January 9

Last Time

1. Last time we ended with saying that the following four systems "are equivalent" in the sense that we can move from one system to the other by a special move we discussed.

   (a)
   $$\begin{aligned} x + 3y &= 6 \\ -x + 2y &= -1 \end{aligned}$$

   (b)
   $$\begin{aligned} x + 3y &= 6 \\ y &= 1 \end{aligned}$$

   (c)
   $$\begin{aligned} 2x + 4y &= 8 \\ x - 2y &= 1 \end{aligned}$$

   (d)
   $$\begin{aligned} x + 3y &= 6 \\ -2x - y &= -7 \end{aligned}$$

Lecture 2

1. Which one of the above equations is quicker to solve? The second one, of course. Because you already know what y is, you can just plug it in the first equation and you have the x value. So, it is a reasonable thing to want to turn our system to that form. What do we do? We add the first equation to the second and replace the second equation with this. Then, we divide the second equation by 5. Easy.

2. If we understand how this worked, we can easily generalize this to more variables! It is easy to solve the system:
   $$\begin{aligned} x + 2y + 3z &= 6 \\ 3y + z &= 4 \\ z &= 1 \end{aligned}$$
   if you start from the bottom. You start from the very bottom equation and when you go to the top you be like "Started from the bottom now we here". If we do some manipulations as above, we see that this system and the following system have the same solution set:
   $$\begin{aligned} x + 2y + 3z &= 6 \\ x + 5y + 4z &= 10 \\ 2x + 4y + 7z &= 13 \end{aligned}$$

and this has the same solution set with

$$\begin{aligned} x + 2y + 3z &= 6 \\ x + 5y + 4z &= 10 \\ 3x + 9y + 11z &= 23 \end{aligned}$$

If we are given this last system, we should want to go back to the first system and solve that.

3. Imagine you are a mathematician in the beginning of 20th century and you are solving a lot of equations of this form with 10, 15 variables. The amount of work is already huge when you deal with 4 or more variables. What do you do? One answer: You invent computers. Another answer: You invent a new way to write these equations so that you don't have to keep writing the x, y, z all the time! Since all we do is to add/multiply the coefficients in front of x, y, z I only need the coefficients. As long as I remind myself which number is a coefficient for which variable in which equation, I am okay! So, I will transform the last system above to

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 4 \\ 3 & 9 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ 23 \end{bmatrix}$$

This is just a notation! We have an array - or a matrix of numbers on the left of the equation that we manipulate with the operations we discussed above. However, we should be careful and note that when we perform the abovementioned operations on this matrix of coefficients, we also perform them on the array of numbers on the right of the equation. In order not to forget this, we will completely forget x, y, z and we will write

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 1 & 5 & 4 & 10 \\ 3 & 9 & 11 & 23 \end{array}\right]$$

with an augmentation line and call it an augmented matrix.

4. Now, we know what we should be doing. **Example.** Solve the system

$$\begin{aligned} x + 2y + 3z &= 6 \\ x + 5y + 4z &= 10 \\ 3x + 9y + 11z &= 23 \end{aligned}$$

Solution. We start with putting the system to augmented form:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 1 & 5 & 4 & 10 \\ 3 & 9 & 11 & 23 \end{array}\right]$$

and we perform the row operations we discussed last time in order to reduce it to an equivalent

- but easier - form

$$\left[\begin{array}{ccc|c}1 & 2 & 3 & 6\\ 1 & 5 & 4 & 10\\ 3 & 9 & 11 & 23\end{array}\right] \xrightarrow{R_2-R_1\to R_2} \left[\begin{array}{ccc|c}1 & 2 & 3 & 6\\ 0 & 3 & 1 & 4\\ 3 & 9 & 11 & 23\end{array}\right]$$
$$\xrightarrow{R_3-3R_1\to R_3} \left[\begin{array}{ccc|c}1 & 2 & 3 & 6\\ 0 & 3 & 1 & 4\\ 0 & 3 & 2 & 5\end{array}\right]$$
$$\xrightarrow{R_3-R_2\to R_3} \left[\begin{array}{ccc|c}1 & 2 & 3 & 6\\ 0 & 3 & 1 & 4\\ 0 & 0 & 1 & 1\end{array}\right]$$

This augmented form corresponds to the system

$$\begin{aligned} x+2y+3z &= 6\\ 3y+z &= 4\\ z &= 1\end{aligned}$$

which can be solved easily to find

$$z=1, y=1, x=1$$

or in other words the solution set is

$$\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix} \right\}$$

5. Now, we will do all of this in a more general setting. We will start with introducing some vocabulary. Note that you should start getting comfortable with new vocabulary - in this course you will learn a lot of new words. You should get comfortable with them quickly -via practice- because each new word depends on the previous ones.

