + & z & = & 1 & (2') \\
& x & - & 3y & + & 3z & = & 14 & (3)
\end{array}$$

Now we can proceed as usual, dividing the first equation by 2, and then subtracting this from (Eq. $3'$) to eliminate x:

$$\begin{array}{rrcrcrcrl}
\text{(Eq. } 1') \nabla\cdot 2\text{:} & x & + & 3y & + & 2z & = & 5 & (1'') \\
& & & 2y & + & z & = & 1 & (2') \\
\text{(Eq. 3)} - \text{(Eq. } 1''\text{):} & & - & 6y & + & z & = & 9 & (3')
\end{array}$$

Now there is only one equation containing x. Of the remaining two equations, we want only one of them to contain y, so we can divide (Eq. $2'$) by 2 and then add 6 times (Eq. $2''$) to (Eq. $3'$):

$$\begin{array}{rrcrcrcrl}
& x & + & 3y & + & 2z & = & 5 & (1'') \\
\text{(Eq. } 2') \nabla\cdot 2\text{:} & & & y & + & \frac{1}{2}z & = & \frac{1}{2} & (2'') \\
\text{(Eq. } 3') + 6\times \text{(Eq. } 2''\text{):} & & & & & 4z & = & 12 & (3'')
\end{array}$$

Now, we have the "triangular" structure from the previous examples. To proceed, we can divide the last equation by 4 to solve for $z = 3$, and use this to eliminate z from the remaining equations:

$$\begin{array}{rrcrrcrl}
\text{(Eq. } 1''\text{)} -2\times \text{(Eq. } 3'''\text{):} & x & + & 3y & & = & -1 & (1''') \\
\text{(Eq. } 2''\text{)} - \frac{1}{2}\times \text{(Eq. } 3'''\text{):} & & & y & & = & -1 & (2''') \\
\text{(Eq. } 3''\text{)} \nabla\cdot 4\text{:} & & & & z & = & 3 & (3''')
\end{array}$$

Finally, we can subtract 3 times (Eq. $2'''$) from (Eq. $1'''$) to solve for x:

$$\begin{array}{rrrrcrl}
\text{(Eq. } 1'''\text{)} -3\times \text{(Eq. } 2'''\text{):} & x & & & = & 2 & (1'''') \\
& & y & & = & -1 & (2''') \\
& & & z & = & 3 & (3''')
\end{array}$$

1.2.4 Gaussian Elimination with Matrices

Now, let's try to apply our previous approach for solving linear equations on the matrix representation of a linear system.

For convenience, rather than using the system of equations presented above, let's look at the simpler system, we saw in the previous note can be represented as an augmented matrix:

$$\begin{bmatrix} 5x & + & 3y & = & 5 \\ -4x & + & y & = & 2 \end{bmatrix} \qquad \left[\begin{array}{cc|c} 5 & 3 & 5 \\ -4 & 1 & 2 \end{array}\right]$$

In the examples we have seen, there are three basic operations that we can perform to a system of equations, that we know will preserve the solution set of the associated system of linear equations. Let's see how they work when applied to the augmented matrix representation of a system of linear equations:

1. Multiplying a row by a scalar. For example, we can multiply the first row by 2:

$$\begin{bmatrix} 10x & + & 6y & = & 10 \\ -4x & + & y & = & 2 \end{bmatrix} \qquad \left[\begin{array}{cc|c} 10 & 6 & 10 \\ -4 & 1 & 2 \end{array}\right]$$

2. Swapping rows. For example, we swap the 2 rows:

$$\begin{bmatrix} -4x & + & y & = & 2 \\ 5x & + & 3y & = & 5 \end{bmatrix} \qquad \left[\begin{array}{cc|c} -4 & 1 & 2 \\ 5 & 3 & 5 \end{array}\right]$$

3. Adding a scalar multiple of a row to another row. For example, we can modify the second row by adding 2 times the first row to the second:

$$\begin{bmatrix} 5x & + & 3y & = & 5 \\ 6x & + & 7y & = & 12 \end{bmatrix} \qquad \left[\begin{array}{cc|c} 5 & 3 & 5 \\ 6 & 7 & 12 \end{array}\right]$$

Our procedure so far has been to successively eliminate variables using the above steps. A bit more precisely, if we number the variables 1 through n in the order they appear from left to right, to begin Gaussian elimination we eliminate a variable i with the following steps, beginning with $i = 1$ and ending when $i = n$:

1. Swap rows if needed so that an equation containing variable i is contained in row i (in the augmented matrix, this means column i and row i should be nonzero).
2. Divide row i by the coefficient of variable i in this row such that the i^{th} row and column of the augmented matrix is 1.
3. For rows $j = i+1$ to n, subtract row i times the entry in row j and column i to cancel variable i.

