Linear Algebra I
MA1S1

Tristan McLoughlin

October 17, 2014

Anton & Rorres: Ch 1.3
Hefferon: Ch One, sec I.2

What is linear and not linear?

Here are some examples of equations that are plausibly interesting from some practical points of view

$$\begin{aligned} x^2 + 3x - 4 &= 0 \quad (i) \\ x^2 + 3x + 4 &= 0 \quad (ii) \\ x^2 &= 2 \quad (iii) \\ x - \sin x &= 1 \quad (iv) \end{aligned}$$

but **none of these are linear!**

What is linear and not linear?

Lets consider the first equation (i):

$$x^2 + 3x - 4 = 0$$

This equation is easy to solve by factorisation.

$$x^2 + 3x - 4 = (x-1)(x+4) = 0$$

so $x = 1$ or $x = -4$.

The second equation (ii):

$$x^2 + 3x + 4 = 0$$

is harder to solve— it can't be factored and if you use the quadratic formula to get the solutions, you get complex roots:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-3 \pm \sqrt{9-16}}{2} = \frac{-3 \pm \sqrt{-7}}{2} = \frac{-3 \pm \sqrt{7}\sqrt{-1}}{2}.$$

So the need for complex numbers emerged from certain quadratic equations.

The third equation (iii)

$$x^2 = 2$$

is not that complicated but it was upsetting for the Greek mathematicians of a few thousand years ago to realise that there are no solutions that can be expressed as $x = \frac{p}{q}$ with p and q whole numbers. (Such numbers are called *rational numbers* and $\sqrt{2}$ is an *irrational number* because it is not such a fraction.)

For the last equation (iv)

$$x - \sin x = 1$$

perhaps it is not so obvious that $\sin x$ is not linear, but it is not.

Linear equations with a single unknown

So what *are* linear equations?

They are very simple equations like

$$3x + 4 = 0.$$

There is not much to solving this type of equation (for the unknown x). For reasons we will see later, let's explain the simple steps involved in solving this example in a way that may seem unnecessarily long-winded.

Add -4 to both sides of the equation to get

$$3x = -4.$$

Then multiply both sides of this by $\frac{1}{3}$ (or divide both sides by 3 if you prefer to put it like that) to get

$$x = -\frac{4}{3}.$$

These simple steps transform the problem (of solving the equation) to a problem with all the same information.

If x solves

$$3x + 4 = 0$$

then it has to solve

$$3x = -4.$$

On the other hand, we can go back from $3x = -4$ to the original $3x + 4 = 0$ by adding $+4$ to both sides.

In this way we see that any x that solves $3x = -4$ must also solve $3x + 4 = 0$. So the step of adding -4 is reversible. Similarly the step of multiplying by $\frac{1}{3}$ is reversible by multiplying by 3.

Linear equations with two unknowns

Well, that little explanation seems hardly necessary, and how could we be having a course about such a simple thing?

We can start to make things a little more complicated if we introduce a linear equation in two unknowns

$$ax + by = c$$

where neither a or b are not zero. In the case where $c = 0$ the equation is called a homogeneous linear equation.

Consider the example:

$$5x - 2y = 1.$$

What are the solutions to this equation? We can't really "solve" it because one equation is not enough to find 2 unknowns. What we can do is solve for y in terms of x.

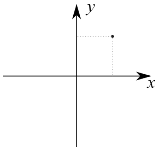
$$\begin{aligned} 5x - 2y &= 1 \\ -2y &= -5x + 1 && \text{(add } -5x \text{ to both sides)} \\ y &= \frac{5}{2}x - \frac{1}{2} && \text{(multiply both sides by } -\tfrac{1}{2}\text{)} \end{aligned}$$

When we look at it like this we can understand things in at least 2 ways.

(A) We are free to choose any value of x as long as we take $y = \frac{5}{2}x - \frac{1}{2}$. So we have infinitely many solutions for x and y.

(B) Another way is to think graphically. $y = \frac{5}{2}x - \frac{1}{2}$ is an equation of the type $ay + bx = c$ or the even more familiar form $y = mx + c$.
That is to say, the equation of a line in the x-y plane with slope $m = \frac{5}{2}$ that crosses the y-axis at the place where $y = c = -\frac{1}{2}$.

Thus we can think of solutions as points in a plane labeled by two coordinates (x, y).



In summary: *The solutions of a single linear equation in 2 unknowns can be visualised as the set of all the points on a straight line in the plane.*

Going back to the mechanism of solving, we could equally solve $5x - 2y = 1$ for x in terms of y. We'll write that out because it is actually the way we will do things later. (We will solve for the variable listed first in preference to the one listed later.)

$$
\begin{aligned}
5x - 2y &= 1 \\
x - \frac{2}{5}y &= \frac{1}{5} \text{ (multiply both sides by } \tfrac{1}{5}\text{)} \\
x &= \frac{1}{5} + \frac{2}{5}y \qquad \text{(add } \tfrac{2}{5}y \text{ to both sides)}
\end{aligned}
$$

We end up with y a *free variable* and once we choose any y we like to get a solution as long as we take $x = \frac{1}{5} + \frac{2}{5}y$. Also we get all the solutions this way — different choices of y give all the possible solutions. In this sense we have described *all* the solutions in a way that is *as uncomplicated* as we can manage.

