MA1042 Module 1 - Solving Linear Systems

One of the most common applications of mathematics is solving several equations for several variables. In many cases, these equations are *linear*, where each variable is multiplied by some number (we will be more precise later). This type of problem arises in a wide variety of fields, including physics, engineering, chemistry, economics, logistics, biology, cryptography, etc. (The text shows many interesting applications.) However, due to the compressed nature of this course, we cannot take the time to explore particular applications. Instead, we focus on the mathematical techniques for solving these and related problems.

Our basic tool for solving such a problem is just a formalized version of how you would normally solve it by hand. In overview, the process has two main stages:

1. the **forward, elimination** stage, where at each step we use one equation to eliminate one variable from the remaining equations, to simplify them
2. the **backward, substitution** stage, where at each step we solve the simplest resulting equation for one variable and substitute that back into the other equations, again simplifying them

The end result gives the solution.

This basic method is the *single most important tool* you will learn in this course. We will apply it over and over, to a wide variety of problems, so it pays to master it right away. In this module, we develop a formal, structured approach, that works with any number of equations in any number of variables. We will find that the "bookkeeping" can be simplified by representing all the equations as a single *matrix*, a rectangular array of numbers.

Mod1.1 Systems of Linear Equations

When you think of two or three variables, you might think of x and y, or x, y and z, as your preferred notation. This gets less convenient as you consider more variables. Since we need to handle any number of variables, we will use subscript notation, so for five variables we might call them x_1, x_2, x_3, x_4, and x_5. For some generic number n of variables, we would abbreviate this list as $x_1, x_2, \ldots, x_n$.

A *linear combination* of variables is where we multiply each variable by a number and add those products, for example $3x_1 - 4x_2 + x_4 + 47x_5$ (where x_2 was multiplied by -4, and x_3 by 0). A *linear equation* is one that can be put in the form of a linear combination of the variables on the left and a number on the right, for example $3x_1 - 4x_2 + x_4 + 47x_5 = 493$.

The general form of a linear equation in the n variables $x_1, x_2, \ldots, x_n$, where the numbers multiplying each are $a_1, a_2, \ldots, a_n$ and the other number is b, is

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

If an equation can be put in this form, it is linear; if not, we call it nonlinear. For example, $2x_1 + x_2x_3 = 4$ cannot be put in that form due to the product x_2x_3; it is nonlinear. Similarly $x_1^2 - 2x_2 = 0$ is nonlinear.

A collection of one or more linear equations in the same variables is called a *system of linear equations*, or a *linear system* for short. A *solution* is an ordered list of numbers $(s_1, s_2, \ldots, s_n)$ that, when substituted for the corresponding variables, make all the equations true simultaneously. We will see that a linear system might have only one solution, or it might have more than one solution (infinitely many, if more than one). If a linear system has no solutions at all, we call it *inconsistent*; a system with one or more solutions is *consistent*.

If two different systems in the same variables both have the same set of solutions, we call those systems *equivalent*. Our goal in solving a linear system is to find an equivalent system that has the simplest possible form. As an example, the system

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ x_1 & + & x_2 & = & 5 \end{array}$$

is equivalent to the (simplest) system

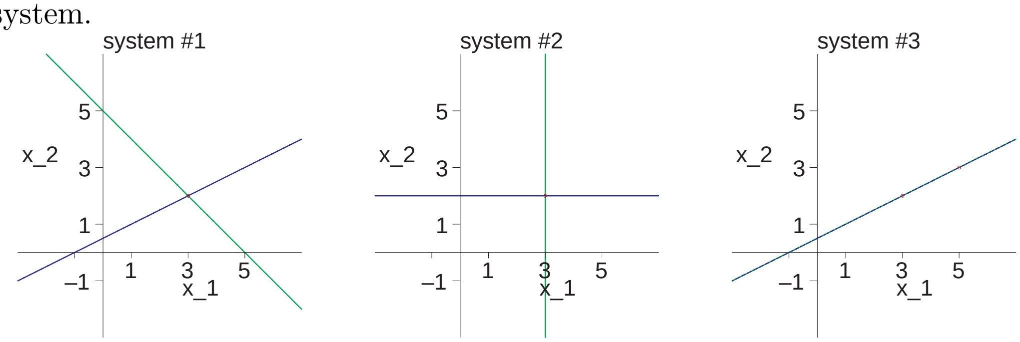
$$\begin{array}{rcrcr} x_1 & & & = & 3 \\ & & x_2 & = & 2 \end{array}$$

But note that neither of those is equivalent to the system

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ -2x_1 & + & 4x_2 & = & 2 \end{array}$$

because, even though the solution $(3, 2)$ satisfies this last system, so does $(5, 3)$, so the *solution set* is different. The first two systems only have one solution, $(3, 2)$; the last has many.