So far, the above steps eliminating variables from left to right (operating on equations from top to bottom) proceeded until we found a "triangular" system of linear equations with an explicit equation at the bottom, which we could then propagate upwards to solve for the remaining variables.

This "triangular form" is known as **row echelon form**. More precisely, a matrix is in row echelon form when the following criteria are met:

- All nonzero rows are above all zero rows.
- The leading coefficient of a non-zero row is always to the right of the leading coefficient of the row above it.

Some textbooks will require a third property to be true for row echelon form:

- The leading coefficient of every non-zero row (which we call the *pivot*, and say is in the *pivot position*) is 1.

In addition to row echelon form, there is also **reduced row echelon form**. This requires that: In addition, after the upwards propagation of variables in step (3), we will obtain a matrix with the following properties, in addition to the two mentioned above:

- The matrix is in row echelon form.
- The leading coefficient of every non-zero row (which we call the *pivot*, and say is in the *pivot position*) is 1.
- Each *column* with an element that is in the pivot position of some row has 0s everywhere else.

When all of the above properties are met, we say a matrix is in **reduced row echelon form**, sometimes abbreviated (especially in programming) as `rref`.

Note that, depending on the source, row echelon form is sometimes defined to also require all the leading coefficients of non-zero rows to be normalized to be 1 - however, this is not a requirement in our definition.

We are introducing the terminology of echelon form just for consistency with textbooks. However, such jargon and definitions are not emphasized in this class; what is most important is that you understand the fundamental principles.

For illustrative purposes, the following represents a matrix in row echelon form

$$\left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \tag{1}$$

where $*$ denotes that any value may be entered in that location.

Step 2 corresponds to the “back-substitution" of variables performed in the previous examples. At the conclusion of Step 2, we are left with a matrix satisfying the following two properties:

- The matrix is in row echelon form.
- Each leading entry of a nonzero row is the only nonzero entry in its column.

Such a matrix is said to be in **reduced row echelon form**, sometimes abbreviated (especially in programming) as `rref`. As another illustration, the following represents a matrix in reduced row echelon form

$$\left[\begin{array}{cccc|c} 1 & 0 & * & 0 & * \\ 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{2}$$

By construction, the Gaussian elimination algorithm always results in a matrix that is in reduced row echelon form. Once an augmented matrix is reduced to reduced row echelon form, variables corresponding to columns containing leading entries are called **basic variables**, and the remaining variables are called **free variables**. For example, if we consider the augmented matrix

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

then it corresponds to the system of equations

$$\begin{array}{rrrrl} x_1 & & +2x_3 & & = 3 \\ & x_2 & +4x_3 & & = 5 \\ & & & x_4 & = -8. \end{array} \tag{4}$$

In this system of equations, x_1, x_2, x_4 are all basic variables, and x_3 is a free variable. As will be discussed shortly, the distinction between basic and free variables allows us to characterize all solutions to the system of linear equations (if any exist!).

Why does Gaussian elimination always result in a matrix in reduced row echelon form? And will this result *always* allow us to determine a single explicit solution for any system of equations? The next few examples explore what might happen after these steps are applied. On the left hand side, we will show the system of equations, and on the right hand side, we show the corresponding augmented matrix.

1.2.4.1 Gaussian Elimination Examples

Example 1.4 (Equations with exactly one solution):

$$\begin{bmatrix} 2x & + & 4y & + & 2z & = & 8 \\ x & + & y & + & z & = & 6 \\ x & - & y & - & z & = & 4 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 2 & 4 & 2 & 8 \\ 1 & 1 & 1 & 6 \\ 1 & -1 & -1 & 4 \end{array}\right]$$

First, divide row 1 by 2, the scaling factor on x in the first equation.

$$\begin{bmatrix} x & + & 2y & + & z & = & 4 \\ x & + & y & + & z & = & 6 \\ x & - & y & - & z & = & 4 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 1 & 1 & 1 & 6 \\ 1 & -1 & -1 & 4 \end{array}\right]$$

To eliminate x from the two remaining equations, subtract row 1 from row 2 and 3.