6. We say that an equation of the form

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

where $a_1, \ldots, a_n$ are (real) numbers is a **linear equation** in n variables. The numbers $a_1, \ldots, a_n$ are called coefficients and the number b is called the constant term.

7. We say that

$$\begin{array}{ccccccccc} a_{11}x_1 &+& a_{12}x_2 &+& \ldots &+& a_{1n}x_n &=& b_1\\ a_{21}x_1 &+& a_{22}x_2 &+& \ldots &+& a_{2n}x_n &=& b_2\\ \vdots && \vdots && \ldots && \vdots &=& \vdots\\ a_{k1}x_1 &+& a_{k2}x_2 &+& \ldots &+& a_{kn}x_n &=& b_k\end{array}$$

is a **system of linear equations** with k equations and n unknowns/variables. **The augmented matrix** of this system is

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1\\ a_{21} & a_{22} & \ldots & a_{2n} & b_2\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ a_{k1} & a_{k2} & \ldots & a_{kn} & b_k\end{array}\right]$$

8. When we are given such a system, as we discussed earlier our algorithm is to reduces it to an "easier" but equivalent system with the following **elementary row operations**:
   - Interchange two rows ($R_i \leftrightarrow R_j$).
   - Multiply a row by a *nonzero* number k ($kR_i \to R_i$).
   - Add a multiple of one row to another ($R_i + kR_j \to R_i$).

   Note that you can always combine the last two and have the operation $aR_i + bR_j \to R_i$ where $a \neq 0$.

9. **Important.** When you apply these row operations, apply one operation at a time and **always** write down the exact operations at each step as we did in the previous example. Not only it minimizes the risk of a simple error, but also it allows you to check your work easily after you are done. Also, it is less work for the graders in the exams!

10. Every elementary row operation preserves the set of solutions to the system - we discussed this in a simpler case in the first lecture.

11. During this course, you should frequently stop, go back, and remember what we have been trying to do. We would like to solve systems of linear equations. We saw that we can apply some *row operations* in order to get an "easier" system without changing the solution set. What do we exactly mean by an "easier" system? If we would liket to teach our algorithm to a computer, where should it stop?

12. An augmented matrix is in **Row Echelon Form - REF** if all the following properties hold:
    (a) All rows consisting entirely of zeroes are at the bottom of the matrix.
    (b) The first nonzero entry from the left in each nonzero row is equal to 1.
    (c) Each leading 1 is to the right of the leading 1s above it.

    Exercise. Give several examples of augmented matrices which are in echelon form, which are not in echelon form.
    If in addition to the above properties, we have the fourth property

    (d) Each leading 1 is the only nonzero entry in its column.

    then we say that our matrix is in **Reduced Row Echelon Form - RREF**. As an **exercise**, give examples of matrices which are in REF but not in RREF.

13. Make sure that you understand why RREF is much more better than REF. Indeed, start with an augmented matrix which is in RREF and write down the corresponding system. What do you see?

14. **Fact.** Every augmented matrix has a **unique** RREF. (Note: REF doesn't have to be unique).

15. Well, then, our main purpose is to start with a system and converting it to its RREF. This will allow us to solve our system and also we can tell a computer to do it for us! How do we convert an augmented matrix to its RREF? **Gaussian Algorithm.** It is really a simple algorithm, here is the recipe:
    - If the matrix is all zeroes: Stop.

- Find the first column with a nonzero c entry and move the row containing c to the top row via interchanging two rows.
- Multiply that row with $1/c$ to get a leading 1.
- Use this leading one to clear its column.
- Now apply the algorithm to the matrix without the first row.