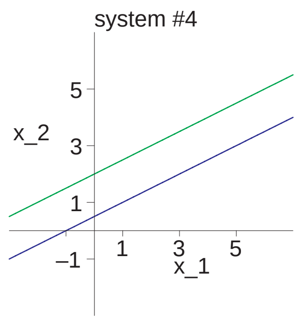
We can look at these systems graphically. Each linear equation in two variables represents a line in the plane; each point on the line satisfies the equation. (Typically, to plot the line, you find the intercepts, where the line crosses the axes, where one of the variables is zero.) A system of two such equations represents two lines; the solution is where they intersect (the point or points that satisfy both equations). The above three systems are plotted below. In the third system, both equations describe the same line; every point on the line is a solution to the system.



What happens if we change one number on the right of the last system, to get

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ -2x_1 & + & 4x_2 & = & 8 \end{array}$$

The resulting system is inconsistent; there are no solutions. So this system is not equivalent to any of the above ones. Graphically, the two lines are parallel and never intersect, as shown below:



As we will see, the above possibilities are the rule; regardless of the number of variables, every linear system has either:

1. no solutions (inconsistent), or
2. only one solution (unique), or
3. an infinite number of solutions.

Mod1.2 Solving Systems of Linear Equations

You may think that the goal of solving a linear system is to find what value each variable must have. That's correct for a system with a unique solution. But a better way to look at it, which works for all three cases, is: the goal of solving a linear system is to find the simplest possible equivalent system (like system 2 above).

We will use three elementary operations to simplify a system. The most useful one is adding a multiple of one equation to another equation; the text calls this *replacement* because we replace an equation by the new combination. Also, we can multiply an equation by a nonzero number; this is called *scaling* the equation. Lastly, we can *interchange* (swap) two equations, so we have the same equations in a new order. The text proves that these operations yield equivalent systems, which should be no surprise.

Example 1

Here's an example system:

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\ -x_1 & - & x_2 & + & x_3 & = & 3 \\ 2\,x_1 & + & 3\,x_2 & - & 3\,x_3 & = & -7 \end{array}$$

Let's use the first equation to eliminate x_1 from the other two, using the replacement operation. (This is basically equivalent to solving the first equation for x_1 and plugging that into the other two.) First, add row 1 to row 2 to cancel the x_1's:

$$\begin{array}{rcrcrcrcrc}
 & (& x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 &) \\
+ & (& -x_1 & - & x_2 & + & x_3 & = & 3 &) \\ \hline
= & (& 0 & + & x_2 & - & 3\,x_3 & = & -5 &)
\end{array}$$

This replaces equation 2, giving the equivalent system

$$\begin{array}{rcrcrcr}
x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\
 & & x_2 & - & 3\,x_3 & = & -5 \\
2\,x_1 & + & 3\,x_2 & - & 3\,x_3 & = & -7
\end{array}$$

Similarly, to cancel the $2x_1$ in the third equation, we can subtract twice the first:

$$\begin{array}{rcrcrcrcrc}
(-2)\times & (& x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 &) \\
+ & (& 2\,x_1 & + & 3\,x_2 & - & 3\,x_3 & = & -7 &) \\ \hline
= & (& 0 & + & -x_2 & + & 5\,x_3 & = & 9 &)
\end{array}$$

which replaces the third equation, giving

$$\begin{array}{rcrcrcr}
x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\
 & & x_2 & - & 3\,x_3 & = & -5 \\
 & & -x_2 & + & 5\,x_3 & = & 9
\end{array}$$

Now the last two equations only contain two variables, much simpler. Working with them, we do the same sort of thing we just did: we use the second equation to eliminate x_2 from the last, by adding it:

$$\begin{array}{rcrcrcrc}
 & (& x_2 & - & 3\,x_3 & = & -5 &) \\
+ & (& -x_2 & + & 5\,x_3 & = & 9 &) \\ \hline
= & (& 0 & + & 2\,x_3 & = & 4 &)
\end{array}$$

to replace the third row:

$$\begin{array}{rcrcrcr}
x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\
 & & x_2 & - & 3\,x_3 & = & -5 \\
 & & & & 2\,x_3 & = & 4
\end{array}$$

This completes the *forward elimination* stage. We have eliminated as many variables as possible from the last equation. We can see at a glance that we will be able to solve the last equation for x_3, substitute that back into the other two, then solve the middle equation for x_2 and substitute that into the first, which we can then solve for x_1. (OK, maybe a long glance.) Let's do it!

