GEOMETRY OF SYSTEMS OF LINEAR EQUATIONS

Consider a linear equation $a_1x_2 + \cdots + a_nx_n = b$. The solutions to this systems are vectors $\mathbf{x}$ in $\mathbb{R}^n$. Collecting the coefficients a_i into another vector in $\mathbb{R}^n$, we can write this equation as $\mathbf{a} \cdot \mathbf{x} = b$. Thus the solutions to the given equation lie in a subset of $\mathbb{R}^n$, called a *hyperplane*, for which $\mathbf{a}$ is a normal vector. Does the scalar b have an equally geometric meaning?

Example. Let $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $b = 6$. Writing out the equation $\mathbf{a} \cdot \mathbf{x} = b$ in coordinates, we get $x_1 + x_2 = 6$. The solution set is the line through $\binom{6}{0}$ and $\binom{0}{6}$. The difference between any two points in the solution set is a vector in $\mathbf{a}^\perp$. The line $\{t\mathbf{a} \mid t \in \mathbb{R}\}$ is perpendicular to the solution set, and intersects it at $\binom{3}{3}$. This is the closest point to the origin, and its distance to $\mathbf{0}$ is $3\sqrt{2}$. Where does this appear in the equation? It's b divided by the length of $\mathbf{a}$. So if we "scale" the equation to replace $\mathbf{a}$ with a unit vector, we get a constant term that measures the distance from the line to the origin.

One of the homework exercises asks you to examine the general case of the above example in $\mathbb{R}^n$.

It should come as no surprise at this point that solving systems of linear equations is deeply connected to the geometry of $\mathbb{R}^n$. Define

$$H_{\mathbf{a},b} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \cdot \mathbf{x} = b\}.$$

This is called a *hyperplane* in $\mathbb{R}^n$. A system of linear equations looks like

$$\begin{cases} \mathbf{a}_1 \cdot \mathbf{x} = b_1 \\ \mathbf{a}_2 \cdot \mathbf{x} = b_2 \\ \vdots \\ \mathbf{a}_k \cdot \mathbf{x} = b_k \end{cases}.$$

Any solution must solve all of the equations *simultaneously*—i.e., it must lie in

$$H_{\mathbf{a}_1,b_1} \cap H_{\mathbf{a}_2,b_2} \cap \cdots \cap H_{\mathbf{a}_k,b_k}.$$

The linear relations between the $\mathbf{a}_i$s can thus tell us something about the size (read: dimension) of the solution set. (Look at picture in textbook: p. 184.)

Proposition 0.1. *If the vectors $\mathbf{a}_i$ are linearly independent, then a solution always exists.*

The geometric reason, loosely speaking, is that the hyperplanes are all "oriented" in different directions, so they must intersect. This corresponds to the case where the coefficient matrix reduces to an echelon form where every row has a pivotal 1.

Proposition 0.2. *If the vectors $\mathbf{a}_i$ are linearly dependent, then at least one of the following is true:*

- *one of the equations is* redundant, *meaning it is a consequence of some of the others;*
- *the equations are* inconsistent, *meaning no solutions exist.*

Both of these situations correspond to having a reduced coefficient matrix that contains a row of zeroes. As you've seen in class, all bets are off in that case; you must study the values of the b_is to know if solutions exist. The easiest examples to think of are where two of the equations determine the same hyperplane (i.e., $\mathbf{a}_i = t\mathbf{a}_j$ and $b_i = tb_j$ for some i, j,

t)—then one of the equations is redundant—or when two of the hyperplanes are *parallel* (i.e., $\mathbf{a}_i = t\mathbf{a}_j$ for some i, j, t, but $b_i \neq tb_j$)—then no vector can possibly be in both.

We can turn this whole picture around by collecting the b_is into a single vector $\mathbf{b} \in \mathbb{R}^k$, and collecting corresponding elements of the a_is into n vectors $\mathbf{v}_j$ in $\mathbb{R}^k$. Thus,

$$\mathbf{v}_1 = \begin{bmatrix} a_{1,1} \\ \vdots \\ a_{k,1} \end{bmatrix}, \ldots, \mathbf{v}_n = \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{k,n} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix},$$

and the system becomes

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n = \mathbf{b},$$

a *single* equation in $\mathbb{R}^k$. (You're still looking for a solution in $\mathbb{R}^n$, however.) You may recall that this was pointed out in class when defining matrix multiplication: each column of the product matrix is a linear combination of the columns of the first factor matrix. Now you're asking: can you "build" the vector $\mathbf{b}$ by a linear combination of the $\mathbf{v}_j$s, i.e., does $\mathbf{b}$ lie in their span? Again, the key to knowing the answer is row reduction.

Why is row reduction so important? The former view of systems of equations gives a partial answer: any equation obtained as a consequence of the given equations must also be true for a solution of the system. That is, *row operations do not change the solution set.* The latter view of systems of equations also gives an answer: because the set of x_i solving the system in $\mathbb{R}^k$ is invariant under row operations on the $\mathbf{v}_j$, *row operations do not change linear relations among the coefficient vectors.* Effectively, row operations are a "change of basis," and row echelon form allows you to see exactly which columns of a matrix can form a basis, and how the remaining columns look in that basis.

Example. Consider the linear map $\mathbb{R}^5 \to \mathbb{R}^3$ defined by the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 0 & -2 & -2 \\ 3 & -3 & 1 & -2 & 1 \\ 1 & -1 & 1 & -1 & 0 \end{bmatrix}.$$

Given $\mathbf{b} \in \mathbb{R}^3$, can we always find $\mathbf{x} \in \mathbb{R}^5$ such that $\mathbf{A}\mathbf{x} = b$? IOW, is this map *onto*? We want to know if the columns of $\mathbf{A}$ span $\mathbb{R}^3$. We guess they might, because there are so many of them, but let's be sure. $\mathbf{A}$ row reduces to

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

Now we see that the 1st, 3rd, and 4th columns form a basis for $\mathbb{R}^3$, and we even know how to express the 2nd and 5th columns in that basis. Thus the given map is onto $\mathbb{R}^3$. In a couple of weeks, it will be very important to know when a linear map is onto.

What does the reduced form of $\mathbf{A}$ tell us about the first perspective on linear systems? Well, it also shows that all three rows are independent, so for any particular vector $\mathbf{b} \in \mathbb{R}^3$, the solution set of $\mathbf{A}\mathbf{x} = \mathbf{b}$ will be the intersection of three non-parallel hyperplanes; i.e., there will be two free parameters. This relates to a theorem you'll see in class tomorrow, on the various dimensions of subspaces associated to a linear map.