LECTURES 4/5: SYSTEMS OF LINEAR EQUATIONS

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

1. Linear equations

We now switch gears to discuss the topic of solving linear equations, and more interestingly, systems of them.

Definition. A linear equation in the variables $x_1, \dots x_n$ is an equation of the form

$$a_1x_1 + \dots a_nx_n = b$$

with $a_1, \dots a_n, b \in \mathbb{R}$.

There is not much to say about the solution to a single equation. It is always the set of points on an $(n-1)$-dimensional "hyperplane" in $\mathbb{R}^n$, for example it is a point is $n = 1$, a line if $n = 2$, and a plane if $n = 3$. The situation is more interesting if we discuss the solution sets of **systems** of equations. For example, we may want to find all $x_1, \dots x_n$ which satisfy the following m linear equations simultaneously:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots a_{m,n}x_n = b_m. \end{cases}$$

It is somewhat tedious to write the x_i's and the plus signs, etc. at every step of our analysis, as well as to always write the commas in the subscripts of the $a_{i,j}$'s. We will thus use an (augmented) matrix (a matrix is simply a rectangular array of numbers, and augmented just means that we will insert a bar conveniently separating different types of information) the as a bookkeeping tool to represent the system above by

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

One can also think of this geometrically, in that we are trying to identify the intersection points of m hyperplanes. Of course, the algebraic interpretation is more convenient for

Date: October 4, 2016.

working with, and we are interested in finding a simple algorithm which will solve all such systems.

Example. *Consider the system*

$$\begin{cases} x + y + z = 2 \\ 2x - y + z = 6 \\ 3z + y - z = 4. \end{cases}$$

This is represented by the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 6 \\ 3 & 1 & -1 & 4 \end{array}\right),$$

and solving it is equivalent to finding the intersection points between three planes in $\mathbb{R}^3$*. We could already start reducing these equations by hand, but the question is, how can we reduce this to a routine, simple method?*

2. Row reduction

The basic idea is to use what are known as **elementary row operations**. These can be applied to any (augmented or ordinary) matrix, and consist of the following:

(1) Switch any two rows. If we swap rows (i) and (j), we will denote this by $(i) \leftrightarrow (j)$.
(2) Add a multiple of one row to another. If we add $c(j)$ to the row (i), we will denote this operation by $(i) \mapsto (i) + c(j)$.
(3) Multiply a row by a **non-zero** constant. If we multiply row (i) by c, we will denote this operation by $(i) \mapsto c(i)$.

The point is (and you should check!) that if we perform an elementary row operations on an augmented matrix, then it gives a new augmented matrix which corresponds to a system of equations with the same solution set.

Example. *In the example above, we can use elementary row operations to simplify the system as follows:*

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 6 \\ 3 & 1 & -1 & 4 \end{array}\right) \xrightarrow{(2)\mapsto(2)-2(1),\ (3)\mapsto(3)-3(1)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 2 \\ 0 & -2 & -4 & -2 \end{array}\right)$$

$$\xrightarrow{(3)\mapsto-\frac{1}{2}(3)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right) \xrightarrow{(2)\leftrightarrow(3)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & -3 & -1 & 2 \end{array}\right)$$

$$\xrightarrow{(1)\mapsto(1)-(2),\ (3)\mapsto(3)+3(2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 5 & 5 \end{array}\right) \xrightarrow{(3)\mapsto\frac{1}{5}(3)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{(1)\mapsto(1)+(3),\ (2)-2(3)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

From this, we can directly read off the solution to our system as

$$x = 2, \qquad y = -1, \qquad z = 1.$$

Of course, you can go back and plug these numbers into the original system to check that they do indeed solve it.

3. Gauss-Jordan elimination

The ideas behind this example can be extended to a general procedure, called **row reduction**.