We scale equation 3 by $\frac{1}{2}$ (divide out the 2) to solve for x_3, giving

$$\begin{array}{rcrcrcr}
x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\
 & & x_2 & - & 3\,x_3 & = & -5 \\
 & & & & x_3 & = & 2
\end{array}$$

Now we substitute it back to cancel x_3 out of the other equations, using the replacement operation. So we add 3 times equation 3 to equation 2:

$$\begin{array}{rcrcrcrc}
3\times & (& & & x_3 & = & 2 &) \\
+ & (& x_2 & - & 3\,x_3 & = & -5 &) \\ \hline
= & (& x_2 & + & 0 & = & 1 &)
\end{array}$$

and add 4 times equation 3 to equation 1:

$$\begin{array}{rcrcrcrcrl} 4\times & (& & & & & x_3 & = & 2 &) \\ + & (& x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 &) \\ \hline = & (& x_1 & + & 2\,x_2 & + & 0 & = & 0 &) \end{array}$$

giving the simpler system

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & & & = & 0 \\ & & x_2 & & & = & 1 \\ & & & & x_3 & = & 2 \end{array}$$

Finally, we substitute x_2 from equation 2 back into the first, by

$$\begin{array}{rcrcrcrl} -2\times & (& & & x_2 & = & 1 &) \\ + & (& x_1 & + & 2\,x_2 & = & 0 &) \\ \hline = & (& x_1 & + & 0 & = & -2 &) \end{array}$$

giving the *simplest* system

$$\begin{array}{rrrcr} x_1 & & & = & -2 \\ & x_2 & & = & 1 \\ & & x_3 & = & 2 \end{array}$$

This completes the *backward substitution* stage, and there is the solution: $(-2, 1, 2)$.

Mod1.3 Using Matrices to Streamline the Process

We went through the last example in full detail, to make sure all the steps are clear and the reasons make sense. But now that we understand the process, we will streamline the steps. Notice that for each step in the example, we lined up all the x_1 terms in one column, the x_2's in another ...which helped in the bookkeeping. So instead of writing x_1, x_2 and x_3 every time, let's just write each *coefficient* (the number multiplying the variable) in the appropriate column. This saves a lot of writing; we just have to remember that the first column corresponds to x_1, etc.

So we write the coefficients as a rectangular array of numbers, called a *matrix* (plural: matrices). It is standard to put big square brackets around the block of numbers, to signify that this is a matrix. For our original system

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & - & 4\,x_3 & = & -8 \\ -x_1 & - & x_2 & + & x_3 & = & 3 \\ 2\,x_1 & + & 3\,x_2 & - & 3\,x_3 & = & -7 \end{array}$$

the *coefficient matrix* is

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

Each *column* gives the coefficients of a particular *variable*; each *row* gives the coefficients in a particular *equation.*

To keep track of the numbers on the right-hand side of each equation, we add another column, and call the result the *augmented matrix* for the system:

$$\left[\begin{array}{rrr|r} 1 & 2 & -4 & -8 \\ -1 & -1 & 1 & 3 \\ 2 & 3 & -3 & -7 \end{array}\right]$$

Here, I put a vertical bar before the augmented column, which I will also call the *data column*. This vertical bar reminds us where the equal sign goes in the equations; the bar separates the left side from the right side of each equation. The text does not use this notation, but I find it helpful at the beginning and at the end of the solution process, where we convert between a system of equations and an augmented matrix.

