LINEAR ALGEBRA: LECTURES 17-18

LEO GOLDMAKHER

Last time we proved:

Proposition 1. *Given $\vec{v}_1, \vec{v}_2 \in V^2$, let $f : V^2 \to V^2$ be the linear map such that $f(\vec{e}_1) = \vec{v}_1$ and $f(\vec{e}_2) = \vec{v}_2$. If f is nonsingular, then every vector is a linear combination of $\vec{v}_1, \vec{v}_2$. In other words, given any $\vec{v} \in V^2$, there exist $\alpha, \beta \in \mathbb{R}$ such that*

$$\vec{v} = \alpha\vec{v}_1 + \beta\vec{v}_2.$$

The linear map f in the above proposition turns out to be pretty useful, so we give it a name:

Definition. The *change-of-basis map* (also called the *change-of-basis matrix*) from $\vec{e}_1, \vec{e}_2$ to $\vec{v}_1, \vec{v}_2$ is the linear map $f : V^2 \to V^2$ which sends $\vec{e}_1 \mapsto \vec{v}_1$ and $\vec{e}_2 \mapsto \vec{v}_2$.

The change-of-basis map is very easy to write down in matrix form. If $\vec{v}_1 = p\vec{e}_1 + q\vec{e}_2$ and $\vec{v}_2 = r\vec{e}_1 + s\vec{e}_2$, then the change-of-basis matrix is $\begin{pmatrix} p & r \\ q & s \end{pmatrix}$. (Do not memorize this! Instead, make sure you can explain it.) This in turn makes it very easy to find the expression predicted in the above proposition.

Example 1. Let $\vec{v}_1 := \vec{e}_1 + 2\vec{e}_2$ and $\vec{v}_2 := 3\vec{e}_1 + 4\vec{e}_2$. How can one express $\vec{e}_1 + 6\vec{e}_2$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$? The change-of-basis matrix is $f := \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ and we immediately deduce that

$$f^{-1} = \frac{1}{\det f}\begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix}.$$

Thus,

$$f^{-1}(\vec{e}_1 + 6\vec{e}_2) = \begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix}\begin{pmatrix} 1 \\ 6 \end{pmatrix} = 7\vec{e}_1 - 2\vec{e}_2$$

and it follows that

$$\vec{e}_1 + 6\vec{e}_2 = f(7\vec{e}_1 - 2\vec{e}_2) = 7\vec{v}_1 - 2\vec{v}_2.$$

The above example might seem more complicated than it needs to be, but it demonstrates a purely mechanical approach to expressing any vector $\alpha\vec{e}_1 + \beta\vec{e}_2$ as a linear combination of two other vectors $\vec{v}_1, \vec{v}_2$. Namely:

Step 1. Let f be the change-of-basis matrix. (As explained above, this can be written down immediately without any work.)

Step 2. Compute the matrix of f^{-1}. (We have a formula for this, so it also doesn't require any thought.)

Step 3. Compute $f^{-1}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$; the coordinates tell you how many copies of $\vec{v}_1$ and $\vec{v}_2$ to take.

I encourage you to revisit Example 1, viewing it through the lens of this three-step approach.

Change of basis might seem like a formalism, but in fact it's a ubiquitous process in math and day-to-day life. The underlying idea is simple: if you have to perform some complicated transformation, change your point of view (your 'basis') in such a way that the transformation becomes simple to carry out, then perform your (now simple) transformation, then change your point of view back to your original perspective, and voilà! You've successfully performed your complicated transformation.

You've already encountered (on your midterm) one mathematical example of change of basis: the linear map $\sigma_{\mathcal{L}}$ which reflects points across a line $\mathcal{L}$ through the origin. Finding the matrix of $\sigma_{\mathcal{L}}$ directly is nontrivial, as many of you discovered. Instead, we change our point of view, as follows. First, tilt your head (i.e., rotate

Date: March 16 & 18, 2016.

the plane) until $\mathcal{L}$ become the horizontal axis. Next, reflect across this horizontal line using the simple matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Finally, untilt your head (i.e., rotate the plane back to where it was). Putting this all together, we find that $\sigma_{\mathcal{L}} = R_\alpha \circ \rho \circ R_{-\alpha}$, where α is the angle $\mathcal{L}$ forms with the positive horizontal axis. Thus, our change of basis is the rotation: we use it as a dictionary to translate back and forth between points of view (one in which the reflection is hard to describe, the other in which it's easy).

