Chapter 3

Ben Zelditch

August 26, 2017

Eigenvalues and Eigenvectors

In the last chapter we learned all about linear transformations from one vector space to another. Now we're going to focus our attention on the study of linear operators, which are linear transformations from a vector space to itself. The deepest results in linear algebra are related to linear operators, so this chapter is going to be a good one. Stay tuned.

Invariant Subspaces

Before we can think about eigenvalues and eigenvectors, we need a refresher on linear operators. Remember from last chapter that a linear operator $T \in \mathcal{L}(V)$ on a vector space V is a linear map from V to itself. Now imagine that T is some particularly nasty linear map and is complicated to visualize because V is a high dimensional vector space. Remember from Chapter 1 that we could decompose

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$$

where $\oplus$ is the direct sum. If we could understand how T behaves on each subspace U_j individually, then we would understand how it behaves on V by linearity. Let $T|_{U_j}$ denote the *restriction* of T to the subspace U_j, i.e. we're considering U_j to be the domain of T instead of V. Since U_j is smaller than V it makes sense to think that it would easier to deal with T on U_j than on V.

But what if $T|_{U_j}$ doesn't map to U_j to itself, i.e. what if there exists some $u \in U_j$ such that $Tu \notin U_j$? In other words, what if $T|_{U_j}$ is not an operator on U_j?

Well that would screw things up. Instead we'll only consider decompositions of V such that $T|_{U_j}$ is an operator on U_j. This has a name: we say that a subspace $U \subset V$ is *invariant* under a linear operator T if for all $u \in U$ we have $Tu \in U$, equivalently if $T|_U$ is an operator on U. So from now on, we'll only consider direct sum decompositions of invariant subspaces.

Let's drive this point home by looking at a few examples of invariant subspaces. Let $T \in \mathcal{L}(V)$. Then $\{0\}$ and V itself are invariant subspaces (the so called trivial cases). Additionally, $\operatorname{null} T$ and $\operatorname{range} T$ are both invariant under T (the proofs for both cases are easy). For a more concrete example, suppose that $T \in \mathcal{L}(\mathbb{R}^2)$ is given by the map

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 3x_1 \\ x_2 \end{pmatrix}$$

Then the x-axis is invariant under T because if $(x_1, 0)^T$ is a point on the x-axis then $T((x_1, 0)^T) = (3x_1, 0)^T$ is also a point on the x-axis.

That's about it for invariant subspaces. We'll see why they're so important in the next few sections.

Eigenvalues and Eigenvectors

Now that we're pretty familiar with invariant subspaces, let's look at a special case of them: one-dimensional invariant subspaces. Let's first define the one-dimensional subspace. Let V be a vector space, say $V = \mathbb{R}^n$, and let $u \in V$. Then the set

$$U = \{\alpha u : \alpha \in \mathbb{R}\}$$

is a one-dimensional subspace of V. In fact, all one-dimensional subspaces are of this form.

Now let $T \in \mathcal{L}(V)$. The subspace U is invariant under T if for all $u \in U$ we have $Tu \in U$, in other words if

$$Tu = \lambda u$$

for some $\lambda \in \mathbb{R}$. If this is the case then we say that λ is an *eigenvalue* of T and that u is the corresponding *eigenvector*. Observe that if u is an eigenvector of T corresponding to the eigenvalue λ then any scalar multiple αu of u is also an eigenvector of T with the eigenvalue $\alpha\lambda$ because $T(\alpha u) = \alpha(Tu) = \alpha\lambda u$. This is another way of looking at the invariance of U.

Another way of looking at eigenvalues and eigenvectors is this: the equation $Tu = \lambda u$ is equivalent to $(T - \lambda I)u = 0$, where I is the identity. This second equation is telling us that u is an eigenvector of T precisely when u is in the null space (or kernel) of $T - \lambda I$. In other words, the set of eigenvectors of T corresponding to λ equals the null space of $T - \lambda I$, written $\text{null}(T - \lambda I)$. Remember from last chapter that this is a subspace.

This is a revelation. It means that we can determine whether or not λ is an eigenvalue of T by checking if $T - \lambda I$ has a non-trivial kernel, i.e. if $\dim \text{null}(T - \lambda I) > 0$. Equivalently, we can check if $T - \lambda I$ is not injective and therefore not invertible. If $T - \lambda I$ is not invertible then we say that $T - \lambda I$ is *singular*. More on singular transformations later.