Systems of linear equations

If we now move to systems of equations (also known as simultaneous equations) where we want to understand the (x, y) that solve all the given equations simultaneously, we can have examples like

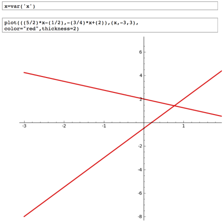
$$\left\{ \begin{array}{rcrcr} 5x & - & 2y & = & 1 \\ 3x & + & 4y & = & 8 \end{array} \right. \tag{1}$$

or

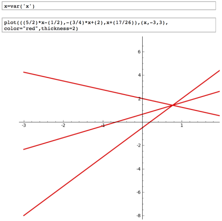
$$\left\{ \begin{array}{rcrcr} 5x & - & 2y & = & 1 \\ 3x & + & 4y & = & 8 \\ 26x & - & 26y & = & -17, \end{array} \right. \tag{2}$$

We can think about the problem graphically. One linear equation describes a line and so the solutions to the system (**??**) should be the point (or points) where the two lines meet. The solutions to (**??**) should be the point (or points) where the three lines meet.

Here is a graph (drawn by SAGE) of the lines (**??**)



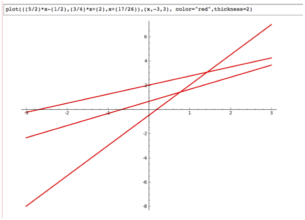
and this is the picture for (**??**)



You can see that, in this particular case, the two systems (**??**) and (**??**) have the same solution. There is just one solution $x = 10/13$, $y = 37/26$, or one point $(x, y) = (10/13, 37/26)$ on all the lines in (**??**). Same for (**??**).

However, you can also start to see what can happen in general. If you take a system of two linear equations in two unknowns, you will typically see two lines that are not parallel and they will meet in one point. But there is a chance that the two lines are parallel and never meet, or a chance that somehow both equations represent the same line and so the solutions are all the points on the line.

When you take a 'typical' system of 3 equations in 2 unknowns, you should rarely expect to see the kind of picture we just had. It is rather a fluke that 3 lines will meet in a point. Typically they will look like

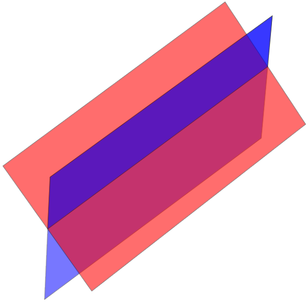


and will not meet in a point. You can see geometrically or graphically what is going on when you solve systems of equations in just 2 unknowns.

If we move to equations in 3 unknowns and consider two equations

$$\begin{cases} 5x - 2y + z = 4 \\ 3x + 4y + 4z = 10 \end{cases}$$

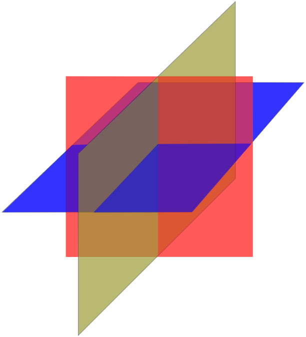
it is still possible to visualise what is going on. Recall that the solutions of a single linear equation in 3 variables (like $5x - 2y + z = 4$) can be visualised as the points on a (flat) plane in space. The common solutions of two such equations will usually be the points on the line of intersection of the two planes (unless the planes are parallel).



If we move to equations in 3 unknowns and consider three equations

$$\left\{\begin{array}{rcrcrcl} 5x & - & 2y & + & z & = & 4 \\ 3x & + & 4y & + & 4z & = & 10 \\ 26x & - & 26y & - & z & = & -7, \end{array}\right. \tag{3}$$

The common solution of three such equations will be a point



This idea of visualisation is good for understanding what sort of things can happen, but it is not really practical for actually calculating the solutions. Worse than that, we are pretty much sunk when we get to systems of equations in 4 or more unknowns. They need to be pictured in a space with at least 4 dimensions. And that is not practical.

Here is an example of a single linear equation in 4 unknowns x_1, x_2, x_3 and x_4

$$5x_1 - 2x_2 + 6x_3 - 7x_4 = 15$$

Solving systems of equations, preliminary approach

We turn instead to a recipe for solving systems of linear equations, a step-by-step procedure that can always be used. It is a bit harder to see what the possibilities are (about what can possibly happen) and a straightforward procedure is a valuable thing to have.

We'll explain with the examples (**??**) and (**??**).

Examples (I)

We'll start with the example (**??**).

$$\left\{\begin{array}{rcrcl} 5x & - & 2y & = & 1 \\ 3x & + & 4y & = & 8 \\ 26x & - & 26y & = & -17, \end{array}\right.$$

Step 1: Make first equation start with $1 \times x$ i.e. just x. In our example we do this by the step

$$\text{New Equation No. 1} = (\text{Old Equation No. 1}) \times \frac{1}{5}$$

and the result of this is

$$\left\{\begin{array}{rcrcl} x & - & \frac{2}{5}y & = & \frac{1}{5} \\ 3x & + & 4y & = & 8 \\ 26x & - & 26y & = & -17, \end{array}\right.$$

The idea of writing down the unchanged equations again (in their old places) is that we have all the information about the solutions in the new system.