There are lots of real-world situations in which it's useful to have a dictionary translating between different points of view. One beautiful example, which I learned about from an online posting by Alon Amit, is about older-style showers with two separate taps: one for hot water and one for cold water. The natural way to describe a state of this shower is (h, c), where h is the number of twists of the hot tap and c is the number of twists of the cold. These are the two parameters we control, but the parameters we're actually interested in are pressure and temperature. As we discussed in class, pressure is related to $h + c$, while temperature is related to $h - c$. Thus, when we fiddle with the hot and cold taps, we're secretly applying a change of basis in our heads to translate our fiddling into desired changes to pressure and temperature. You'll explore this example more on your homework.

Given two vectors $\vec{v}_1, \vec{v}_2 \in V^2$, let f denote the change-of-basis map from $\vec{e}_1, \vec{e}_2$ to $\vec{v}_1, \vec{v}_2$. Proposition 1 asserts that if f is nonsingular, then every vector can be expressed as a linear combination of $\vec{v}_1, \vec{v}_2$. What if f is singular?

Proposition 2. *Let $f : V^2 \to V^2$ be the change-of-basis map from $\vec{e}_1, \vec{e}_2$ to $\vec{v}_1, \vec{v}_2$. If f is singular, then there exists $\vec{v} \in V^2$ which* cannot *be expressed as a linear combination of $\vec{v}_1$ and $\vec{v}_2$.*

Proof. We abuse notation and write a vector $p\vec{e}_1 + q\vec{e}_2$ in the form $\begin{pmatrix} p \\ q \end{pmatrix}$; this is OK, since we know that any linear map treats points and vectors in an indistinguishable way. Suppose f is singular. Then problem **3.6** tells us that the image of f (viewed as a subset of $\mathbb{R}^2$) lives on some line. In particular, there must exist some point $\begin{pmatrix} r \\ s \end{pmatrix} \notin \operatorname{im}(f)$. In particular, for any $\alpha, \beta \in \mathbb{R}$, we have $\begin{pmatrix} r \\ s \end{pmatrix} \neq f(\alpha\vec{e}_1 + \beta\vec{e}_2)$. But this immediately implies that $\begin{pmatrix} r \\ s \end{pmatrix} \neq \alpha\vec{v}_1 + \beta\vec{v}_2$ for any $\alpha, \beta \in \mathbb{R}$, as claimed. $\square$

1. Singular Value Decomposition

Next, we moved on to an important topic within linear algebra: the *Singular Value Decomposition* of a linear map. This is a fundamental tool in numerous applications, particularly in analyzing big data sets (Principal Component Analysis), in statistics (best least squares fit), in signal processing, and in weather prediction. We will develop a baby case of SVD which won't tell us much about these applications, but is interesting in its own right – it will give us a good geometric intuition for what a linear map does to the plane.

Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a nonsingular linear map. Earlier, we discussed what the image of the unit square looks like. What about the image of the unit circle? i.e., if we set

$$U := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

what does $f(U)$ look like? Let's reverse engineer. Pick some point $(x, y) \in f(U)$; by definition, we have $f^{-1}(x, y) \in U$. To make this more concrete, let's write $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and set

$$\Delta := \det f.$$

Since f is nonsingular, $\Delta \neq 0$. We have

$$f^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

whence

$$f^{-1}(x, y) = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \left(\frac{dx - by}{\Delta}, \frac{-cx + ay}{\Delta} \right).$$

