Math 1180, fall, 2009

notes #2

We have seen several examples of systems of linear equations. The last example in the previous notes consisted of two equations involving three unknown quantities. This leads us to discuss the general case, with m equations and n unknowns , where m and n are positive integers. We could have $m = n$, $m > n$, or $m < n$. We can write these equations as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + .\cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + .\cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + .\cdots + a_{mn}x_n &= b_m, \end{aligned}$$

where $x_1, ..., x_n$ are the unknowns. The numbers a_{ij}, with $1 \le i \le m$ and $1 \le j \le n$ are called the "coefficients."

We now discuss how to solve such a system of m equations in n unknowns, or else determine that there is no solution.

We start with a system of 2 equations in 3 unknowns. We have discussed how this is the intersection of two planes in R^3 and therefore is a line, if the normal vectors to the planes are not parallel. We will show how to get the so-called "parametric" form of this line.

Consider the system

$$\begin{aligned} x - y + z &= 1 \\ x + y - 2z &= -2 \end{aligned} \tag{1}$$

The normal vectors are $(1, -1, 1)$ and $(1, 1, -2)$, and neither is a scalar multiple of the other, so these planes intersect in a line. Subtract the first equation from the second to get a new system

$$\begin{aligned} x - y + z &= 1 \\ 2y - 3z &= -3. \end{aligned}$$

Having recognized that our solution is a line, we don't expect there to be a unique solution, which would be a single point on the intersection. So, we realize that some variable can be selected arbitrarily. We will assume this value is z. We solve for x and y in terms of z. We find from the last equation that

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

and then from the first equation that

$$x = 1 + y - z = 1 - \frac{3}{2} + \frac{3}{2}z - z = -\frac{1}{2} + \frac{1}{2}z.$$

Thus, the "general" solution can be written as

$$(x, y, z) = \left(-\frac{1}{2} + \frac{1}{2}z, -\frac{3}{2} + \frac{3}{2}z, z\right).$$

Choosing any number for z, we get a solution $(x, y, z.)$ For example, one solution is $\left(-\frac{1}{2}, -\frac{3}{2}, 0\right)$, and another is $(0, 0, 1)$.

The text introduces "vectors" in section 1.5, but does not exploit this notation in giving solutions to systems of equations. But I will do so now.

Definition 1 *A vector in R^n is an ordered set of real numbers $(x_1, ..., x_n)$.*

The word "ordered" simply means that it matters which of the numbers is called the first, which the second, and so forth. We can write sets of numbers using curly brackets, such as $\{1, 3, -7\}$. The order in which we list them doesn't matter. This is the same set as $\{-7, 1, 3\}$. But the vectors $(1, 3, -7)$ and $(-7, 1, 3)$ are different.

In discussing systems of linear equations, it is convenient to write vectors in columns, such as $\begin{pmatrix} 1 \\ 3 \\ -7 \end{pmatrix}$. We will usually use columns to write our vectors. Only occasionally is it necessary to discuss a "row vector", such as $(1, 3, -7)$. For our purposes these are the same thing, but columns turn out to be a more convenient way of writing them.

We can add vectors component wise. This means that

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}.$$

We can also multiply a vector by a scalar. The definition is:

$$c\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} cx \\ cy \\ cz \end{pmatrix}.$$

Hence, we can give our general solution to the system (1) as

$$\mathbf{x} = \begin{pmatrix} \frac{1}{2} + \frac{1}{2}z \\ -\frac{3}{2} + \frac{3}{2}z \\ z \end{pmatrix}$$

or as

$$\mathbf{x} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{3}{2} \\ 0 \end{pmatrix} + z\begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}.$$

We will always use bold face for a vector in these notes, while on the board in class, they will be underlined. Bold face is used in the text, also, as in example 7 on page 52.

Here z is the parameter. We could replace it with any other letter, such as t or s , so that the solution is given as

$$\mathbf{x} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{3}{2} \\ 0 \end{pmatrix} + t\begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{pmatrix}.$$

Again, t can be any real number, and we get every possible solution in this way.

Now an example in R^4. (This means that the number of unknowns is 4, so that a solution can be written as a vector $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ in R^4.) For dimensions greater than 3, a new geometrical term is used:

Definition 2 *A "hyperplane" in R^4 is the set of all vectors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ in R^4 which satisfy a single linear equation*

$$ax_1 + bx_2 + cx_3 + dx_4 = e.$$

Suppose we have two equations in four variables, such as

$$\begin{array}{c} x_1 - x_2 + x_3 - 3x_4 = 1 \\ 2x_1 + x_2 - x_3 - x_4 = 2 \end{array} .$$

The set of solutions is the intersection of two hyperplanes. Subtract twice the first equation from the second to get a new second equation and give the system

$$\begin{array}{c} x_1 - x_2 + x_3 - 3x_4 = 1 \\ 3x_2 - 3x_3 + 5x_4 = 0 \end{array} \tag{2}$$

Then we can solve the second equation for x_2 in terms of x_3 and x_4 and substitute this into the first equation to get x_1. We obtain

$$\begin{aligned} x_2 &= x_3 - \frac{5}{3}x_4 \\ x_1 &= x_3 - \frac{5}{3}x_4 - x_3 + 3x_4 + 1 = 1 + \frac{4}{3}x_4. \end{aligned}$$

Hence

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} \frac{4}{3} \\ -\frac{5}{3} \\ 0 \\ 1 \end{pmatrix} .$$

Now x_3 and x_4 are the parameters (or we could use s and t, for example). This is the general solution of the system.

