

Linear Algebra notes (3/5)

March 5, 2020

1 Linear systems

Linear algebra begins with the study of linear systems of equations. A single linear equation takes a form we are used to — like

$$2x - y + 3z = 3;$$

it defines a plane in 3-space. A system of linear equations is a collection of more than one of these, such as

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

You may be familiar with how to solve these from your high school algebra classes. We are allowed to:

I: Add or subtract any two of these equations from each other; for instance, I may subtract the third equation from the first one, leaving me with

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

II: we can scale any one of the rows by a nonzero constant. So this system of equations is equivalent to, for instance,

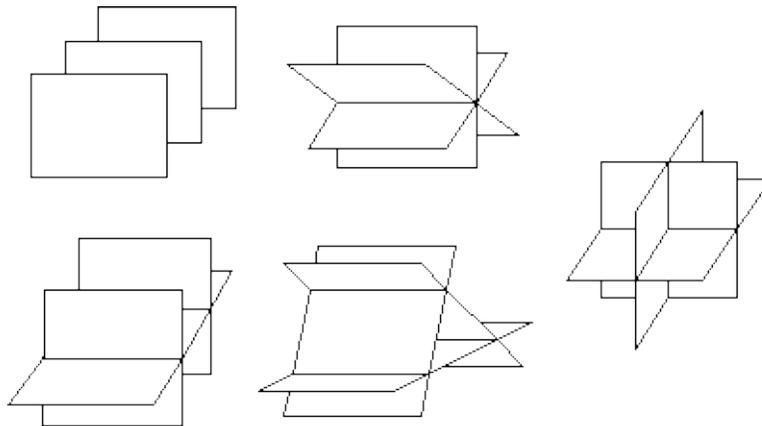
$$\begin{aligned}3y + 2z &= 2 \\2x + y + 3z &= 5 \\2x - 2y + 2z &= 4.\end{aligned}$$

Every linear system of equations can be solved (meaning, simplified as much as possible) using only these two operations. For the sake of space, I won't proceed to do so with this system; but applying these operations five or six more

times, one finds that this system has the unique solution $(x, y, z) = (1, 0, 1)$.

One can proceed in an ad hoc way to solve these systems like we did above, but a more systematic approach was needed, especially for systems with many variables (oftentimes in machine learning or economics you'll have thousands of variables). This systematization is what became linear algebra.

Before moving on, let me point out how one can view these geometrically. In the above case, we had three equations. Each equation represented a plane in space. To demand that all three equations hold means that we are looking for the *intersection* of all three of these planes — the set of points where all of their defining equations are true. In the above case, the intersection was just a point. We can visualize this situation as being similar to the case of the right-most configuration here:



In the first configuration and both bottom configurations, we would find that we actually have no solutions to the system of equations (there are no points lying in all three planes at once). In the top-right configuration, we would have a *line* of solutions.

2 Linear maps and transformations

The first step to working systematically with systems of equations is realizing what's special about them. In the first section, we saw that given any system of equations, we can add two of the equations together or scale the equations without changing the solution set, just like we can add or scale vectors. This ends up being the property we need to focus on.

Definition 1. A linear function is a function $L : \mathbb{R}^n \rightarrow \mathbb{R}$ which satisfies

$$\begin{aligned}L(c\vec{v}) &= cL(\vec{v}) \text{ for any } c \in \mathbb{R} \\L(\vec{v} + \vec{w}) &= L(\vec{v}) + L(\vec{w}).\end{aligned}$$

Example 1. The function $L(x, y, z) = 3x - y + 2z$ is linear. Seeing this is a calculation:

$$\begin{aligned}L(cx, cy, cz) &= 3cx - cy + 2cz \\&= c(3x - y + 2z) \\&= cL(x, y, z),\end{aligned}$$

while

$$\begin{aligned}L(x_1 + x_2, y_1 + y_2, z_1 + z_2) &= 3(x_1 + x_2) - (y_1 + y_2) + 2(z_1 + z_2) \\&= 3x_1 - y_1 + 2z_1 + 3x_2 - y_2 + 2z_2 \\&= L(x_1, y_1, z_1) + L(x_2, y_2, z_2).\end{aligned}$$

So L satisfies both of the requirements to be a linear map; L is linear.