Since $f^{-1}(x,y)$ is supposed to be a point on the unit circle, we deduce that

$$\left(\frac{dx-by}{\Delta}\right)^2+\left(\frac{-cx+ay}{\Delta}\right)^2=1$$

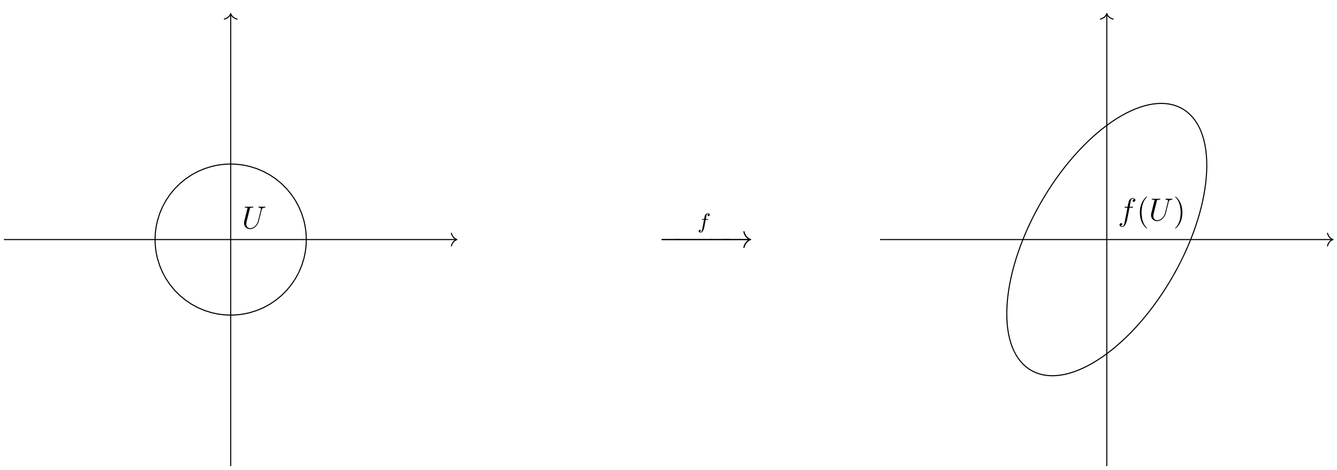
Simplifying this gives

$$\frac{c^2+d^2}{\Delta^2}x^2-\frac{2(ac+bd)}{\Delta^2}xy+\frac{a^2+b^2}{\Delta^2}y^2=1 \tag{1}$$

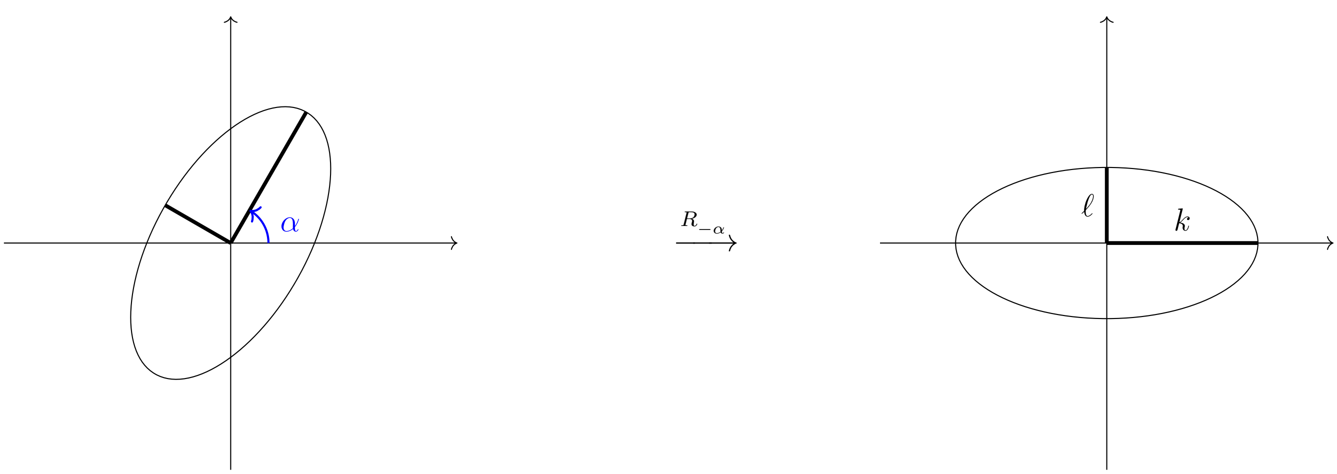
Moreover, it's a straightforward exercise to verify that any point $(x,y)\in\mathbb{R}^2$ satisfying (1) must live in the image of f. (Make sure you can do this.) Thus, we conclude $\exists A,B,C\in\mathbb{R}$, with A and C positive, such that