Procedure (Gauss-Jordan method)**.** *To solve a system of linear equations:*

(1) *Write down the corresponding augmented matrix.*
(2) *Use elementary row operations to put it in* **reduced row echelon form (RREF)***. This is a matrix which satisfies:*
  (a) *All rows of only* 0*'s are grouped together at the bottom of the matrix.*
  (b) *In every non-zero row, the left-most entry is* 1*. We refer to this entry as a* **pivot** *of the matrix.*
  (c) *Each pivot is in a column with all other values in the column equal to zero.*
  (d) *The pivot in any row is to the left of any pivots below it.*
(3) *Determine the solution set as follows.*
  (a) *If the last non-zero row is of the form* $(0 \ldots 0|1)$*, i.e there is a pivot in the last column, then the system is* **inconsistent***, and there are no solutions.*
  (b) *Otherwise, the system is* **consistent***, there will be two possibilities; there is either one solution or infinitely many solutions. If there is a pivot in each column before the bar* $|$ *(like in the last example), then the last column directly gives the* **unique** *solution. If there isn't a pivot in each "variable column", there will be infinitely many solutions. In this case, call each variable which*

corresponds to a pivot a **pivotal unknown** *and each other variable a* **free unknown**. *The free unknowns can then be assigned arbitrary values, after which we solve for the pivotal variables.*

As always, this is made more clear by looking at explicit examples. We have already seen in the example above how to row reduce and interpret a system with no free unknowns and a unique solution. We now want to solve one example of the other possibilities: when there are infinitely many solutions, and when there are no solutions.

As is often said, mathematics is not a spectator sport. In particular, with things like row reduction, the only way to learn it well is to do enough examples by hand until it feels like second nature.

Example. *We solve the system of equations*

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 + x_5 = 1 \\ -3x_1 - 6x_2 - 2x_3 - x_5 = 0 \\ 2x_1 + 4x_2 + 2x_3 + x_4 + 3x_5 = -3. \end{cases}$$

We first represent this by the matrix

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 1 \\ -3 & -6 & -2 & 0 & -1 & 0 \\ 2 & 4 & 2 & 1 & 3 & -3 \end{array}\right)$$

and row reduce

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 1 \\ -3 & -6 & -2 & 0 & -1 & 0 \\ 2 & 4 & 2 & 1 & 3 & -3 \end{array}\right) \xrightarrow{(2)\mapsto(2)+3(1),\ (3)\mapsto(3)-2(1)} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 2 & 3 \\ 0 & 0 & 0 & -1 & 1 & -5 \end{array}\right)$$

$$\xrightarrow{(1)\mapsto(1)+(3),\ (2)\mapsto(2)+3(3),\ (3)\mapsto-(3)} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 0 & 2 & -4 \\ 0 & 0 & 1 & 0 & 5 & -12 \\ 0 & 0 & 0 & 1 & -1 & 5 \end{array}\right)$$

$$\xrightarrow{(1)\mapsto(1)-(2)} \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & -3 & 8 \\ 0 & 0 & 1 & 0 & 5 & -12 \\ 0 & 0 & 0 & 1 & -1 & 5 \end{array}\right).$$

We have now reached the reduced row echelon form. There are pivots in the first, third, and fourth columns, and so x_1, x_3, x_4 are our pivotal variables, and x_2 and x_5 are both free. Thus, we set $x_2 = t_2$, $x_5 = t_5$ where t_2 and t_5 are arbitrary real numbers. We now solve for the pivotal variables, yielding

$$x_1 + 2t_2 - 3t_5 = 8, \qquad x_3 + 5t_5 = -12, \qquad x_4 - t_5 = 5.$$

Thus, the final solution set is

$$\begin{cases} x_1 = 3t_5 - 2t_2 + 8 \\ x_2 = t_2 \\ x_3 = -12 - 5t_5 \\ x_4 = 5 + t_5 \\ x_5 = t_5. \end{cases}$$

These are parametric equations for a plane in five-dimensional space.

Example. *We solve the system*

$$\begin{cases} x + 2y + z = 2 \\ 2x + y + 2z = 1 \\ 5x + 4y + 5z = 2. \end{cases}$$

This corresponds to the matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 5 & 4 & 5 & 2 \end{array}\right),$$

which reduces as

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 5 & 4 & 5 & 2 \end{array}\right) \xrightarrow{(2)\mapsto(2)-2(1) \ \ (3)\mapsto(3)-5(1)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & 0 & -3 \\ 0 & -6 & 0 & -8 \end{array}\right)$$
$$\xrightarrow{(3)\mapsto(3)-2(2)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & -3 & 0 & -3 \\ 0 & 0 & 0 & -2 \end{array}\right).$$

We now stop, as the last row corresponds to the equation $0x + 0y + 0z = 0 = -2$, which is clearly nonsense. Hence, there are no solutions.