Let's look at an example. Suppose that $T \in \mathcal{L}(\mathbb{R}^2)$ is the linear map that we looked at above, with $T((x_1, x_2)^T) = (3x_1, x_2)^T$. We want to determine if T has any eigenvalues and eigenvectors. To do this, let's go back to the definition. Suppose T does have an eigenvalue and call it λ. Then we have

$$\begin{pmatrix} 3x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

for some non-zero vector, i.e. x_1 and x_2 aren't both zero. In the first coordinate we have $3x_1 = \lambda x_1$ and in the second coordinate we have $x_2 = \lambda x_2$. Clearly there's no value of λ that satisfies both of these equations simultaneously for all values of x_1, x_2. But remember that we can choose one of x_1, x_2 to be zero (but not both) and go from there. Let's choose $x_2 = 0$. Then the second equation holds no matter what λ is and the first equation holds for all x_1 when $\lambda = 3$. Therefore *any* non-zero vector of the form $(x_1, 0)^T$ is an eigenvector of T corresponding to the eigenvalue 3. In other words,

$$\text{null}(T - 3I) = \{(x_1, 0)^T : x_1 \in \mathbb{R}\}.$$

But we also could have chosen $x_1 = 0$. In this case the first equation equation holds no matter what λ is and the second equation holds for all x_2 when $\lambda = 1$. Therefore any non-zero vector of the form $(0, x_2)^T$ is an eigenvector of T corresponding to the eigenvalue 1, i.e.

$$\text{null}(T - I) = \{(0, x_2)^T : x_2 \in \mathbb{R}\}.$$

These are the only eigenvalues of T and their only corresponding sets of eigenvectors. Observe that both $\text{null}(T - 3I)$ and $\text{null}(T - I)$ are invariant

subspaces of dimension one as promised.

Here are a few interesting theorems that follow from what we've learned so far:

1. Let $T \in \mathcal{L}(V)$. Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct eigenvalues of T and suppose that $v_1, v_2, \ldots, v_m$ are corresponding non-zero eigenvectors. Then $v_1, v_2, \ldots, v_m$ are linearly independent.

2. Each operator on V has at most $\dim V$ distinct eigenvalues.

One last thing: As we've seen from this example, while a linear map T may have a few eigenvalues, it always has infinitely many eigenvectors corresponding to its eigenvalues. Therefore it's usually convenient to pick the simplest possible eigenvector corresponding to a particular eigenvalue. For example, in the case we just looked at, we would probably say that $(1, 0)^T$ is the eigenvector corresponding to $\lambda = 3$ and $(0, 1)^T$ is the eigenvector corresponding to $\lambda = 1$. Strictly speaking any scalar multiple of those vectors is also an eigenvector, but picking the simplest one makes our lives easier.

Upper Triangular Matrices

Let's take a step back. This chapter has been dedicated to the study of linear operators and, in particular, the study of eigenvalues and eigenvectors. So far we've been able to compute concrete examples of eigenvalues and eigenvectors only by reasoning through a system of equations, like we did in the example right above. This method generally won't work for more complex transformations because the system of equations almost always becomes intractable. To work with more complex transformations, we need matrices.

Remember from last chapter that any linear transformation $T \in \mathcal{L}(V, W)$ corresponds to a rectangular matrix, so long as we choose a basis for V and for W. Now that we're working with linear operators, all of the matrices that we'll be working with will be square so we only need to choose one basis. Also remember that a fixed linear operator $T \in \mathcal{L}(V)$ can have many different matrix representations depending on the choice of basis for V. A central goal of linear algebra is to show that, given a linear operator on V, we can pick a basis of V that makes the matrix corresponding to the linear operator as simple as possible.

What does it mean for a matrix to be "as simple as possible"? One natural way to define it might be to say that a matrix is simple if it has many zeros. But this is still a bit vague because having zeros scattered all throughout a matrix doesn't necessarily make it that much simpler to work with. On the other hand, if a matrix looked like this:

$$\begin{pmatrix} a_{11} & * & * \\ 0 & a_{22} & * \\ \vdots & 0 & * \\ 0 & \cdots & a_{nn} \end{pmatrix}$$

then that would be much nicer. Note that the stars above the diagonal indicate any value and the entries below the diagonal are all zero. This is called an *upper triangular* matrix.

Luckily we can get any matrix to look like this! This brings us to one of the major theorems of this chapter. It goes: Let V be a complex vector space and let $T \in \mathcal{L}(V)$. Then T has an upper triangular matrix with respect to some basis of V.

This is a pretty powerful statement and it is more subtle than it looks at first glance. The proof of it relies on the fact that any operator on a complex vector space has at least one eigenvalue, which we won't prove here. Using this fact, one can prove the theorem by induction on the size of the dimension of V. Note that the theorem says that V is a *complex* vector space, so we're working over the field $\mathbb{C}$ instead of $\mathbb{R}$ now. This is an important distinction because it is *not* true in general that an operator over a real vector space has a real eigenvalue. The vast majority of them do not.

Diagonal Matrices