We now notice that in the solutions above we are doing a lot of unnecessary writing. We do not have to keep writing $x_1, x_2,$ and so forth. We only have to keep track of the coefficients of the equations as we change from one equation to another, and also of the b's on the right-hand side. For a general system, we consider the "matrix" of coefficients

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

and the right side

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot\cdot \\ b_n \end{pmatrix} .$$

We combine these into what is called the "augmented" matrix for the system:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot\cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right)$$

The vertical line is just meant to separate the coefficients from the right side, and show that this is an augmented matrix. It does not affect any of the calculations.

Having formed the augmented matrix, we then "add and subtract" equations. Subtracting a multiple of one equation from another is the same as doing these operations with the corresponding rows of the augmented matrix.

There are clearly many ways of doing this. But the method of Gaussian elimination requires that you do it in a particular order, as follows.

1. Subtract multiples of the first row from the other rows in such a way as to get all zeros in the first column below the first row.

2. Subtract multiples of the new second row from the rows below it in order to get all zeros in the second column below the second row.

3. Continue in this way as long as you can, moving from column to column. You may reach a situation where you can't eliminate the desired non-zero terms. We will see how this can happen below, and what to do about it. But usually you will be able keep going until you reach either the last row or the next to last column of the augmented matrix, or both. You want to get $a_{ij} = 0$ for each entry where $i > j$, i.e. below each "diagonal" term, which we will take to be a term a_{ii}, even if the matrix is not square. But you do not do this in the last column – the one which makes the matrix "augmented".

It is very important that you include the last column in your additions and subtractions. It's just that you stop the procedure of eliminating nonzero terms after the next to last column.

The resulting matrix is the augmented matrix for a new system of equations. But the solutions to this new system will be exactly the same as the solutions of the original system, because all you are really doing is adding or subtracting equations, which does not change the solutions.

Here is an example:

$$\begin{gathered} x_1 + 3x_2 + x_4 = 1 \\ x_3 + 4x_4 = 6 \\ x_1 + 3x_2 + x_3 + 6x_4 = 7. \end{gathered}$$

The augmented matrix is

$$\begin{pmatrix} 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 1 & 3 & 1 & 6 & 7 \end{pmatrix}$$

and the first Gaussian elimination step is

$$\begin{pmatrix} 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 1 & 3 & 1 & 6 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 1 & 5 & 6 \end{pmatrix}$$

This takes care of the first column, and by luck, also handles the second column. In these examples it happens that at this stage $a_{22} = 0$. This won't always be the case, and affects the answer.

The next Gaussian elimination step is to subtract the 2nd row from the third.

$$\begin{pmatrix} 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 1 & 5 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Now let's write down the new equations.

$$\begin{aligned} x_1 + 3x_2 + x_4 &= 1 \\ x_3 + 4x_4 &= 6 \\ x_4 &= 0. \end{aligned}$$

The answer is now obvious. We work backwards from x_4, in a process called "back substitution." We get $x_4 = 0$, $x_3 = 6$. This leaves us with the equation

$$x_1 + 3x_2 = 1.$$

We make x_2 the parameter, with

$$x_1 = 1 - 3x_2.$$

The general solution is then

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 6 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Remark: *It is very important that you do the operations in the order described above. It is possible to solve the equations in other ways, but if you do so, some features of the solutions will not be the same, and also, some other operations done by Gaussian elimination will simply give the wrong answer.*

In particular, if you add (or subtract) a multiple of the i^{th} row to (or from) the j^{th} row, the result must always go in the j^{th} row, never in the i^{th} row.

Here's another example. I will not bother with the equations, but go straight to the augmented matrix and show the steps:

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 8 & 10 \\ 3 & 3 & 10 & 13 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 \\ 3 & 3 & 10 & 13 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can write down the set of equations corresponding to the last matrix:

$$\begin{aligned} x + y + 2z &= 3 \\ 4z &= 4 \\ 0 &= 0. \end{aligned}$$

The last equation is irrelevant. We get $z = 1$ and then

$$x = 1 - y.$$

We can therefore write the general solution as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

In this notation y is the parameter instead of s or t or x_3, etc.

You will see a similar set of operations in the text, on page 11. But there, the author does a few more steps. This does not matter – he will get the same answer as we will. But, as he points out eventually, the extra steps are more work than are needed, and the answer can be gotten more quickly. We will discuss this further later on.

Reading: Sections 1.2, 1.3 (through pg. 22) of the text.

Homework, due at the beginning of class on September 9.

Problems 1-3: pg. 25, # 2, 6, 32

4. On pages 21 and 22 of the text there are several pictures of how three planes can intersect. Find specific examples of systems of three equations for each of these pictures. Hint: The system

$$\begin{aligned} x_1 &= 0 \\ x_2 &= 0 \\ x_3 &= 0 \end{aligned}$$

fits figure 1.2 , because the three planes meet at exactly one point, as in that figure.

5. Now find a specific set of equations in which no two of the three planes are parallel, but still, there is no solution to the system. (This situation is not pictured in the examples in the text.)

Hint: Find the equations of three lines in the x, y plane such that no two are parallel, but the three lines don't meet at a common point. Now turn this into equations involving three unknowns, possibly with some zero coefficients.