Here is the reasoning. If we had numbers x and y that solved (**??**) then the same numbers have to solve the new system of 3 equations we got above. That is because we got the new equations by combining the old ones. But, we can also undo what we have done. If we multiplied the first of the 3 new equations by 5, we would exactly get back the information we had in (**??**).

Step 2: Eliminate x from all but the first equation.
We do this by subtracting appropriate multiples of the first equation from each of the equations below in turn. So, in this case, we do

$$\text{New Equation No. 2} \quad = \quad \text{Old Equation No. 2} - 3(\text{Old Equation No. 1})$$

$$\begin{cases} x & - & \frac{2}{5}y & = & \frac{1}{5} \\ & & \frac{26}{5}y & = & \frac{37}{5} \\ 26x & - & 26y & = & -17 \end{cases}$$

And then

$$\text{New Equation No. 3} \quad = \quad \text{Old Equation No. 3} - 26(\text{Old Equation No. 1}).$$

$$\begin{cases} x & - & \frac{2}{5}y & = & \frac{1}{5} \\ & & \frac{26}{5}y & = & \frac{37}{5} \\ & - & \frac{78}{5}y & = & -\frac{111}{5} \end{cases}$$

The next time we do this, we will do both these operations together. It is important that we do the steps so that we **could** do them one step at a time, but we can save some writing by doing both at once.

Step 3: Leave equation 1 as it is for now (it is now the only one with x in it). Manipulate the remaining equations as in step 1, but taking account of only the remaining variable y.

Like Step 1 Make the second equation start with $1y$.

$$\text{New Equation No. 2} = (\text{Old Equation No. 2}) \times \frac{5}{26}$$

$$\begin{cases} x - \frac{2}{5}y = \frac{1}{5} \\ y = \frac{37}{26} \\ -\frac{78}{5}y = -\frac{111}{5} \end{cases}$$

Like Step 2 Eliminate y from equations below the second.

$$\text{New Equation No. 3} = \text{Old Equation No. 3} + \frac{78}{5} \times (\text{Old Equation No. 2})$$

$$\begin{cases} x - \frac{2}{5}y = \frac{1}{5} \\ y = \frac{37}{26} \\ 0 = 0 \end{cases}$$

We are practically done now. The last equation tells us no information (it is true that $0 = 0$ but imposes no restrictions on x and y) while the second equation says what y is.

Now that we know both y we can find x from the first equation

$$x = \frac{1}{5} + \frac{2}{5}y = \frac{1}{5} + \frac{2}{5}\left(\frac{37}{26}\right) = \frac{10}{13}$$

So we end up with the solution

$$x = \frac{10}{13}, \; y = \frac{37}{26}$$

Example (II)

Example (III)

The point of the steps we make is partly that the system of equations we have at each stage has exactly the same solutions as the system we had at the previous stage. We already tried to explain this principle with the simplest kind of example above.

If x_1, x_2, x_3 and x_4 solve the original system of equations, then they must solve all the equations in the next system because we get the new equations by combining the original ones. But all the steps we do are reversible by a step of a similar kind. So any x_1, x_2, x_3 and x_4 that solve the new system of equations must solve the original system.

What we have ended up with is a system telling us what x_1, x_3 and x_4 are, but nothing to restrict x_2. We can take any value for x_2 and get a solution as long as we have

$$x_1 = \frac{4}{9} - \frac{1}{3}x_2, x_3 = \frac{1}{3}, x_4 = 4.$$

We have solved for x_1, x_3 and x_4 in terms of x_2, but x_2 is then a *free variable.*

Solving systems of equations, Gaussian elimination

We will now go over the same ground again but using a new and more concise notation.

Take an example again

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

In the last example, we were always careful to write out the equations keeping the variables lined up (x_1 under x_1, x_2 under x_2 and so on). As a short cut, we will write only the numbers, that is the coefficients in front of the variables and the constants from the right hand side. We write them in a rectangular array like this

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

and realise that we can get the equations back from this table because we have kept the numbers in such careful order.

A rectangular table of numbers like this is called a matrix in mathematical language. When talking about matrices (the plural of matrix is matrices) we refer to the rows, columns and entries of a matrix. We refer to the rows by number, with the top row being row number 1. So in the matrix

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

the first row is

$$[1 \quad 1 \quad 2 \quad 5] \ .$$

Similarly we number the columns from the left one. So column number 3 is

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \ .$$

Augmented matrices

The entries are the numbers in the various positions in the matrix. So, in our matrix above the number 5 is an entry. We identify it by the row and column it is on, in this case row 1 and column 4. So the $(1, 4)$ entry is the number 5.

In fact, though it is not absolutely necessary now, we often make an indication that the last column is special. We mark the place where the "=" sign belongs with a dotted vertical line, like this:

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

A matrix like this is called an augmented matrix.

Elementary row operations

The process of solving the equations can be translated into a process where we just manipulate the rows of an augmented matrix.

