Systems of Linear Equations

You've probably found the solution of two linear equations in two variables in school, but let's review it briefly. Say you have two equations like

$$2x - y = 3, \qquad 3x + 2y = 4,$$

and suppose we want to find which values of x and y solve this system. We can rewrite the first equation as

$$y = 2x - 3,$$

and then plug this expression for y into the second equation to get

$$3x + 2(2x - 3) = 7x - 6 = 4 \Rightarrow x = \frac{10}{7}.$$

Now plug $x = 10/7$ into $y = 2x - 3$ to get $y = -1/7$, and we have our solution.

In fact, we can do the same sort of thing if we have three equations in three unknowns, or 46 equations in 46 unknowns, but the notation starts to get awkward. Also, this process is difficult to program into a computer. So, maybe we'd like a better way to write down our system of linear equations.

Warning: The following will seem overly complicated at first; you may want to skim it and then get to the general rules for elementary row operations later on in these notes, and then come back and read it again. You can also make up your own examples to work through, as there's really nothing special about the example we do immediately below.

You may have also learned in school how to solve these equations by eliminating variables; basically you add together multiples of the original equation (in the case above, twice the first equation plus the second equation) to eliminate one of the variables. It turns out that the natural setting for eliminating variables like this is an equation with matrices. Let's see how this works with our original set of equations. We started with

$$2x - y = 3 \qquad 3x + 2y = 4.$$

If we write this as a column of two equations, we have

$$\begin{bmatrix} 2x - y \\ 3x + 2y \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix},$$

which we can write as the following matrix equation:

$$\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Now let's eliminate the varaible y from the second equation. We replace the second equation with twice the first equation plus the second equation, which transforms the left hand side of our matrix equation into

$$\begin{bmatrix} 2 & -1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

(You should check that the product above is correct.)

Ok, how did we get that matrix $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$? We can see this by remembering how matrix multiplication works. Recall that we multiply matrices by multiplying a row of the matrix on the left with a column of the matrix on the right. So, to effect the second row of our given matrix we have to concentrate on the second row. We want twice the first equation plus the second equation, which means we should put a 2 in the first column (so it multiplies the first row) and a 1 in the second column (so it multiplies the second row). As for the row 1   0 in the matrix, we want to leave the first equation alone, which is the same thing as replacing the first equation with itself. We do that by having a 1 in the first colum of the first row (to get 1 times the first row) and a 0 in the second column of the first row (to get 0 times the second row). As we said earlier, this looks overly-complicated, but it's a lot easier once we've practiced with some examples. See the rules for elementary row operations below, and work through the tutorials.

Now, we always have to do the same thing to the right hand side of an equation that we do to the left hand side (otherwise it wouldn't be an equation anymore). So we have

$$\begin{bmatrix} 2 & -1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 10 \end{bmatrix},$$

which is really the system of equations

$$2x - y = 3, \qquad 7x = 10.$$

Next we want to eliminate the 2 in the upper left corner of the first coefficient matrix, so we replace the first row (of both sides of the equation) with $-2/7$ times the second row plus the first row. This gives us

$$\begin{aligned} \begin{bmatrix} 0 & -1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 & -2/7 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 1 & -2/7 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 10 \end{bmatrix} = \begin{bmatrix} 1/7 \\ 10 \end{bmatrix}. \end{aligned}$$

Now, this last equation might look a little odd, because the first equation involves only y while the second equation involves only x, but we can fix that by swapping the two row to get

$$\begin{aligned} \begin{bmatrix} 7 & -0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/7 \\ 10 \end{bmatrix} = \begin{bmatrix} 10 \\ 1/7 \end{bmatrix}. \end{aligned}$$

This last matrix equation says

$$7x = 10, \qquad -y = \frac{1}{7}.$$

We can get rid of those inconvenient coefficients in front of the x and the y by multiplying the first row by $1/7$ and the second row by -1 to get

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1/7 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 1/7 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 10 \\ 1/7 \end{bmatrix} = \begin{bmatrix} 10/7 \\ -1/7 \end{bmatrix}, \end{aligned}$$

which is really the set of equations

$$x = 10/7, \qquad y = -1/7.$$

Whew! That seemes like a lot of work for such a simple system of two equations in two unknowns. Actually, it was a little more work than it should have been, but the real advantage in solving systems of linear equations in this way lies in

- it's a method you can easily program into a computer and
- it adapts very very well to systems of many equations in many unknowns.

Now let's list the general rules for elementary row operations. First suppose we have a system of m equations in n unknows. Call the unknowns $x_1, x_2, \dots, x_n$, and say the equations look like

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ \vdots &= \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m, \end{aligned}$$

which we can rewrite as the matrix equation

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

or

$$Ax = b,$$

where A is the $m \times n$ matrix with entries A_{ij}, x is the vector $(x_1, \dots, x_n)$, and b is the vector $(b_1, \dots, b_m)$. (You might want want to think about why the matrix equation is the same thing as the systems of equations.) There are two main operations we'll discuss. The first one is replacing an equation with the sum of two equations. Let's say we want to replace the third equation with the sum of the two times the first and three times the third equation. Then we multiply by an $m \times m$ matrix which is the identity, except in the third row. In the third row we have a 2 in the first column (to get the 2 times the first row) and a 3 in the third column (to get 3 times the third row. The result is that we multiply the matrix A by the $n \times n$ matrix

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 2 & 0 & 3 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}.$$

In general, if we want to repalce the ith row with c times the ith row plus d times the jth row, our matrix E will look like the identity except in the ith row. In the ith row, we will have c in the ith column, d in the jth column, and 0 in all the other columns. Here i and j can be any natural numbers from 1 to n, and c and d can be any real numbers. Then multiplying the matrix A by this choice of E (with E on the left) will have the desired effect.