$$f(U)=\{(x,y):Ax^2+Bxy+Cy^2=1\}.$$

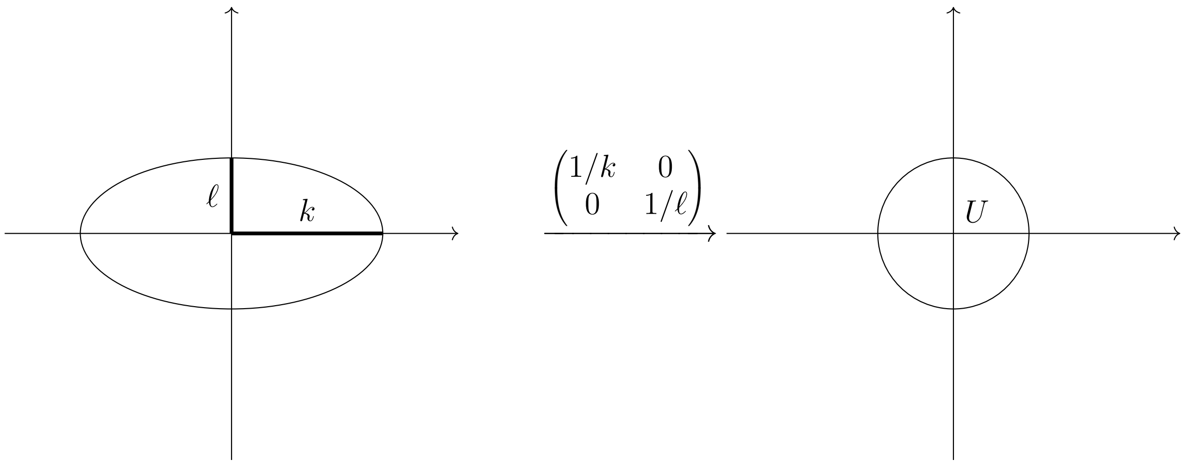
This set of points is a (possibly tilted) ellipse:



This already tells us something about how f acts geometrically, but we can be more precise by applying a few simple transformations renormalize the ellipse. First, let's say the ellipse is tilted from horizontal by some angle α counterclockwise. Then rotating $f(U)$ by $-\alpha$ yields an ellipse whose major and minor axes line up on the x- and y-axes:



Finally, we can simplify this ellipse by applying the diagonal matrix which rescales it:



Putting these three maps together, we see that the linear map $\begin{pmatrix} 1/k & 0 \\ 0 & 1/\ell \end{pmatrix} \circ R_{-\alpha} \circ f$ sends the unit circle to itself. What can we deduce from this? Quite a lot, it turns out:

Lemma 3. *Suppose $g : \mathbb{R}^2 \to \mathbb{R}^2$ is linear and satisfies $g(U) = U$. If $\det g > 0$, then $g = R_\theta$ for some θ, while if $\det g < 0$, then $g = R_\theta \circ \rho$ for some θ.*

Remark. If $\det g = 0$, then all the outputs of g live on some line, so if $g(U) = U$ then g must be nonsingular.