The steps that are allowed are called elementary row operations:

(i) multiply all the numbers is some row by a nonzero factor (and leave every other row unchanged)

(ii) replace any chosen row by the difference between it and a multiple of some other row.
That is, one can replace each entry in the chosen row by the sum of itself and the negative of a multiple of a corresponding entry of another row.

(iii) Exchange the positions of some pair of rows in the matrix.

We see that the first two operations, in terms of of linear equations, are just the steps we did to the equations above. The third operation, row exchange, just corresponds to writing the same list of equations in a different order. We need this to make it easier to describe the step-by-step procedure we are going to describe now.

Gaussian elimination

The procedure to corresponding to solving the equations is called *Gaussian elimination*. We will describe it so that there is a clear recipe to be followed, one that does not involve any choices.

Step 1: Organise top left entry of the matrix to be 1.

How:

- if the top left entry is already 1, do nothing
- if the top left entry is a nonzero number, multiply the first row across by the reciprocal of the top left entry
- if the top left entry is 0, but there is some nonzero entry in the first column, swap the position of row 1 with the first row that has a nonzero entry in the first column; then proceed as above
- if each entry in the first column is 0, ignore the first column and move to the next column; then proceed as above

Gaussian elimination

Step 2: Organise that the first column has all entries 0 below the top left entry.

How:

- subtract appropriate multiples of the first row from each of the rows below in turn

Step 3: Ignore the first row and first column and repeat steps **1**, **2** and **3** on the remainder of the matrix until there are either no rows or no columns left.

Once this is done, we can write the equations corresponding to the final matrix we found and solve them by back-substitution.

An Example

We'll carry out this process on our example, in hopes of explaining it better. So starting with the equations

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

we abstract to the following augmented matrix

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

Then we perform **step 1** and **step 2** for the first time.

$$\Rightarrow \quad \begin{bmatrix} 1 & 1 & 2 & : & 5 \\ 0 & -1 & -1 & : & -7 \\ 0 & -1 & -1 & : & -6 \end{bmatrix} \quad \begin{array}{l} \text{Old Row No. 2} - \text{Old Row No. 1} \\ \text{Old Row No. 3} - 2 \times \text{Old Row No.1} \end{array}$$

We now close our eyes to the first row and column and work essentially with the remaining part of the matrix.

$$\left[\begin{array}{c|cccc} 1 & 1 & 2 & : & 5 \\ \hline 0 & -1 & -1 & : & -7 \\ 0 & -1 & -1 & : & -6 \end{array}\right]$$

We don't actually discard anything, we keep the whole matrix, but we only look at rows and columns after the first for our subsequent procedures.

So we do **step 1** (now with reference to the smaller matrix).

$$\begin{bmatrix} 1 & 1 & 2 & : & 5 \\ 0 & 1 & 1 & : & 7 \\ 0 & -1 & -1 & : & -6 \end{bmatrix} \qquad \text{Old Row No. } 2\times(-1)$$

and **step 2** again

$$\Rightarrow \begin{bmatrix} 1 & 1 & 2 & : & 5 \\ 0 & 1 & 1 & : & 7 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \qquad \text{Old Row No. } 3 + \text{Old Row No. } 2$$

We should now cover over row 2 and column 2 (as well as row 1 and column 1) and focus our process on what remains.

All that remains is $0 : 1$. We skip the column of zeros and the next entry is just 1 and so we are done.

That is we have finished the Gaussian elimination procedure.

What we do now is write out the equations that correspond to this matrix. We thus get

$$\left\{\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 5 \\ & & x_2 & + & x_3 & = & 7 \\ & & & & 0 & = & 1 \end{array}\right.$$

As the last equation in this system is **never true** (no matter what values we give the unknowns) we can see that there are no solutions to the system.

They are called an inconsistent system of linear equations. So that is the answer: No solutions for this example.

Another example

Let's try an example that does have solutions. Consider the system of equations:

$$\begin{cases} & & 2x_2 & + & 3x_3 & = & 4 \\ x_1 & - & x_2 & + & x_3 & = & 5 \\ x_1 & + & x_2 & + & 4x_3 & = & 9 \end{cases}$$

To start we write it as an augmented matrix:

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

We now carry out the process of Gaussian elimination.

We first reorder the rows:

$$\left[\begin{array}{ccccc} 1 & -1 & 1 & : & 5 \\ 0 & 2 & 3 & : & 4 \\ 1 & 1 & 4 & : & 9 \end{array}\right] \qquad \begin{array}{l} \text{Old Row No. 2} \\ \text{Old Row No. 1} \\ \end{array}$$

Now we remove the "1" from the subsequent rows:

$$\left[\begin{array}{ccccc} 1 & -1 & 1 & : & 5 \\ 0 & 2 & 3 & : & 4 \\ 0 & 2 & 3 & : & 4 \end{array}\right] \qquad \begin{array}{l} \\ \\ \text{Old Row No. 3 - Old Row No. 1} \end{array}$$