We solve a system in matrix form using the same steps as with equations, but now each equation corresponds to a *row*. So we call the three basic operations we applied to equations the *elementary row operations*:

1. *replacement*: add a multiple of one row to another row
2. *scaling*: multiply a row by a nonzero number
3. *interchange*: swap two rows

Example 1, revisited

Let's solve that same system again, using matrices. Our starting system:

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

For the forward elimination stage, we start with the first column, and use the top 1 to cancel out the -1 and 2 below it (eliminate x_1 from the other two equations). So we add row 1 to row 2:

$$\begin{array}{cc} & [\ 1 \quad 2 \quad -4 \quad -8\ ] \\ + & [\ -1 \quad -1 \quad 1 \quad 3\ ] \\ \hline = & [\ 0 \quad 1 \quad -3 \quad -5\ ] \end{array}$$

which replaces row 2, and we subtract twice row 1 from row 3:

$$\begin{array}{rc} (-2)\times & [\ 1 \quad 2 \quad -4 \quad -8\ ] \\ + & [\ 2 \quad 3 \quad -3 \quad -7\ ] \\ \hline = & [\ 0 \quad -1 \quad 5 \quad 9\ ] \end{array}$$

which replaces row 3. So now our new matrix, which is called *row equivalent* to the first matrix, since it represents an equivalent system, looks like:

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

Next, ignoring the first row, we work on the second column, and use the 1 in the second row to cancel out the -1 below it (eliminate x_2 from the last equation), that is, add row 2 to row 3:

$$\begin{array}{rl} & [\ 0 \quad 1 \quad -3 \quad -5\ ] \\ + & [\ 0 \quad -1 \quad 5 \quad 9\ ] \\ \hline = & [\ 0 \quad 0 \quad 2 \quad 4\ ] \end{array}$$

to replace row 3, giving

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

With some practice, you can probably do these row replacement operations in your head. We would summarize the steps so far by

$$\begin{bmatrix} 1 & 2 & -4 & -8 \\ -1 & -1 & 1 & 3 \\ 2 & 3 & -3 & -7 \end{bmatrix} \begin{matrix} \\ + \text{row } 1 \\ -2 \times \text{row } 1 \end{matrix} \sim \begin{bmatrix} 1 & 2 & -4 & -8 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 5 & 9 \end{bmatrix} \begin{matrix} \\ \\ + \text{row } 2 \end{matrix} \sim \begin{bmatrix} 1 & 2 & -4 & -8 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

where the symbol $\sim$ means "is row equivalent to" (which is *not* the same as $=$).

At this point, we have finished the forward elimination stage, and we say the matrix is in *row echelon form*. Now, proceeding with the backward substitution phase, we start at the bottom row, scale the leading 2 (first nonzero in the row) into 1 (solve for x_3), and use that to cancel the -4 and -3 above (substitute x_3 back into the first two equations):

$$\begin{bmatrix} 1 & 2 & -4 & -8 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 2 & 4 \end{bmatrix} \begin{matrix} \\ \\ \times\frac{1}{2} \end{matrix} \sim \begin{bmatrix} 1 & 2 & -4 & -8 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 1 & 2 \end{bmatrix} \begin{matrix} +4 \times \text{row } 3 \\ +3 \times \text{row } 3 \\ \\ \end{matrix} \sim \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

And finally, we use the leading 1 in the second row to cancel out the 2 above (substitute x_2 back into the first equation):

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \begin{matrix} -2 \times \text{row } 2 \\ \\ \\ \end{matrix} \sim \left[\begin{array}{ccc|r} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

This completes the back substitution phase, and this matrix is the simplest row equivalent form, called *reduced* row echelon form. Also, I put the vertical bar back in before the data column, to remind us it represents the (simplest equivalent) system:

$$\begin{array}{ccccr} x_1 & & & = & -2 \\ & x_2 & & = & 1 \\ & & x_3 & = & 2 \end{array}$$

Using matrices did not give us any new information, but it did eliminate a lot of writing, especially if you use the summary-type notation and do the row operations in your head. And it gave us a very structured approach, which works for any number of equations in any number of variables.

On the next page, we give a summary of the solution method, in full generality. You may want to print that page out for easy reference. It introduces some terminology and some situations we have not seen yet, but will in the subsequent examples. If it seems confusing at first, go on to the examples, to see how it works in practice.

Mod1.4 Method Summary

Here is a summary of the standard method for solving a linear system.

1. Convert the system to an augmented matrix.
2. Start with the first (leftmost) column that is not all zeros. Choose a nonzero entry in this column, and move it to the top (if it's not already there) by a row swap. In this position, the entry is called a *pivot*, and its position is called a *pivot position*.
3. Use the pivot, with row replacement operations, to zero out all the entries below it in that column.
4. Now work on the *submatrix* consisting of all rows *below* the last pivot row; go back to step 2 for the submatrix. Repeat until you run out of pivots.