$$\begin{bmatrix} x & + & 2y & + & z & = & 4 \\ & - & y & & & = & 2 \\ & - & 3y & - & 2z & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & -1 & 0 & 2 \\ 0 & -3 & -2 & 0 \end{array}\right]$$

To ensure y is scaled by 1 in the second equation, multiply row 2 by -1. Then, to eliminate y from the final equation, subtract -3 times row 2 from row 3.

$$\begin{bmatrix} x & + & 2y & + & z & = & 4 \\ & & y & & & = & -2 \\ & & & - & 2z & = & -6 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & -2 & -6 \end{array}\right]$$

To scale z by 1 in the final equation, divide row 3 by -2.

$$\begin{bmatrix} x & + & 2y & + & z & = & 4 \\ & & y & & & = & -2 \\ & & & & z & = & 3 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Notice that our matrix is now in *row echelon* form, since the leading coefficient of each nonzero row is to the right of the leading coefficient of the row above it. However, it is not yet in *reduced row echelon form*, since the second and third columns, which each contain an element in the pivot position of a row, have a nonzero element in another row.

Continuing, we then subtract row 3 from row 1 to eliminate z from the first equation.

$$\begin{bmatrix} x & + & 2y & & & = & 1 \\ & & y & & & = & -2 \\ & & & & z & = & 3 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Finally, subtract 2 times row 2 from row 1 to obtain an explicit equation for all variables.

$$\begin{bmatrix} x & & & = & 5 \\ & y & & = & -2 \\ & & z & = & 3 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Observe that our matrix is now in *reduced row echelon form*, since in addition to still being in row echelon form, each column with an element in pivot position (which, in this case, are all the columns) has only one nonzero element, which equals 1 and is in the pivot position of a row.

This system of equations has a unique solution — x, y, and z can take on only *one* value in order for each equation to be true.

Example 1.5 (Equations with an infinite number of solutions):

$$\begin{bmatrix} x & + & y & + & 2z & = & 2 \\ & & y & + & z & = & 0 \\ 2x & + & y & + & 3z & = & 4 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 3 & 4 \end{array}\right]$$

To eliminate x from the third equation, subtract 2 times row 1 from row 3.

$$\begin{bmatrix} x & + & y & + & 2z & = & 2 \\ & & y & + & z & = & 0 \\ & - & y & - & z & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \end{array}\right]$$

To eliminate y from the third equation, add row 2 to row 3.

$$\begin{bmatrix} x & + & y & + & 2z & = & 2 \\ & & y & + & z & = & 0 \\ & & & & 0 & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

At this point, the third equation no longer contains z so we cannot “eliminate” it. We can, however, proceed by eliminating y from the first equation. To do this, subtract row 2 from row 1.

$$\begin{bmatrix} x & & & + & z & = & 2 \\ & & y & + & z & = & 0 \\ & & & & 0 & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We are now in reduced row echelon form, with basic variables x and y, and z a free variable. For any choice of z, setting $x = 2 - z$ and $y = -z$ yields a solution to this system of equations. This explains the *free variable* terminology; the values of free variables can be chosen arbitrarily, and then the resulting basic variables can be (uniquely) solved for to yield a solution to the system of equations.

This is the best we can do. Notice that the third equation is redundant (it is simply $0 = 0$), so we are left with two equations but three unknown variables. One solution is $x = 1, z = 1, y = -1$. Another possible solution would be $x = 2, z = 0, y = 0$. In fact, this system of equations has an infinite number of solutions — we could choose any value for z, set y to be $-z$ and x to be $2 - z$ and the two equations would still be true.

Despite the lack of a solution, notice that we have still placed this matrix into row echelon form, since the first two nonzero rows are both above the zero row, and the leading coefficient of the second row is to the right of the leading coefficient of the first row. Indeed, it is in fact even in *reduced* row echelon form, since the two elements in pivot position are both 1 and the first two columns, containing these two elements, contain no other elements.

Notice that the third column has two nonzero elements. Doesn't this violate the requirements of reduced row echelon form? No, since the third column does not contain either of the two entries in pivot position, so it is permitted to have more than one nonzero elements.

A key takeaway from this example is that placing a matrix in reduced row echelon form does not imply that it has a unique solution! However, it makes finding the *set* of possible solutions a lot easier, as will be discussed in future notes.

In a later note, we will discuss this situation in more detail.