Now ignoring the first row and column and repeating **step 1** and **step 2** on the remaining part of the matrix

$$\left[\begin{array}{ccccc} 1 & -1 & 1 & : & 5 \\ 0 & 1 & \frac{3}{2} & : & 2 \\ 0 & 2 & 3 & : & 4 \end{array}\right] \qquad \begin{array}{l} \\ \text{Old Row No. } 2\times\left(\frac{1}{2}\right) \\ \end{array}$$

$$\Rightarrow \quad \left[\begin{array}{ccccc} 1 & -1 & 1 & : & 5 \\ 0 & 1 & \frac{3}{2} & : & 2 \\ 0 & 0 & 0 & : & 0 \end{array}\right] \qquad \begin{array}{l} \\ \\ \text{Old Row No. } 3 - 2\times\text{Old Row No. 2} \end{array}$$

This matrix is now in the required form so we can write out the equations for this; we get

$$\left\{\begin{array}{rcrcrcl} x_1 & - & x_2 & + & x_3 & = & 5 \\ & & x_2 & + & \frac{3}{2}x_3 & = & 2 \\ & & & & 0 & = & 0 \end{array}\right.$$

or as the last line is trivial we can just drop it and we have

$$\left\{\begin{array}{rcl} x_1 & = & 5 + x_2 - x_3 \\ x_2 & = & 2 - \frac{3}{2}x_3 \end{array}\right.$$

Using back-substitution, we put the value for x_2 from the second equation into the first

$$\left\{\begin{array}{rcl} x_1 & = & 5 + \left(2 - \frac{3}{2}x_3\right) - x_3 \\ & = & 7 - \frac{5}{2}x_3 \\ x_2 & = & 2 - \frac{3}{2}x_3 \end{array}\right.$$

Thus we end up with this system of equations, which must have all the same solutions as the original system.

We are left with a free variable x_3, which can have any value as long as we take $x_1 = 7 - \frac{5}{2}x_3$ and $x_2 = 2 - \frac{3}{2}x_3$. So

$$\begin{cases} x_1 & = & 7 - \frac{5}{2}x_3 \\ x_2 & = & 2 - \frac{3}{2}x_3 \\ x_3 & \text{free} & \end{cases}$$

describes all the solutions. This is called a general solution of the system.

Row echelon form

The term row echelon form refers to the kind of matrix we must end up with after successfully completing the Gaussian elimination method.

Here are the features of such a matrix

- The first nonzero entry of each row is equal to 1 (unless the row consists entirely of zeroes). We refer to these as *leading ones*.
- The leading one of any row (after the first) must be at least one column to the right of the leading ones of any rows above.
- Any rows of all zeroes are at the end.

It follows that for a matrix in row echelon form there are zeroes directly below each leading one.

Note: In the following (and in the book) we will no longer single out the last column by using ":" marks. Consider as matrices there is nothing special about the last column.

Thus the matrix

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

is in row echelon form and so is

$$\begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

All the following matrices are also in row echelon form:

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

Gauss-Jordan elimination

In the previous lecture we calculated the row echelon form of the augmented matrices associated with a linear equation and then we returned to the equation to perform the back-substitution.

We mentioned the alternative to using back-substitution called Gauss-Jordan elimination which is almost equivalent to doing the back substitution steps on the matrix.

- **Steps 1–3** Follow the steps for Gaussian elimination (resulting in a matrix that has row-echelon form).
- **Step 4** Starting from the bottom, arrange that there are zeros above each of the leading ones.

Step 4

How:

- Starting with the last nonzero row, subtract appropriate multiples of that row from all the rows above so as to ensure that there are all zeros above the leading 1 of the last row. (There must already be zeros below it.)
- Leave that row aside and repeat the previous point on the leading one of the row above, until all the leading ones have zeroes above.

Following completion of these steps, the matrix should be in what is called *reduced row-echelon form*. That means that the matrix should be in row-echelon form **and** the column containing each leading one must have zeroes in every other entry.

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

Some more examples of matrices in reduced row echelon form are:

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

and

$$\begin{bmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

All the following matrices are also in reduced row echelon form:

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

Let us return to the example

$$\begin{cases} & & 2x_2 & + & 3x_3 & = & 4 \\ x_1 & - & x_2 & + & x_3 & = & 5 \\ x_1 & + & x_2 & + & 4x_3 & = & 9 \end{cases}$$

and finish it by the method of Gauss-Jordan. We had the row-echelon form

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

$$\to \begin{bmatrix} 1 & 0 & \frac{5}{2} & 7 \\ 0 & 1 & \frac{3}{2} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Old Row No. 1} + \text{Old Row No. 2}$$

and we already have the matrix in reduced row-echelon form.

When we write out the equations corresponding to this matrix, we get

$$\begin{cases} x_1 \quad\quad\quad + \frac{5}{2}x_3 = 7 \\ \quad\quad x_2 + \frac{3}{2}x_3 = 2 \\ \quad\quad\quad\quad\quad\quad 0 = 0 \end{cases}$$

and the idea is that each equation tells us what the variable is that corresponds to the leading one.

Writing everything but that variable on the right we get

$$\begin{cases} x_1 = 7 - \frac{5}{2}x_3 \\ x_2 = 2 - \frac{3}{2}x_3 \end{cases}$$

and no equation to tell us x_3. So x_3 is a free variable.