Note: at this point, you have completed the forward elimination stage (called *Gaussian elimination*), and the matrix is in **row echelon form** (or just *echelon form*):

  (a) the leading entry (first nonzero) in any row is in a column to the left of the leading entry in the row below;

  (b) any rows of all zeros are below all nonzero rows.

In row echelon form, the leading entry of each row is called a *pivot*.

5. If there is a pivot in the data column, **stop**: the system is **inconsistent**. Otherwise, continue.
6. Start with the lowest pivot. Scale the row to make the pivot = 1.
7. Use that pivot, with row replacement operations, to cancel out all the entries above it in that column.
8. Now work on the *submatrix* consisting of all rows *above* the last pivot row; go back to step 6 for the submatrix. Repeat until you run out of pivots.

Note: at this point, you have completed the backward substitution stage (the whole process up to here is called *Gauss-Jordan elimination*), and the matrix is in **reduced row echelon form** (or just *reduced echelon form*), which is row echelon form with additional structure:

  (c) the leading entry in each row is 1, and every other entry in that column is zero.

9. Convert this augmented matrix back to a system of equations.
10. Solve each equation for the first variable in it; these variables correspond to pivot columns, and are called *basic variables*.
11. Any variables that are not basic variables (i.e., whose column has no pivot) are called *free variables*; for each free variable, write down that it is free.
12. That gives the *general solution*. If there are no free variables, the solution is **unique**. If there are any free variables, then each of them can take on any value (they act as parameters) and the number of solutions is **infinite**.

Sometimes the arithmetic can be simplified using extra row operations, but always keep the goals of row echelon form and reduced row echelon form in mind.

This is not the only possible method, but for the purposes of this course, it is the only method we will use. To solve any linear system, we first find an equivalent row echelon form. There are many different possible row echelon matrices for a given system, but they all have the same pivot positions, which tell us how many solutions there are: if there is a pivot in the data column, there are no solutions; otherwise, if there is a variable column with no pivots (a free variable), there are an infinity of solutions; otherwise, there is a unique solution. To get the solution(s), we find the equivalent *reduced* row echelon matrix; there is only one such matrix for a given system.

Mod1.5 More Examples

Example 2

Let's try another system of three equations in three variables (called a 3×3 system):

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & + & 3\,x_3 & = & -1 \\ -3\,x_1 & - & 6\,x_2 & - & 11\,x_3 & = & 4 \\ 2\,x_1 & + & 4\,x_2 & + & 10\,x_3 & = & -3 \end{array}$$

which corresponds to the augmented matrix:

$$\left[\begin{array}{rrr|r} 1 & 2 & 3 & -1 \\ -3 & -6 & -11 & 4 \\ 2 & 4 & 10 & -3 \end{array}\right]$$

Now zero out column 1 below the pivot in row 1:

$$\left[\begin{array}{rrrr} 1 & 2 & 3 & -1 \\ -3 & -6 & -11 & 4 \\ 2 & 4 & 10 & -3 \end{array}\right] \begin{array}{l} \\ +3 \times \text{row } 1 \\ -2 \times \text{row } 1 \end{array} \sim \left[\begin{array}{rrrr} 1 & 2 & 3 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 4 & -1 \end{array}\right]$$

Ignoring row 1, the first nonzero column in the submatrix is now column 3; use the pivot -2 to zero out the 4 below:

$$\left[\begin{array}{rrrr} 1 & 2 & 3 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 4 & -1 \end{array}\right] \begin{array}{l} \\ \\ +2 \times \text{row } 2 \end{array} \sim \left[\begin{array}{rrr|r} 1 & 2 & 3 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Now, ignoring rows 1 and 2, the submatrix is just row 3 so there's nothing to zero out below it. We're done with the forward stage, and the matrix is in echelon form.

Before continuing, we check where the pivots are: in columns 1, 3, and 4. But column 4 is the data column, and a pivot in the data column means the system is inconsistent. Why? Let's convert back to equations to see:

$$\begin{array}{rcrcrcr} x_1 & + & 2\,x_2 & + & 3\,x_3 & = & -1 \\ & & & & -2\,x_3 & = & 1 \\ & & & & 0 & = & 1 \end{array}$$

There are no values of x_1, x_2, and x_3 that make all three equations true simultaneously, because the last equation is always false! So the original system is also *inconsistent*, and we're done (skip the backward stage).