Example 1.6 (Equations with no solution):

$$\begin{bmatrix} x & + & 4y & + & 2z & = & 2 \\ x & + & 2y & + & 8z & = & 0 \\ x & + & 3y & + & 5z & = & 3 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 1 & 2 & 8 & 0 \\ 1 & 3 & 5 & 3 \end{array}\right]$$

To eliminate x from all but the first equation, subtract row 1 from row 2 and row 3.

$$\begin{bmatrix} x & + & 4y & + & 2z & = & 2 \\ & - & 2y & + & 6z & = & -2 \\ & - & y & + & 3z & = & 1 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 0 & -2 & 6 & -2 \\ 0 & -1 & 3 & 1 \end{array}\right]$$

To make 1 the leading coefficient in row 2, divide row 2 by -2.

$$\begin{bmatrix} x & + & 4y & + & 2z & = & 2 \\ & & y & - & 3z & = & 1 \\ & - & y & + & 3z & = & 1 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & -1 & 3 & 1 \end{array}\right]$$

To eliminate y from the final equation, add row 2 to row 3.

$$\begin{bmatrix} x & + & 4y & + & 2z & = & 2 \\ & & y & - & 3z & = & 1 \\ & & & & 0 & = & 2 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

Now the third equation gives a contradiction, $0 = 2$. No choice of x, y, and z will change the rules of mathematics such that $0 = 2$, so there is no solution to this system of equations. If these were experimentally measured results, this contradiction might indicate that our modeling assumptions are incorrect or that there is noise in our measurements. In fact, this situation comes up frequently in real experiments, and later in this course we'll investigate techniques for dealing with noisy measurements.

Since we ran into a contradiction, we stopped Gaussian elimination early, so the matrix is not yet in reduced row echelon form. We can see this by inspecting the second column, which contains the element in the pivot position of the second row, but also contains a nonzero element in the first row! Still, we proceeded far enough in Gaussian elimination to put the matrix in row echelon form - can you see why?

Example 1.7 (Canceling intermediate variables):

$$\begin{bmatrix} x & + & y & + & 3z & = & 2 \\ 2x & + & 2y & + & 7z & = & 6 \\ -x & - & y & - & 2z & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 2 & 2 & 7 & 6 \\ -1 & -1 & -2 & 0 \end{array}\right]$$

To eliminate x from all but the first equation, subtract 2 times row 1 from row 2 and add row 1 to row 3.

$$\begin{bmatrix} x & + & y & + & 3z & = & 2 \\ & & & & z & = & 2 \\ & & & & z & = & 2 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Canceling x from rows 2 and 3 has also canceled y, so we eliminate the next variable, z. To do this, we can subtract row 2 from row 3, but this gives a zero row because the rows are identical:

$$\begin{bmatrix} x & + & y & + & 3z & = & 2 \\ & & & & z & = & 2 \\ & & & & 0 & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Because we now have fewer non-zero rows than variables, this system of equations has an infinite number of solutions, but we can still subtract 3 times row 2 from row 1 to eliminate z from the first equation.

$$\begin{bmatrix} x & + & y & & = & -4 \\ & & & z & = & 2 \\ & & & 0 & = & 0 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 0 & -4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Here, x and z are basic variables (since each of the corresponding columns contains a leading entry), and y is a free variable (since the corresponding column does not contain a leading entry of any row). For any choice of y, setting $x = -(4+y)$ and $z = 2$ will yield a solution.

While we can solve explicitly for z, there are an infinite number of possible values for x and y: for any choice of x, setting y to be $-(4+x)$ will provide a valid solution.

We see that this matrix is in fact in reduced row echelon form. This is to be expected, since we are at the stopping point for Gaussian elimination.

1.2.4.2 Algorithm Stopping Point

Based on the previous examples, we have seen that running Gaussian elimination does *not* guarantee that we will be able to find a solution to the system of equations. However, running the algorithm will tell us whether or not there is **one**, **zero**, or **infinitely many** solutions.

If a single solution exists, we will have an explicit equation for each variable. From the augmented matrix perspective, this means that the portion of the matrix corresponding to the coefficient weights will have 1's on the diagonal and 0's everywhere else, as in this example for a system of three equations with three unknowns (the first three columns are the coefficient weights):

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

If we think of Gaussian elimination as a way to rewrite our system of m equations with n variables as a set of explicit equations for each variable, intuitively there must be at least one equation for each variable ($m \geq n$) for a solution to exist. What happens if $m > n$? If the system of equations is consistent, the extra rows of the final augmented matrix should be all zeros — running Gaussian elimination will set the variable coefficients in these rows to zero, so the corresponding result entry should also be zero if a solution exists.