Yet Another Example

Consider the matrix:

General solution

This matrix is now in reduced row echelon form and we could now rewrite it so that the general solution to the corresponding system of linear equations is found.

$$
\begin{aligned}
&x_1 + 2x_2 + 3x_4 = 7 \\
&x_3 = 1 \\
&x_5 = 2
\end{aligned}
$$

It is common to introduce the arbitrary parameters s and t such that $x_2 = s$ and $x_4 = t$ so that we can write the equations in the form

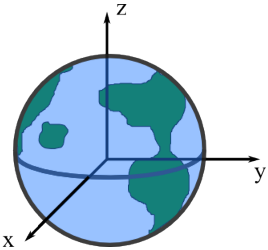
$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}
$$

This is the general solution written as a linear combination of column vectors (here in $\mathbb{R}^5$).

- Every matrix has a unique reduced row echelon form.
- Row echelon forms of matrices are not unique, however the every row echelon form has its leading ones in the same positions - these positions are called pivot positions and the columns that have leading ones are called pivot columns.
- Every row echelon form also has the same number of rows of zeros at the bottom. The rank of a matrix is the number of non-zero rows in its row echelon form.
- For a system of linear equations the number of free parameters in the general solution is the number of original variables minus the rank of the augmented matrix for the linear system.

An Application: GPS

As a basic explanation of the Global Positioning System, let us assume the earth is an exact sphere and let us take our coordinate system to have its origin at the centre of the earth:



Let us assume that there is some ship with some coordinates (x, y, z) at some time t. We will take all distances to be measured in units of the earth's radius, R_e, so that (assuming the ship is on the surface of the earth!)

$$x^2 + y^2 + z^2 = 1 \ .$$

The GPS identifies the ship's coordinates using a triangulation computed from satellites which are in orbits chosen so that between five and eight are always in line of sight of every point on earth.

An Application: GPS

Distances from each satellite can be computed using the speed of light $3 \times 10^8 m/s$ or 0.469 Earth radii per hundredth of a second. For example if a ship receives a signal at time t and the signal indicates it was sent at t_0 then the distance travelled is

$$d = 0.469(t - t_0) \ .$$

In principle only three ship-to-satellite distances are sufficient to determine the ships position. However as most ships do have have atomic clocks that can measure the time accurately, the time t must be considered a separate unknown and four satellites are needed.

For example if the ship receives a signal from four satellites encoding the satellites position, (x_0, y_0, z_0) and the time the signal is sent t_0 then the time can be computed from

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

and to three significant digits

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = 0.22(t - t_0)^2 \ .$$

While this equation is second order we will see that by solving a system of three linear equations and a single quadratic equation we can determine the position of the ship.

An Application: GPS

Suppose a ship with unknown position (x, y, z) at an unknown time t receives the following data from four satellites (giving position in terms of Earth radii and time from midnight in hundredths of second):

Satellite	Satellite position	Time
1	$(1.12, 2.10, 1.40)$	1.06
2	$(0.00, 1.53, 2.30)$	2.30
3	$(1.40, 1.12, 2.10)$	1.16
4	$(2.30, 0.00, 1.53)$	0.75

From the data for the first satellite we get the equation:

$$\begin{aligned} & (x - 1.12)^2 + (y - 2.10)^2 + (z - 1.4)^2 = 0.22(t - 1.06)^2 \\ \Rightarrow \quad & 2.24x + 4.20y + 2.8z - 0.466t = x^2 + y^2 + z^2 - 0.22t^2 + 7.377 \; . \end{aligned}$$

From the remaining three signals we get

$$\begin{aligned} 0x + 3.06y - 4.6z - 0.246t &= x^2 + y^2 + z^2 - 0.22t^2 + 7.562 \\ 2.8x + 2.24y + 4.2z - 0.51t &= x^2 + y^2 + z^2 - 0.22t^2 + 7.328 \\ 4.6x + 0y + 3.06z - 0.33t &= x^2 + y^2 + z^2 - 0.22t^2 + 7.507 \; . \end{aligned}$$

An Application: GPS

The quadratic parts of all these equations are the same so we can subtract each of the last three from the first we find the three linear equations:

$$\begin{aligned} 2.24x + 1.14y - 1.8z - 0.22t &= -0.185 \\ -0.56x + 1.96y - 1.4z + 0.044t &= 0.049 \\ -2.36x + 4.2y - 0.26z - 0.136t &= -0.13 \ , \end{aligned}$$

or equivalently the augmented matrix

$$\begin{bmatrix} 2.24 & 1.14 & -1.8 & -0.22 & -0.185 \\ -0.56 & 1.96 & -1.4 & 0.044 & 0.049 \\ -2.36 & 4.2 & -0.26 & -0.136 & -0.13 \end{bmatrix}$$

As there are three equations in four unknowns we expect to find a one-parameter family of solutions. This solution can then be substituted back into one of the quadratic equations to fix the parameter uniquely and so determine (x, y, z).