Example 3

Here's another 3×3 system:

$$\begin{array}{rcrcrcl} & & 4\,x_2 & - & 4\,x_3 & = & 0 \\ x_1 & - & 3\,x_2 & + & x_3 & = & 1 \\ 3\,x_1 & - & 4\,x_2 & - & 2\,x_3 & = & 3 \end{array} \quad \text{or} \quad \left[\begin{array}{rrr|r} 0 & 4 & -4 & 0 \\ 1 & -3 & 1 & 1 \\ 3 & -4 & -2 & 3 \end{array}\right]$$

Here, we can't use the current row 1 to zero out the rest of the first column, so let's swap row 2 into that position first (we could have used row 3, but putting a 1 in the pivot position is convenient for hand calculation):

$$\begin{bmatrix} 0 & 4 & -4 & 0 \\ 1 & -3 & 1 & 1 \\ 3 & -4 & -2 & 3 \end{bmatrix} \begin{array}{l} \text{swap w/2} \\ \text{swap w/1} \\ \end{array} \sim \begin{bmatrix} 1 & -3 & 1 & 1 \\ 0 & 4 & -4 & 0 \\ 3 & -4 & -2 & 3 \end{bmatrix} \begin{array}{l} \\ \\ -3 \times \text{row } 1 \end{array} \sim \begin{bmatrix} 1 & -3 & 1 & 1 \\ 0 & 4 & -4 & 0 \\ 0 & 5 & -5 & 0 \end{bmatrix}$$

Now we ignore row 1, and just consider the last two rows. And for convenience, scale the second row to get a leading 1 in the new pivot position, before zeroing out the 5 below it:

$$\begin{bmatrix} 1 & -3 & 1 & 1 \\ 0 & 4 & -4 & 0 \\ 0 & 5 & -5 & 0 \end{bmatrix} \begin{array}{l} \\ \times\frac{1}{4} \\ \end{array} \sim \begin{bmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 5 & -5 & 0 \end{bmatrix} \begin{array}{l} \\ \\ -5 \times \text{row } 2 \end{array} \sim \left[\begin{array}{rrr|r} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We've run out of pivots for the forward stage, and reached echelon form. We can see that columns 1 and 2 are pivot columns. There is no pivot in the data column, so the system is consistent. But there is also no pivot in column 3, so x_3 is a free variable, and we will get infinitely many solutions.

Now for the backward stage, our lowest pivot is in row 2, and it's already a 1, so we only need one more operation:

$$\begin{bmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{array}{l} +3 \times \text{row } 2 \\ \\ \end{array} \sim \left[\begin{array}{rrr|r} 1 & 0 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and have reached the reduced echelon form. Converting back:

$$\begin{array}{rrcrcl} x_1 & & - & 2\,x_3 & = & 1 \\ & x_2 & - & x_3 & = & 0 \\ & & & 0 & = & 0 \end{array}$$

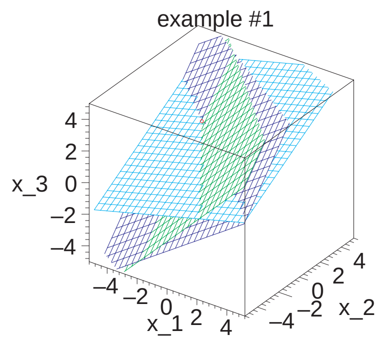
(the last equation is true but not informative) and solving gives the general solution:

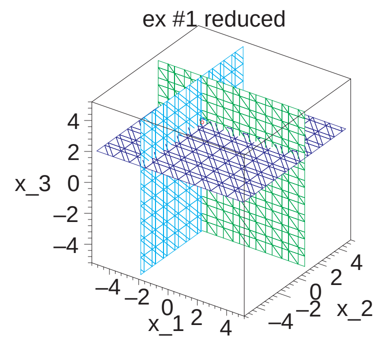
$$\begin{array}{lcl} x_1 & = & 1 + 2\,x_3 \\ x_2 & = & x_3 \\ x_3 & \text{is} & \text{free} \end{array}$$

The general solution shows the free variable x_3 and how the other (basic) variables depend on it, in all of the different possible solutions. If we choose $x_3 = 0$, then $x_1 = 1$ and $x_2 = 0$, so one solution is $(1, 0, 0)$. We get another if we choose $x_3 = 2$, then $x_1 = 5$ and $x_2 = 2$.