Now we can generalize this strategy to an arbitrary number of equations.

1. For a system of m equations and n variables ($m \geq n$), the first n rows of the augmented matrix have a triangular structure — specifically, the leftmost nonzero entry in row i is a 1 and appears in column i for $i = 1$ to n. If $m > n$, exactly $(m-n)$ rows are all-zero, and all correspond to the equation $0 = 0$. With 4 equations and 3 unknowns, this could be an augmented matrix such as

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

   This means that the system of equations has a **unique solution**. Notice that there may still be zero rows after the nonzero rows, but so long as the zero rows have a zero constant entry, they are consistent and so do not pose an issue. We can solve for one unknown by scaling the final row appropriately and eliminating it from every other equation. Repeat this until every equation has one unknown left and the system of equations is solved.

2. There are effectively fewer non-zero rows in the augmented matrix than there are variables, and any rows with all-zero variable coefficients also have a zero result, corresponding to the equation $0 = 0$. This could be an augmented matrix such as

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \text{ or } \left[\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Notice that this situation can occur even if there are no zero rows. If it occurs, there are fewer equations than unknowns and the system of linear equations is underdetermined. There are an **infinite number of solutions**.

3. There is a row in the augmented matrix with all-zero variable coefficients but a nonzero result, corresponding to the equation $0 = a$ where $a \neq 0$. This could be an augmented matrix such as

$$\left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right] \text{ or } \left[\begin{array}{cc|c} 1 & 4 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right] \text{ or } \left[\begin{array}{ccc|c} 1 & 4 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

This means that the system of linear equations is inconsistent and there are **no solutions**. Be aware that this scenario can occur *regardless of the shape of the matrix*, and automatically means that no solutions exist, since there exists an inconsistency in the given system of equations.

1.2.4.3 Formal Algorithm

So far, we have walked through in detail how to implement Gaussian elimination by hand. However, this quickly becomes impractical for large systems of linear equations — realistically, this algorithm will be implemented using a software program instead. Given the wide range of programming languages, each with unique syntax conventions, algorithms are often represented instead as **pseudocode**, a detailed English language description of an algorithm. While there is no single format for pseudocode, it should be general enough that it is free of syntactic dependencies but specific enough that it can be translated to actual code nearly automatically if you are familiar with the proper syntax.

Below is one possible pseudocode description of the Gaussian elimination algorithm. Note that we specify the input to the algorithm ("data") and the expected outcome ("result") at the top. We include explicit iterative statements ("for"), which execute the indented steps for each value of the parameters in the description, and conditional ("if," "if/else") statements, which (as the name implies) execute the indented code only if the conditions listed are met. Also note the "gets" symbol ($\longleftarrow$): $a \longleftarrow b$ means that a "gets" the value of b. In many programming languages, this is implemented as an equals sign, but the directed arrow notation makes it completely clear which variable takes on the value of the other variable.

Data: Augmented matrix $A \in \mathbb{R}^{m\times(n+1)}$, for a system of m equations with n variables
Result: Reduced form of augmented matrix

```
# Forward elimination procedure:
for each variable index i from 1 to n do
    if entry in row i, column i of A is 0 then
        if all entries in column i and row > i of A are 0 then
            proceed to next variable index;
        else
            find j, the smallest row index > i of A for which entry in column i ≠ 0 ;
            # The following rows implement the "swap" operation:
            old_row_j ⟵ row j of A;
            row j of A ⟵ row i of A;
            row i of A ⟵ old_row_j;
        end
    end
    divide row i of A by entry in row i, column i of A;
    for each row index k from i+1 to m do
        scaled_row_i ⟵ row i of A times entry in row k, column i of A;
        row k of A ⟵ row k of A − scaled_row_i;
    end
end
# Back substitution procedure:
for each variable index u from n−1 to 1 do
    if entry in row u, column u of A ≠ 0 then
        for each row v from u−1 to 1 do
            scaled_row_u ⟵ row u of A times entry in row v, column u of A;
            row v of A ⟵ row v of A − scaled_row_u;
        end
    end
end
```

Algorithm 1: The Gaussian elimination algorithm.