The reduced row echelon form of the augmented matrix is

$$\begin{bmatrix} 1 & 0 & 0 & -0.153 & -0.139 \\ 0 & 1 & 0 & -0.127 & -0.118 \\ 0 & 0 & 1 & -0.149 & -0.144 \end{bmatrix}$$

from which we can find the general solution

$$x = 0.153t - 0.139 \ , \quad y = 0.127t - 0.118 \ , \quad z = 0.149t - 0.144 \ .$$

We can now substitute these expressions into any of the quadratic equations, for example

$$\begin{aligned} & 4.6x + 0y + 3.06z - 0.33t + x^2 + y^2 + z^2 - 0.22t^2 + 7.507 = 0 \ , \\ \Rightarrow \quad & 8.641 - 0.944t - 0.158t^2 = 0 \end{aligned}$$

which has the solutions $t = -10.95$ and 4.99. As $t > 0$ (the ship can't receive signals before they were sent!) $t = 4.98$ and so

$$(x, y, z) = (0.623, 0.518, 0.597) \ .$$

Let us check that is indeed a point on the surface of the earth:

$$x^2 + y^2 + z^2 = 1.014 \quad !?!$$

Computer round-off error! Need to carry out the calculation to greater precision!

Homogeneous linear systems

We mentioned that linear equations of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

with zero constant on the r.h.s. are called homogeneous equations. A homogeneous linear system is a set of m homogeneous equations in n variables

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & 0 \end{cases}$$

- Every homogeneous linear system is consistent as there is always at least one solution

$$x_1 = x_2 = \cdots = x_n = 0 \ .$$

- If there is one non-trivial solution there is an infinitely family of solutions. If

$$x_1 = s_1 \ , \quad x_2 = s_2 \ , \ldots , x_n = s_n \ .$$

  is a solution then so is

$$x_1 = ts_1 \ , \quad x_2 = ts_2 \ , \ldots , x_n = ts_n \ ,$$

  for any t.

- The augmented matrix for a homogeneous system has a final column of all zeros. The reduced row echelon form also has a final column of zeros so that the final equations are also homogeneous.

An Example

Consider the system of homogeneous equations

$$\begin{array}{rcrcrcrcrcrcl} x_1 & + & 3x_2 & - & 2x_3 & & & + & 2x_5 & & & = & 0 \\ 2x_1 & + & 6x_2 & - & 5x_3 & - & 2x_4 & + & 4x_5 & - & 3x_6 & = & 0 \\ & & & & 5x_3 & + & 10x_4 & & & + & 15x_6 & = & 0 \\ 2x_1 & + & 6x_2 & & & + & 8x_4 & + & 4x_5 & + & 18x_6 & = & 0 \end{array}$$

The augmented matrix is

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 & 0 \\ 2 & 6 & 0 & 8 & 4 & 18 & 0 \end{bmatrix}$$

while the reduced row echelon form is

$$\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note the last column is still all zeros.

The corresponding system of equations is:

$$\begin{array}{cccccccccccc} x_1 & + & 3x_2 & & & + & 4x_4 & + & 2x_5 & & = & 0 \\ & & & & x_3 & + & 2x_4 & & & & = & 0 \\ & & & & & & & & & x_6 & = & 0 \end{array}$$

Hence assigning the parameters r, s and t respectively to the variables x_2, x_4 and x_5 we have

$$x_1 = -3r - 4s - 2t\ , \quad x_2 = r\ , \quad x_3 = -2s\ , \quad x_4 = s\ , \quad x_5 = t\ , \quad x_6 = 0.$$

- For a homogeneous linear system in n unknowns with a reduced row echelon form with r non-zero rows, the systems has $n - r$ free variables.
- A homogeneous linear system with more unknowns than equations has infinitely many solutions.

More Applications

There are essentially far too many applications of solving linear by Gaussian elimination or Gauss-Jordan to even begin to do justice to the power of the method.

To list just a few common examples

- Studying network flows e.g. traffic patterns
- Analyzing electrical circuits
- Balancing chemical equations

Instead let us consider one example from social sciences and one from data analysis.

Leontief input-output model

Consider a model of an economy comprising three sectors:

- service
- manufacturing
- natural resources

In a given year each sector produces a certain amount of labour, goods or resources (output) and in doing so consumes a certain amount of the same (input). We assume that there is no trade with other economies and that each sector produces as much as it consumes (this is a closed economy in equilibrium). Now if we know how much of each output is consumed by each sector e.g.

		Produced		
		Service	Man.	Nat. Res.
Consumed	Service	0.25	0.33	0.5
	Man.	0.25	0.33	0.25
	Nat. Res.	0.5	0.33	0.25

That is the service sector consumes 1/4 of its own output, 1/3 of the manufactured goods, and 1/2 of the natural resources produced.

Leontief input-output model

Let x_1, x_2 and x_3 denote the output of the service, manufacturing and natural resource sector. The total consumption of each sector must equal its output we have that for the service sector:

$$x_1 = \frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3$$

i.e. the value of services produced must equal the value consumed by the service sector. There are corresponding equations for the other two sectors

$$\begin{aligned} \text{Manufacturing}: \quad & x_2 = \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 \\ \text{Natural Resources}: \quad & x_3 = \tfrac{1}{2}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 \end{aligned}$$

Of course this is nothing but a system of homogeneous linear equations. We can solve these equations and find out what the relative output of the various sectors must be for the economy to be in equilibrium.