4. Linear combinations

An important concept in linear algebra is the following.

Definition. A vector v is a **linear combination** of the vectors $v_1, \ldots, v_n$ if there are constants $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$v = c_1 v_1 + \ldots c_n v_n.$$

Given a collection of vectors $v_1, \ldots v_n$, the **span** of them, written $\operatorname{span}(v_1, \ldots, v_n)$, is the set of all linear combinations of the vectors $v_1, \ldots, v_n$.

We now require a second way to think about vectors. Instead of thinking of vectors as tuples such as $u = (u_1, \dots, u_n)$, we can think of them as $n \times 1$ matrices such as

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Returning to systems of linear equations, we can write a system like

$$\begin{cases} x - y + 3z = 2 \\ x + z = 0 \\ y - z = 1 \end{cases}$$

using vectors as

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

In general, a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots a_{m,n}x_n = b_m. \end{cases}$$

has a solution (i.e., is consistent) if and only if

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \operatorname{span}(a_1, a_2, \dots a_n),$$

where $a_1, \dots a_n$ are the m-dimensional **column vectors** of the corresponding matrix:

$$a_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}.$$

Example. *We will determine whether* $v = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ *is a linear combination of*

$$v_1 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 6 \\ 5 \\ 5 \end{pmatrix}.$$

This is the same as asking whether there is a solution to the equation

$$x_1v_1 + x_2v_2 + x_3v_3 = v \iff \begin{cases} 3x_1 + x_2 + 6x_3 = 2 \\ -x_1 + 2x_2 + 5x_3 = 1 \\ 2x_1 + x_2 + 5x_3 = 1. \end{cases}$$

We row reduce this as follows:

$$\left(\begin{array}{ccc|c} 3 & 1 & 6 & 2 \\ -1 & 2 & 5 & 1 \\ 2 & 1 & 5 & 1 \end{array}\right) \xrightarrow{(1)\leftrightarrow(2),\ (1)\mapsto-(1)} \left(\begin{array}{ccc|c} 1 & -2 & -5 & -1 \\ 3 & 1 & 6 & 2 \\ 2 & 1 & 5 & 1 \end{array}\right)$$

$$\xrightarrow{(2)\mapsto(2)-3(1),\ (3)\mapsto(3)-2(1)} \left(\begin{array}{ccc|c} 1 & -2 & -5 & -1 \\ 0 & 7 & 21 & 5 \\ 0 & 5 & 15 & 3 \end{array}\right) \xrightarrow{(2)\mapsto\frac{1}{7}(2)} \left(\begin{array}{ccc|c} 1 & -2 & -5 & -1 \\ 0 & 1 & 3 & \frac{5}{7} \\ 0 & 5 & 15 & 3 \end{array}\right)$$

$$\xrightarrow{(3)\mapsto(3)-5(2)} \left(\begin{array}{ccc|c} 1 & -2 & -5 & -1 \\ 0 & 1 & 3 & \frac{5}{7} \\ 0 & 0 & 0 & 3-\frac{25}{7} \end{array}\right).$$

Looking at the last row, we don't need to finish finding the RREF, as we already see that the system is inconsistent. Hence, v is not a linear combination of v_1, v_2, v_3.

Finally, there is one special case of spans of vectors which we will be particularly interested in throughout this course.

Definition. Given any matrix A, the **column space** is the span of its column vectors.

Example. *Here we find the column space of*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}.$$

The column vectors of A are $v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. A generic vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ is in the column space if and only if there are constants c_1, c_2 such that

$$c_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \iff \begin{cases} c_1 = x \\ c_2 = y \\ 2c_2 = z, \end{cases}$$

which corresponds to the matrix

$$\left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \\ 2 & 0 & z \end{array}\right).$$

Row reducing this matrix gives

$$\left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & z-2x \end{array}\right).$$

Now this system has a solution if and only if $z - 2x = 0$*, or* $z = 2x$*. Thus, the column space, which is the space of linear combinations of* v_1, v_2 *in* $\mathbb{R}^3$*, is the plane* $z = 2x$*.*