1.2.5 Tomography Revisited

How does what we have learned so far relate back to our tomography example, way back at the start of this note? We know that because our grocer's measurements come from a specific box with a particular assortment of milk, juice, and empty bottles, there must be one underlying solution, but insufficient measurements could give us a system of equations with an infinite number of solutions. So, how many measurements do we need?

Initially, we thought about shining a light vertically and horizontally through the box, giving six total equations because there are three rows and three columns per box. However, there are nine bottles to identify, and therefore nine variables, so we will need nine equations. Based on what you have learned about Gaussian elimination, you now understand that we need at least three more measurements — likely taken diagonally — in order to properly identify the bottles. In coming notes, we will discuss in further detail how you can tell whether or not the nine measurements you choose will allow you to find the solution.

1.3 Practice Problems

1.3.1 Mechanical Practice

Mechanical practice problems are also available in an interactive form on the course website, along with their solutions.

1. True or False: A system of 3 equations with 2 free variables has solutions along a line.

2. How many solutions does the system of equations have if, after performing Gaussian elimination, the row reduced form of the augmented matrix is $\left[\begin{array}{ccc|c} 3 & -1 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right]$?

   (a) One solution

   (b) Infinite solutions

   (c) No solutions

3. How many solutions does the system of equations have if, after performing Gaussian elimination, the row reduced form of the augmented matrix is $\left[\begin{array}{ccc|c} 3 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$?

   (a) One solution

   (b) Infinite solutions

   (c) No solutions

4. True or False: A system of equations with more equations than unknowns will always have either infinite solutions or no solutions.

5. Perform Gaussian elimination on the following set of equations to find x, y, and z. Remember to convert it to matrix form!

$$\begin{cases} 10x - 6y + 2z = 2 \\ 3x + 2y = 10 \\ -5x + 3y - z = 1 \end{cases}$$

   (a) $x = 2, y = 7, z = 4$

   (b) $x = 0, y = 3, z = 16$

   (c) Infinite solutions

   (d) No solutions

6. Solve this system of equations: $\begin{cases} 2x + y + 3z = 1 \\ x - y + 4z = 2 \\ x + 8y + z = 1 \end{cases}$

7. Solve this system of equations: $\begin{cases} 2x - 16y + 4z = -8 \\ x + 12y + 4z = 6 \\ x + 8y - 2z = 4 \end{cases}$

1.3.2 Homework / Exam Practice

1. Pizza and Pirates! (18 points)

You are stuck on a deserted island and you need to find food everyday! From science class, you know that each day you need to eat:

Table 1: Daily doses

Food [grams]	Daily dose
Fat	$4g$
Carbs	$2g$
Protein	$14g$
Vitamins	$6g$

Thankfully, you find a pirate camp on the the island and they have 4 kinds of food; eggs, pineapple pizza, bananas, and carrots. Once again, you thank your science teacher, and remember that you know the composition of each of these foods:

Table 2: Food composition

Food [grams]	1 egg	1 slice of pineapple pizza	1 banana	1 carrot
Fat	$1g$	$2g$	$0g$	$0g$
Carbs	$0g$	$2g$	$1g$	$0g$
Protein	$3g$	$3g$	$1g$	$0g$
Vitamins	$1g$	$0g$	$1g$	$1g$

In order to get enough food, you decide to steal some from the pirates. But, since it is so dangerous to steal food, you want to take *exactly* what you need, no more no less. Each day, you must decide how how much food to steal; number of eggs, x_e, number of pineapple pizza slices, x_p, number of bananas, x_b, and number of carrots, x_c.

1. (2 points) How many unknowns are there in this problem?

2. (6 points) Using Tables 1 and 2, write the equation for your daily dose of food groups in the form $\mathbf{A}\vec{x} = \vec{y}$ where $\vec{x} = [x_e, x_p, x_b, x_c]^T$. Clearly define $\mathbf{A}$ and $\vec{y}$ in your solution.

3. (10 points) Now let $\mathbf{A}$ and $\vec{y}$ be:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 4 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and } \vec{y} = \begin{bmatrix} 4 \\ 2 \\ 14 \\ 6 \end{bmatrix}, \tag{5}$$

where $\vec{y}$ is the daily dose of each food group needed.

Using the values from (5), find the solution or the set of solutions for how much of each type of food you need to steal everyday, i.e. solve for $\vec{x}$ in $\mathbf{A}\vec{x} = \vec{y}$.