16. Now, we know how to solve linear equations. Note that this algorithm can be explained to a computer.

17. Our next goal is to understand how many solutions a system can have. We look at three examples.

(a)

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

One can immediately see that $(1, 2)$ is a solution and it is the only solution. For the sake of exercise, let us solve this equation by reducing to RREF.

$$\begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 1 & | & 2 \end{bmatrix} \xrightarrow{R_1 - R_2 \to R_1} \begin{bmatrix} 1 & 0 & | & -1 \\ 0 & 1 & | & 2 \end{bmatrix}$$

and the system corresponding to this augmented matrix is

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

and the solution set is

$$\left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$$

(b)

$$\begin{aligned} 2x + 2y &= 6 \\ x + y &= 1 \end{aligned}$$

Again, one can immediately see that there is no solution to this system. Indeed, if $x + y = 1$, one must have $2x + 2y = 2$ but this would mean $2 = 6$ which is not true. But, let's find the RREF.

$$\begin{bmatrix} 2 & 2 & | & 6 \\ 1 & 1 & | & 1 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1 \to R_1} \begin{bmatrix} 1 & 1 & | & 3 \\ 1 & 1 & | & 1 \end{bmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{bmatrix} 1 & 1 & | & 3 \\ 0 & 0 & | & -2 \end{bmatrix}$$

which gives us $0\ x + 0\ y = 2$ that is $0 = 2$ which is obviously wrong. So, we conclude there is no solution. But here, the last observation is very important. In our augmented matrix, we have

$$0 \quad 0 \quad | \quad -2$$

and this is the key point which shows "no solutions".

(c)

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

Here, we see that the second equation is equal to two times the first equation. Hence, any solution for the first equation is also a solution for the second and vice versa. So, we don't have any information that we already don't know from the first equation. So, the only information about the solution set is given by the first equation and we know that there are infinitely many solutions to this equation, hence to this system. Let's see how the RREF looks like.

$$\begin{bmatrix} 1 & 1 & \mid & 1 \\ 2 & 2 & \mid & 2 \end{bmatrix} \xrightarrow{R_2 - 2R_1 \to R_2} \begin{bmatrix} 1 & 1 & |1 \\ 0 & 0 & |0 \end{bmatrix}$$

and the corresponding system is

$$\begin{aligned} x + y &= 1 \\ 0 &= 0 \end{aligned}$$

We know that $x = 1 - y$ and for any choice of y, we get a different solution. Namely, $(1 - y, y)$. Hence, the solution set looks like

$$\left\{ \begin{bmatrix} 1 - y \\ y \end{bmatrix} \ : \ y \in \mathbb{R} \right\}$$

So, the solution set is parametrized by one variable. There are infinitely many solutions.

18. Here are more words that you should add to your linear algebra vocabulary: We say that a system is **inconsistent** is there is no solution to the system. So, in b) we have an inconsistent system. We say that a system is **consistent** if it is not inconsistent: that is, if there is at least one solution. So, a) and c) are consistent systems.

19. **Example.** Solve the following system and determine if it is consistent or inconsistent.

$$\begin{aligned} x + 3z &= 4 \\ y + 2z &= 5 \end{aligned}$$

The augmented matrix for this system is

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

and it is already in RREF. We see that $y = 5 - 2z$ and $x = 4 - 3z$. So, we can describe the solution set as

$$\left\{ \begin{bmatrix} 4 - 3z \\ 5 - 2z \\ z \end{bmatrix} \ : \ z \in \mathbb{R} \right\}$$

Notice that in this example for any value of z, we get an x and a y value depending on z. We can change z as we wish and this will change x and y. We have a name for it!

20. Suppose that a system is in RREF. If the column corresponding to a variable has a *leading 1*, then that variable is called a **leading variable**. If the column corresponding to a variable does not have a *leading 1*, then that variable is called a **free variable**.

21. The system in the previous examples has two leading variables: x and y. It has one free variable: z. You can see that we have a 1-*dimensional* set of solutions. We did not define this word yet, but it vaguely means the degree of freedom is 1.

22. Suppose that we have a consistent system. We know that there is at least one solution. We also know that there is either one solution or there are infinitely many solutions (Why?). However, in the case of infinitely many solutions, we have a different notion of magnitude as well. We might have different number of free variables/parameters. We might have different degrees of freedom.

23. The **rank** of a system is the number of leading variables. That is, the number of leading 1s in its REF (or RREF).

24. Notice that

    (a) The rank of a system can not be bigger than the number of variables. (This is obvious by definition).

    (b) The rank of a system can not be bigger than the number of equations. (This is not too obvious, but still easy to see if you spend some time by looking at RREF).

    (c) When the ranks is larger, the number of free variables is smaller. And vice versa.

    (d) **Theorem.** If a system of m equations in n variables is consistent and it (or its augmented matrix) has rank r, then

        i. The set of solutions has $n - r$ parameters.

        ii. If $r < n$, then it has infinitely many solutions.

        iii. If $n = r$, then it has a unique solution.