Example 2. The function $f(x, y) = x^2 + y^2$ is not a linear function: for instance, pick the vectors $\vec{v} = \langle 1, 2 \rangle$ and $\vec{w} = \langle -1, 1 \rangle$; these have $\vec{v} + \vec{w} = \langle 0, 3 \rangle$. Now we have

$$f(\vec{v} + \vec{w}) = f(0, 3) = 9 \neq 5 + 2 = f(\vec{v}) + f(\vec{w}).$$

So f is not linear, or else we would have had $f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w})$ for these two vectors \vec{v} and \vec{w} .

Note that to show that a function is linear we have to show that $L(cx, cy, cz) = cL(x, y, z)$ for all c, x, y, z ; and $L(\vec{v} + \vec{w}) = L(\vec{v}) + L(\vec{w})$ for all vectors \vec{v} and \vec{w} . But to show that a function is not linear, we just need to find two vectors \vec{v} and \vec{w} for which $L(\vec{v} + \vec{w}) \neq L(\vec{v}) + L(\vec{w})$, or a c and \vec{v} so that $L(c\vec{v}) \neq cL(\vec{v})$.

In fact, **every** linear function takes a form similar to the one in the first example:

Theorem 1. Every linear function $L : \mathbb{R}^n \rightarrow \mathbb{R}$ is of the form

$$L(x_1, \dots, x_n) = a_1x_1 + \dots + a_nx_n$$

for suitable constants a_1, a_2, \dots, a_n . In particular, every linear function $L : \mathbb{R}^3 \rightarrow \mathbb{R}$ is of the form

$$L(x, y, z) = ax + by + cz$$

for appropriate constants a, b , and c .

Proof. I will explain why this is true in the case of \mathbb{R}^3 ; nothing changes in the general case.

Given a linear function $L : \mathbb{R}^3 \rightarrow \mathbb{R}$, write

$$\begin{aligned}L(1, 0, 0) &= a \\L(0, 1, 0) &= b \\L(0, 0, 1) &= c.\end{aligned}$$

Now we can write $\langle x, y, z \rangle$ as a combination of these vectors $\hat{i}, \hat{j}, \hat{k}$; in fact,

$$\langle x, y, z \rangle = x\langle 1, 0, 0 \rangle + y\langle 0, 1, 0 \rangle + z\langle 0, 0, 1 \rangle.$$

By definition of linear function, we have

$$L(x, y, z) = xL(1, 0, 0) + yL(0, 1, 0) + zL(0, 0, 1),$$

because we can pull out scalars and separate addition for linear functions. But that just means that

$$L(x, y, z) = ax + by + cz,$$

because those constants are just the names for $L(1, 0, 0)$, $L(0, 1, 0)$, and $L(0, 0, 1)$. \square

Linear functions, then, are reasonably simple, and you know them well. (They're exactly the kind of things we talked about in the linear systems section.) More precisely, every linear equation takes the form

$$ax + by + cz = d$$

for some constants a, b, c, d . Then there is a corresponding linear *function*

$$L(x, y, z) = ax + by + cz;$$

solutions to the linear equation are those vectors $\vec{v} = \langle x, y, z \rangle$ with $L(\vec{v}) = d$.

Another way to say this is that solutions to linear equations are the level sets of linear functions; another way yet is that the level sets of linear functions are all parallel planes.

To go from linear equations to linear systems, we introduce the following.

Definition 2. A linear map or linear transformation is a map $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ satisfying