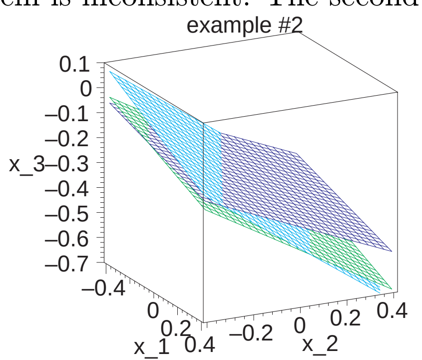
Every choice of the free variable gives a different solution to the original system, so there are an infinite number of solutions.

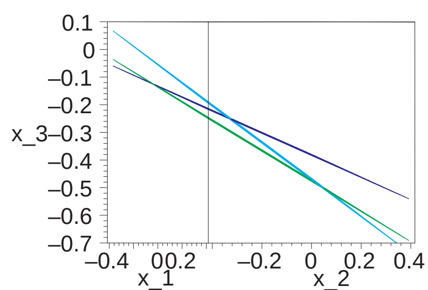
We can look at these three examples graphically. Each linear equation in three variables represents a plane in three-dimensional space; each point on the plane satisfies the equation. A system of three such equations represents three planes; the solution is where all three intersect (the point or points that satisfy all three equations). Below is a plot of Example 1; the three planes intersect in a single point. The second plot is the reduced echelon form of the system.



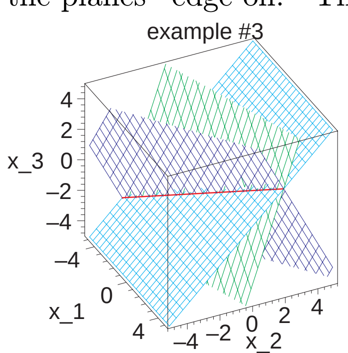


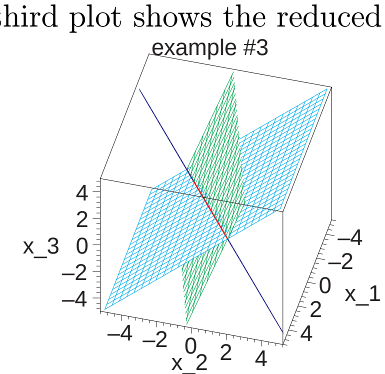
The plot of Example 2 is shown below, from two different viewpoints. Each pair of planes intersects in a line, but those three lines are parallel. The three planes never all intersect, and the system is inconsistent. The second view shows the planes "edge-on."

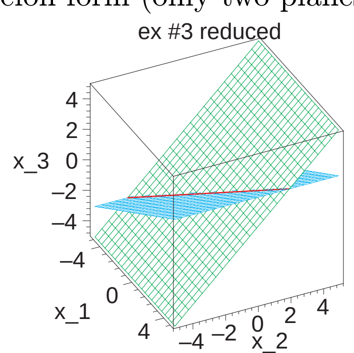




Example 3 is plotted below, again from two viewpoints. Here all three planes intersect in the same line; every point on the line is a solution to the system. The second view shows one of the planes "edge-on." The third plot shows the reduced echelon form (only two planes).







Example 4

This final example is a system of four equations in five variables (a 4×5 system):

$$\begin{array}{rcrcrcrcrcr} -2\,x_1 & + & x_2 & - & x_3 & + & x_4 & - & 3\,x_5 & = & -1 \\ 4\,x_1 & - & 2\,x_2 & + & 2\,x_3 & - & 5\,x_4 & & & = & -4 \\ 2\,x_1 & + & x_2 & + & 3\,x_3 & + & 7\,x_4 & - & x_5 & = & 15 \\ & & x_2 & + & x_3 & + & 3\,x_4 & - & 4\,x_5 & = & 5 \end{array}$$

with the augmented matrix:

$$\left[\begin{array}{rrrrr|r} -2 & 1 & -1 & 1 & -3 & -1 \\ 4 & -2 & 2 & -5 & 0 & -4 \\ 2 & 1 & 3 & 7 & -1 & 15 \\ 0 & 1 & 1 & 3 & -4 & 5 \end{array}\right]$$