The other row operation we'll want to do is swapping two rows. In this case we multiply A on the left with the following matrix F. If we want to swap the ith row with the jth row, our matrix F will look like the identity everywhere except the ith and jth rows. In the ith row, we have a 1 in the j column and a 0 in all the other columns, while in the jth row we have 1 in the ith column and 0 in all the other columns. For instance, if we want to swap the first and third rows our F matrix looks like

$$F = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, let's write out an algorithm to solve a system of linear equations by elementary row operations.

1. Start with the system of m equations in n unknowns

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ \vdots &= \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m, \end{aligned}$$

2. Convert this into the matrix equation $Ax = b$ where A is the $m \times n$ matrix with entries A_{ij}, x is the vector $(x_1, \ldots, x_n)$, and b is the vector $(b_1, \ldots, b_m)$.

3. Swap rows so that the top left entry A_{11} of A is nonzero. This is the same as multiplying A on the left by one of our F matrices.

4. Now use elementary row operations of the first type, multiplying on the left by E matrices, to cancel out the first column in the second, third, and so on rows, leaving the first column to be

$$\begin{bmatrix} A_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where $A_{11} \neq 0$.

5. Now swap rows (among the second through the mth row) to make $A_{22} \neq 0$, and repeat the process you've just done. In this way you can cancel out everything in the second column, in all the rows except the row, without effecting what you just did to the first column. This is because $A_{21} = 0$. Now our first two rows look like

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}.$$

6. Proceed to the third column, fourth column, and so on.

7. While you're doing this, remember that whatever you do to the left hand side of the equation (the A matrix) you also have to do to the right hand side of the equation (the b vector).

8. You stop this process when you hit either the rightmost column or the bottom row of the A matrix, whichever happens first. Now you have a matrix equation with looks like

$$\begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

   This should be easy to solve.

9. Now, if you have more unknowns than equations, that is $n > m$, you'll hit the bottom row first in the process above, and you'll have several columns of nonzero numbers to the right of your nice diagonal matrix. In this case, the system of equations is **underdetermined**, and you'll have some free variables to play with. We'll talk more about this case in a moment.

10. If you have more equations than unknowns, that is $m > n$, then you'll have several rows of zeroes at the bottom of your A matrix. This sort of system of equations is called **overdetermined**. Now you check to see if the last several b_j's corresponding to these rows of zeroes are also zero. If they all are, then you've found the solution to your system of equations. Otherwise, your system of equations has no solutions, and it is inconsistent. This happens ocassionally, and is not the end of the world.

11. Even in the case when you have the same number of equations as unknowns, that is $n = m$, it is possible to get a row of 0's in your A matrix. If this happens in the middle of your computations, you should stop and see if the corresponding b_j is also 0. If it is, then continue as before. If it's not, you can stop computing because you have an inconsistent system of equations. Again, this can happen, but its infrequent.

Let's look at one example of an underdetermined system, and one of an overdetermined system. First consider the system

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

This system has two equations and three unknowns, so it's underdetermined, and translated into a matrix equation it looks like

$$\begin{bmatrix} 2 & 1 & -3 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

We can replace the second equation with the first equation minus twice the second equation. This is the same as multiplying on the left by the 2 matrix $\begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}$, but we can do this operation mechanically, just adding the corresponding rows together without doing the matrix multiplication. After this operation, we get

$$\begin{bmatrix} 2 & 1 & -3 \\ 0 & 3 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}.$$

Next we can replace the first equation with three times the first equation minus the second equation. This is the same as multiplying on the left by the 2 matrix $\begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix}$, but again we don't really need to do the matrix multiplication. Now we have

$$\begin{bmatrix} 2 & 0 & -2 \\ 0 & 3 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ -2 \end{bmatrix}.$$

Now we translate this back into regular scalar equations:

$$2x_1 - 2x_3 = 14, \qquad 3x_2 - 7x_3 = -2,$$

or

$$x_1 = x_3 + 7, \qquad x_2 = \frac{7}{3}x_3 - \frac{2}{3}.$$

We see that we are free to choose x_3 to be whatever we want, and that every choice of x_3 will give us a solution to the system of equations. This is typically what happens in an underdetermined system of equations.

Let's also look at an example of an overdetermined system of linear equations. Consider the system

$$x_1 - x_2 = -1, \qquad x_1 + x_2 = 3, \qquad 3x_1 + 2x_2 = 1,$$

which is equaivalent to the matrix equation

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}.$$

It will be a little convenient to swap the first two rows, which is the same as multiplying on the left by the matrix

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

to get

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}.$$

Now we replace the second row with the first row minus the second row, which is the same as multiplying on the left by

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

to get

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

Next we replace the third row with the three time the first row minus the third row, which is the same as multiplying on the left by the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & -1 \end{bmatrix}$$

to get

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 8 \end{bmatrix}.$$

At this point we stop and look at the last two equations which say

$$2x_2 = 2, \qquad x_2 = 8,$$

which give different values for x_2. Thus, we cannot find a pair of numbers x_1, x_2 which simultaneously solves all three of the given equations, and so the system of equations has no solution. In fact, this is typically what happens when you have more equations than unknowns, *i.e.* an overdetermined system of equations.

We've seen that, in order to do row operations and solve systems of linear equations, we don't really need those odd E and F matrices; so why did we talk about them in the first place? They turn out to be very useful for constructing inverses of a matrix, as we will see in the next set of notes.