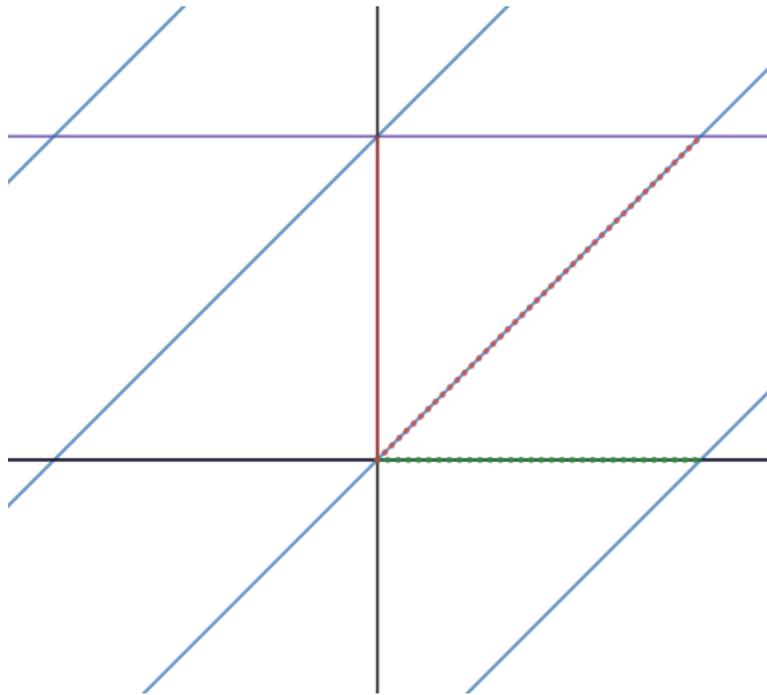
$$\begin{aligned}A(c\vec{v}) &= cA(\vec{v}) \\A(\vec{v} + \vec{w}) &= A(\vec{v}) + A(\vec{w}).\end{aligned}$$

These are the same requirements as for a linear function; the only difference is that we allow the target (codomain) of A to be some higher-dimensional space \mathbb{R}^m . In some sense, these are just as simple as linear functions (more on that next week); in another sense, we get more exotic examples.

Example 3. The transformation of the plane $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ given by

$$A(x, y) = (x + y, y)$$

is called a *shearing map*. We can visualize it as follows:



The purple lines are the images of the lines $y = 0, y = 1$, etc; because $A(x, 0) = (x, 0)$, we see that A fixes the x -axis, and because $A(x, c) = (x + c, c)$, we see that it takes the line $y = c$ and shifts points on it c units to the right.

The green lines are the images of the lines $x = 0, x = 1$, etc. Because $A(0, y) = (y, y)$, we see that the y -axis is sent to the line $x = y$, as shown in the image.

The dotted portions are the image of the vectors $\langle 1, 0 \rangle$ and $\langle 0, 1 \rangle$.

Essentially, this transformation is what you apply when you want to take a letter (like *b*) and make it italic (like *b*).

Example 4. The dot product $\vec{v} \cdot \vec{w}$ is not linear, but it is *linear in each coordinate*

(also called *bilinear*). Linearity in the first coordinate means

$$\begin{aligned}(c\vec{v}) \cdot \vec{w} &= c(\vec{v} \cdot \vec{w}), \\ (\vec{v} + \vec{w}) \cdot \vec{u} &= \vec{v} \cdot \vec{u} + \vec{w} \cdot \vec{u};\end{aligned}$$

linearity in the second coordinate is similar. This property is what makes the dot product so usable, and we use it without even thinking about it in most problems about the dot product

In the exact same way, the cross product $\vec{v} \times \vec{w}$ is a bilinear map.

Example 5. For a more exotic example: the operation "taking the derivative" is linear. After all, given two functions $f(t)$ and $g(t)$, we have

$$\begin{aligned}\frac{d}{dt}(f(t) + g(t)) &= \frac{d}{dt}(f(t)) + \frac{d}{dt}(g(t)) \\ \frac{d}{dt}(cf(t)) &= c \left(\frac{d}{dt}f(t) \right).\end{aligned}$$

This doesn't totally fit into the framework of linear transformations outlined above (it sends functions to functions instead of vectors to vectors), but it goes to show how ubiquitous linear transformations are.

Example 6. The transformation $A(x, y) = (y - x, e^{xy})$ is not linear: we have, for instance,

$$\begin{aligned}A(1, 0) &= (-1, e^0) = (-1, 1) \\ A(0, 1) &= (1, e^0) = (1, 1) \\ A(1, 1) &= (1 - 1, e^1) = (0, e).\end{aligned}$$

But of course $(-1, 1) + (1, 1) = (0, 2) \neq (0, e)$. So $A(x, y)$ is not a linear transformation.