First we zero out the first column below the pivot at the top:

$$\left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 4 & -2 & 2 & -5 & 0 & -4 \\ 2 & 1 & 3 & 7 & -1 & 15 \\ 0 & 1 & 1 & 3 & -4 & 5 \end{array}\right] \begin{array}{l} \\ +\,2 \times \text{row } 1 \\ +\,\text{row } 1 \\ \\ \end{array} \sim \left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 0 & 0 & -3 & -6 & -6 \\ 0 & 2 & 2 & 8 & -4 & 14 \\ 0 & 1 & 1 & 3 & -4 & 5 \end{array}\right]$$

Now, ignoring row 1, the first nonzero column is the second. Let's swap rows 2 and 4 to put a 1 in the pivot position, and also use that 1 to eliminate the 2 in that column:

$$\left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 0 & 0 & -3 & -6 & -6 \\ 0 & 2 & 2 & 8 & -4 & 14 \\ 0 & 1 & 1 & 3 & -4 & 5 \end{array}\right] \begin{array}{l} \\ \text{swap w/4} \\ -\,2 \times \text{row } 4 \\ \text{swap w/2} \end{array} \sim \left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 3 & -4 & 5 \\ 0 & 0 & 0 & 2 & 4 & 4 \\ 0 & 0 & 0 & -3 & -6 & -6 \end{array}\right]$$

Ignoring the first two rows, column 4 is the first nonzero column; use the pivot 2 to eliminate the -3 below:

$$\left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 3 & -4 & 5 \\ 0 & 0 & 0 & 2 & 4 & 4 \\ 0 & 0 & 0 & -3 & -6 & -6 \end{array}\right] \begin{array}{l} \\ \\ \\ +\,\frac{3}{2} \times \text{row } 4 \end{array} \sim \left[\begin{array}{rrrrr|r} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 3 & -4 & 5 \\ 0 & 0 & 0 & 2 & 4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We've reached echelon form, and see the pivot columns are 1, 2, and 4. The data column has no pivot, so the system is consistent, and we have free variables, so we will get an infinite number of solutions.

Now for the backward stage, starting with the lowest pivot (in row 3), we scale it to 1 and zero out the column above it:

$$\left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 3 & -4 & 5 \\ 0 & 0 & 0 & 2 & 4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \times \frac{1}{2} \\ \\ \end{array} \sim \left[\begin{array}{rrrrrr} -2 & 1 & -1 & 1 & -3 & -1 \\ 0 & 1 & 1 & 3 & -4 & 5 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} -\,\text{row } 3 \\ -\,3 \times \text{row } 3 \\ \\ \\ \end{array}$$

$$\sim \left[\begin{array}{rrrrrr} -2 & 1 & -1 & 0 & -5 & -3 \\ 0 & 1 & 1 & 0 & -10 & -1 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Working upward, the next pivot is in column 2; we zero out the 1 above it, then scale the pivot in row 1:

$$\begin{bmatrix} -2 & 1 & -1 & 0 & -5 & -3 \\ 0 & 1 & 1 & 0 & -10 & -1 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} - \text{row } 2 \quad \sim \begin{bmatrix} -2 & 0 & -2 & 0 & 5 & -2 \\ 0 & 1 & 1 & 0 & -10 & -1 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \tfrac{-1}{2}$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -5/2 & 1 \\ 0 & 1 & 1 & 0 & -10 & -1 \\ 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Converting back to equations:

$$\begin{array}{ccccccccc} x_1 & & + & x_3 & & - & 5/2\,x_5 & = & 1 \\ & x_2 & + & x_3 & & - & 10\,x_5 & = & -1 \\ & & & & x_4 & + & 2\,x_5 & = & 2 \\ & & & & & & 0 & = & 0 \end{array}$$

and solving gives the general solution

$$\begin{array}{lll} x_1 & = & 1 - x_3 + 5/2\,x_5 \\ x_2 & = & -1 - x_3 + 10\,x_5 \\ x_3 & \text{is} & \text{free} \\ x_4 & = & 2 - 2\,x_5 \\ x_5 & \text{is} & \text{free} \end{array}$$

The free variables x_3 and x_5 act as parameters, they can be anything, and determine the values of the other variables. So again, we have an *infinite* number of solutions. (But to visualize this we would need five dimensions for the five variables...)

Hopefully, the summary and examples have given you a firm handle on this method for solving linear systems. We will refer to this method as the *row reduction algorithm*. Mathematicians would call it *Gauss-Jordan elimination* (or *Gaussian elimination* if we stop at echelon form).