You will prove this lemma on your next assignment. For the moment, let's take it on faith and run with it. For simplicity, we assume that $\det f > 0$; then

$$\det \left(\begin{pmatrix} 1/k & 0 \\ 0 & 1/\ell \end{pmatrix} \circ R_{-\alpha} \circ f \right) > 0$$

as well. (On your next problem set, you will explore what happens if $\det f < 0$.) Applying the lemma to

$$g := \begin{pmatrix} 1/k & 0 \\ 0 & 1/\ell \end{pmatrix} \circ R_{-\alpha} \circ f$$

yields the existence of some angle β such that

$$\begin{pmatrix} 1/k & 0 \\ 0 & 1/\ell \end{pmatrix} \circ R_{-\alpha} \circ f = R_\beta.$$

It follows that

$$f = R_\alpha \circ \begin{pmatrix} k & 0 \\ 0 & \ell \end{pmatrix} \circ R_\beta.$$

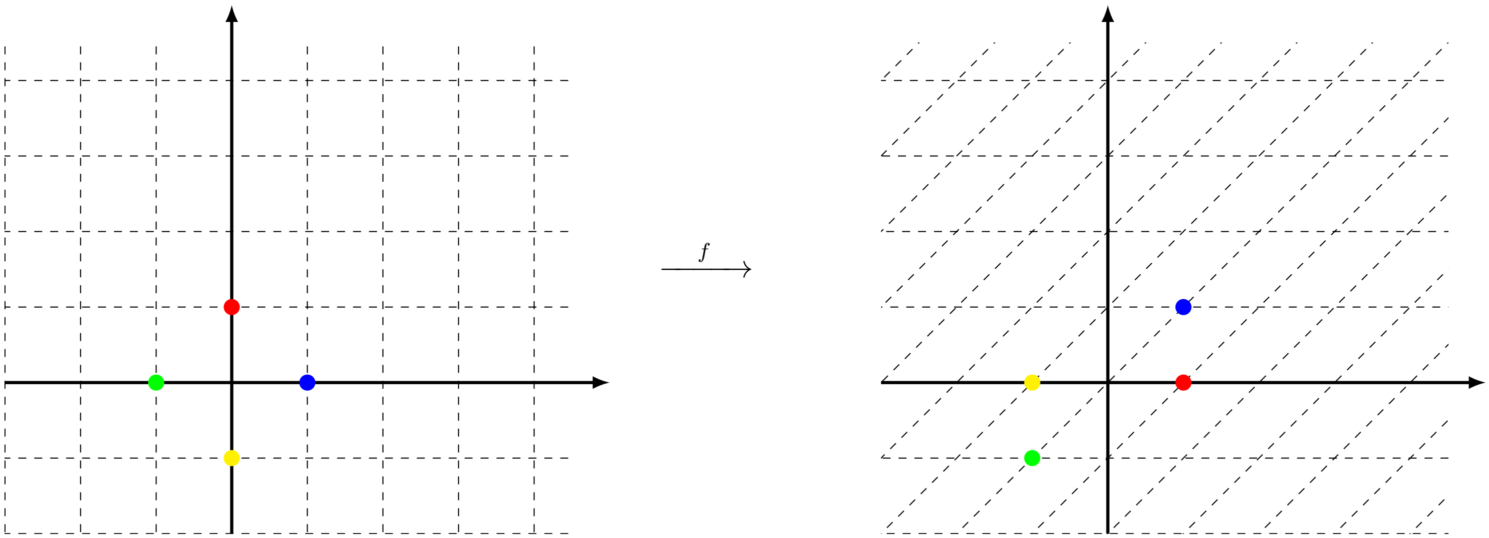
(Make sure you can justify this step!) To summarize, we've proved:

Theorem 4. *Suppose $f : \mathbb{R}^2 \to \mathbb{R}^2$ satisfies $\det f > 0$. Then $\exists \alpha, \beta, k, \ell \in \mathbb{R}$ such that*

$$f = R_\alpha \circ \begin{pmatrix} k & 0 \\ 0 & \ell \end{pmatrix} \circ R_\beta \tag{2}$$