Leontief Closed Economy

Writing the linear equations as an augmented matrix

$$\begin{array}{rcrcrcrcl} -\frac{3}{4} & x_1 & + & \frac{1}{3} & x_2 & + & \frac{1}{2} & x_3 & = & 0 \\ \frac{1}{4} & x_1 & - & \frac{2}{3} & x_2 & - & \frac{1}{4} & x_3 & = & 0 \\ \frac{1}{2} & x_1 & + & \frac{1}{3} & x_2 & - & \frac{3}{4} & x_3 & = & 0 \end{array} \rightarrow \begin{bmatrix} -\frac{3}{4} & \frac{1}{3} & \frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{2}{3} & -\frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{3} & -\frac{3}{4} & 0 \end{bmatrix}$$

we can follow the procedure to put this into reduced row echelon form.

$$\begin{bmatrix} -\frac{3}{4} & \frac{1}{3} & \frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{2}{3} & -\frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{3} & -\frac{3}{4} & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{4}{9} & -\frac{2}{3} & 0 \\ \frac{1}{4} & -\frac{2}{3} & -\frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{3} & -\frac{3}{4} & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{4}{9} & -\frac{2}{3} & 0 \\ 0 & -\frac{5}{9} & \frac{5}{12} & 0 \\ 0 & \frac{5}{9} & -\frac{5}{12} & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & -\frac{4}{9} & -\frac{2}{3} & 0 \\ 0 & 1 & -\frac{3}{4} & 0 \\ 0 & \frac{5}{12} & -\frac{5}{12} & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{4}{9} & -\frac{2}{3} & 0 \\ 0 & 1 & -\frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix now gives the general solution of the linear system of equations. We can see that as we have three unknowns and one row of zeros we will have a one parameter family of solutions.

Leontief Closed Economy

Writing $x_3 = t$ we have

$$x_1 - t = 0$$
$$x_2 - \frac{3}{4}t = 0$$

Thus we see that for the economy to be in equilibrium the outputs of the service, manufacturing and natural resource production sectors need to be in the ration $x_1 : x_2 : x_3 = 4 : 3 : 4$.

Of course this is only the crudest possible model. For example, instead of considering entire sectors we could model individual goods and firms. In a realistic economy the number of variables would become immense, however the basic problem would remain the same.

More profound assumptions include the fact that most economies or countries trade with other countries. We can include this effect by allowing, for example, external consumption (that is products and goods shipped abroad).

Leontief Open Economy

Let us assume that we have the following input-output table

		Produced		
		Service	Man.	Nat. Res.
Consumed	Service	0.2	0.5	0.1
	Man.	0.4	0.2	0.2
	Nat. Res.	0.1	0.3	0.3

That is for example, that of goods produced by the service economy 20% are consumed by the service sector itself, 40% by the manufacturing sector and 10% by the natural resource sector. In this case there will be excess production which can be exported. Similarly there is excess production in the natural resource sector, this sector is productive.

Let us assume that there is €10 billion worth of external demand for services, €10 billion worth of external demand for manufactured goods, and €30 billion worth of external demand for natural resources.

Leontief Open Economy

The total value of services produced must equal those consumed by the service sector plus the value of those services exported, so we have:

$$x_1 = 0.2x_1 + 0.4x_2 + 0.1x_3 + 10$$

Similarly the other sectors must produce as much as they consume plus whatever excess they export:

$$x_2 = 0.4x_1 + 0.2x_2 + 0.2x_3 + 10$$
$$x_3 = 0.1x_1 + 0.3x_2 + 0.3x_3 + 30$$

The augmented matrix for this system is

$$\begin{bmatrix} 0.8 & -0.5 & -0.1 & 10 \\ -0.4 & 0.8 & 0.2 & 10 \\ -0.1 & -0.3 & 0.7 & 30 \end{bmatrix}$$

while we won't go through the steps this has the reduced row echelon form

$$\begin{bmatrix} 1 & 0 & 0 & 61.74 \\ 0 & 1 & 0 & 63.04 \\ 0 & 0 & 1 & 78.70 \end{bmatrix}$$

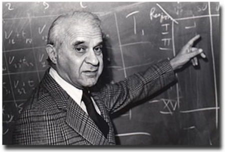
Leontief Open Economy

From the reduced row echelon form we can immediately write down the the solution:

$$x_1 = 61.74\ , \quad x_2 = 63.04\ , \quad x_3 = 78.70\ .$$

That is to say for this open economy to be in equilibrium the service sector must produce €61.74 billion worth of goods, the manufacturing must produce €63.04 billion and the natural resources €78.74 billion.

Even from the form of reduced row echelon form we can see that as there are no zero rows that we will no free parameters and so the solution is uniquely fixed. This is ok as they are not homogeneous equations anymore.



The use of such input-output models (here vastly simplified) was pioneered by Wassily Leontief who compiled large tables in the 1930s of economic data breaking the US economy into ~ 500 sectors. In part for this work he won the Nobel Memorial Prize in Economics in 1973.