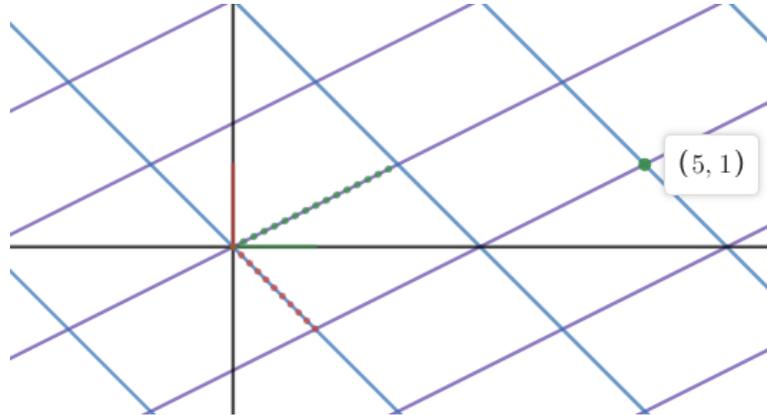
In fact, a transformation A is linear precisely when its coordinate functions (above, $y - x$ and e^{xy}) are all linear functions. Here, $y - x$ was linear, but e^{xy} was not.

Now we move on to talking about linear systems in terms of linear transformations. As an example, consider the linear system

$$\begin{aligned}2x + y &= 5 \\ x - y &= 1.\end{aligned}$$

Corresponding to this linear system is a linear transformation $A(x, y) = (2x + y, x - y)$, from \mathbb{R}^2 to \mathbb{R}^2 . If we set $b = (5, 1)$, then the equation $A(x, y) = b$ is just a very compact way of writing the equations of that linear system above.

Here we give a picture of the above linear transformation A and $b = (5, 1)$. Again, the purple lines are the images of the lines $y = -1, 0, 1,$ and so on (the images of the horizontal lines), while the green lines are the images of $x = -1, 0, 1,$ and so on (the images of the vertical lines). The picture makes clear that $A(2, 1) = (5, 1)$, so $(x, y) = (2, 1)$ is the unique solution to the above system of equations.



In general, a linear system with n variables and m equations gives a linear transformation $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. There is a $\vec{b} \in \mathbb{R}^m$ so that the solutions to the linear system are vectors $\vec{v} \in \mathbb{R}^n$ solving the equation $A(\vec{v}) = \vec{b}$.

For instance, the system

$$\begin{aligned}x + y + z &= 3 \\x - z &= 0\end{aligned}$$

defines a linear map $A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, explicitly $A(x, y, z) = (x + y + z, x - z)$; and solutions are vectors \vec{v} with $A(\vec{v}) = (3, 0)$. In this specific case, that turns out to define a line in 3-space.

As another example, the system

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

defines a linear map $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, given by $A(x, y) = (x - y, x + y, 2x + y)$; solutions to the system of linear equations are vectors $\vec{v} = \langle x, y \rangle$ with $A(\vec{v}) = (3, 1, 2)$. In this case there are no solutions.

Solution sets to systems of linear equations will always look like this; they will be empty, points, lines, or planes.

3 Problems

Determine whether or not the following functions are linear. If they are linear, show it by definition: that is, show that $L(c\vec{v}) = cL(\vec{v})$ for all c and \vec{v} , and that $L(\vec{v} + \vec{w}) = L(\vec{v}) + L(\vec{w})$ for all \vec{v}, \vec{w} . If they are not linear, show it by example (find vectors \vec{v}, \vec{w} so that $L(\vec{v} + \vec{w}) \neq L(\vec{v}) + L(\vec{w})$, or c, \vec{v} so that $L(c\vec{v}) \neq cL(\vec{v})$).

1. $L(x, y) = 2x - 3y$.
2. $L(x, y, z) = xyz$.
3. $L(x, y) = \frac{x^4 + y^4}{x^2 + y^2}$.
4. $L(x, y, z) = x + y + z$.
5. $L(x, y) = x + y + 1$.
6. $L(x, y, z) = xy - x - y + z$.

7. Explain what the linear transformation $A(x, y) = (x + y, x - y)$ does to the plane. Draw a picture, including what happens to the axes $y = 0$ and $x = 0$, as well as a few other vertical and horizontal lines such as $x = 1$ and $y = -1$.

8. Find the solution set to

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

If it is a line, write it in parametric form.

9. Find the solution set to

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

If it is a line, write it in parametric form. How does it relate to the solution set from #8?

10. Find the solution set to

$$x + y + z = 3$$

$$x + 2z = 1$$

$$y - z = 1.$$

If it is a line, write it in parametric form. How does it relate to the solution set from #8?