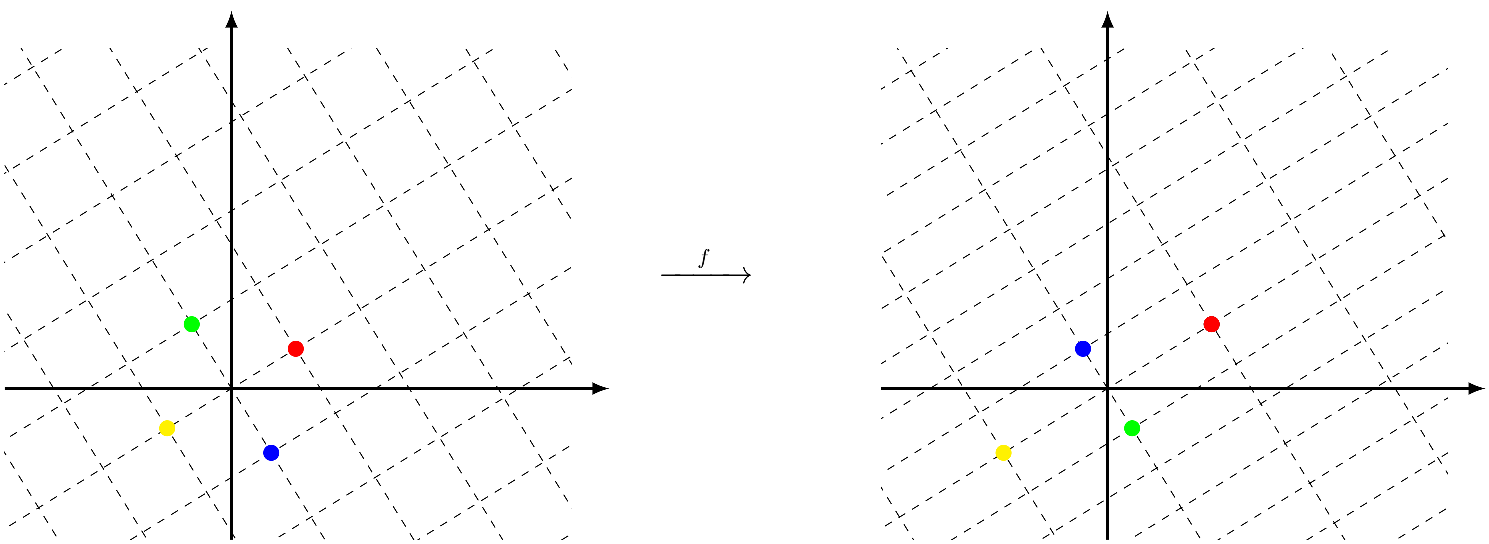
This gives an easy-to-understand geometric description of a linear map: it's a rotation, followed by some nice rescaling in the horizontal and vertical directions, followed by another rotation. The representation (2) is called the *Singular Value Decomposition* of f. Theorem 4 shows that every linear f with positive determinant admits a singular value decomposition, but more is true; on your homework you'll prove that *every* nonsingular linear map admits an SVD.

Next, we view the SVD from a different perspective. A nonsingular linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ sends the square grid (the way we usually view $\mathbb{R}^2$) to some stretched out ('*shear*') grid:



The action of a nonsingular linear map on the rectangular lattice

Rectangular grids are much more intuitive and pleasant to work with than non-rectangular ones, largely because motions (or forces) in perpendicular directions are independent of one another.[1] Thus, one might ask whether it's possible to find a rectangular grid which gets mapped by f to a rectangular grid. The answer turns out to be yes! For example, using the same f as depicted above:



The image of the rectangular grid on the left is also a rectangular grid

What does this have to do with SVD? Given a nonsingular linear map f, consider the ellipse $f(U)$. We can give a complete description of this ellipse using three parameters: the length of the major radius, the length of the minor radius, and the tilt. Note that finding the major and minor radii is a calculus problem – one merely needs to find those points in the image of U which are at a maximal or minimal distance from the origin. The major and minor radii are perpendicular to one another, hence define a rectangular grid. As we shall see, the inverse images of the major and minor radii are also perpendicular to one another. This gives us a rectangular (in fact, a square) grid which gets mapped to a rectangular one. In our above notation, the major radius has length k and the minor radius has length ℓ. These have a special name: they are the *singular values* of the map f. The singular values tell you something about how the map f stretches out the plane. We'll take this up next time.

[1]For example, a bullet fired parallel to the ground will take as long to hit the ground as a bullet dropped from the same height.