LECTURE 9 - EIGENVALUES

Having spent some time looking at the general properties of vector spaces and the generic interplay with linear transformations, we turn our attention now to best understanding a given linear operator on a given vector space. Thus we will assume we are given a finite dimensional vector space V and an operator $L\colon V \to V$. We do not assume we are given a basis! In fact, we'll spend a good bit of time finding a basis.

That being said, we'll begin with some results that will appear as homework problems that help us test whether a linear transformation has a kernel based on a matrix for it. Let $\mathcal{B}$ be any basis for V, and let $A = {}_{\mathcal{B}}[L]_{\mathcal{B}}$.

Defintion. *Let Σ_n denote the set of all bijections of $\{1,\dots,n\}$ with itself. This acts on $\{1,\dots,n\}$, and under composition, this is a group. For an $n \times n$-matrix A, we define the* <u>*determinant*</u> *to be*

$$\det(A) = |A| = \sum_{\sigma \in \Sigma_n} (-1)^{sgn(\sigma)} a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \in \mathbb{F},$$

where $sgn(\sigma)$ is the number of pairs of "out of order" elements in $\sigma((1,\dots,n))$.

This definition is for us now, horribly clunky. We will spend more time talking about determinants in general, and then we will see that this is the most natural definition. Instead, let's look at a standard inductive one.

Defintion. *For a matrix A and for a pair (i,j) indexing a position in A, we can define the minor associated to (i,j), $M_{i,j}(A)$, by deleting the i^{th} row and j^{th} column (resulting in an $(n-1) \times (n-1)$-matrix). We define a new matrix $cof(A)$ (the matrix of cofactors) by*

$$cof(A)_{i,j} = (-1)^{i+j} \det(M_{i,j}).$$

Thus while the determinant is very hard to apply, the matrix of cofactors is the matrix we get by taking determinants of smaller matrices.

Proposition. *For any matrix A,*

$$\det(A) = \sum_{i=1}^{n} a_{i,j} cof(A)_{i,j},$$

for any fixed j, and also

$$\det(A) = \sum_{j=1}^{n} a_{i,j} cof(A)_{i,j},$$

for any fixed i.

This is the usual "expansion by minors": pick a row or column and march along it, scaling the determinant of the associated minors by the deleted element. Though we normally do this over $\mathbb{R}$ or $\mathbb{C}$, it in fact works over any field. Sketching this out is a homework problem. The same holds for the following.

Proposition. *Let $adj(A)$ denote the transpose of $cof(A)$. Then*

$$\det(A) \cdot I_n = A \cdot adj(A) = adj(A) \cdot A.$$

Thus just like for matrices over $\mathbb{R}$ or $\mathbb{C}$, the determinant detects invertibility. If $\det(A)$ is a unit (so in a field, non-zero), then we have an explicit formula for the inverse of a matrix A: $(1/|A|) \cdot adj(A)$. This is actually beautiful for a very different reason that touches on some very deep mathematics. The entries in the $(i,j)^{\text{th}}$ position of A^{-1} are rational functions in the coefficients of A! Finally, as always, determinant has the usual properties:

Theorem. *If A and B are $n \times n$-matrices, then*

$$\det(A \cdot B) = \det(A)\det(B).$$

Back to the matter at hand, we learn the following.

Theorem. *Let A be the matrix associated to L for any basis. The following are equivalent.*

(1) $\det(A) \neq 0$
(2) *L is an isomorphism*
(3) $\ker(L) = \{\vec{0}\}$
(4) $Im(L) = V$.

Here the finiteness of the dimension of V was extra essential.

Proof. We saw above that $\det(A) \neq 0$ is the same as A being invertible. The inverse to A is the matrix for the inverse to L, so this is the same as L being invertible. Since we have a finite dimensional vector space, the rank-nullity theorem establishes the equivalence of points 2, 3, and 4. □

We'll be most interested in looking at particular linear maps that have a kernel.

Defintion. *A vector $\vec{v} \in V$ is an eigenvector for L if $L(\vec{v}) = \lambda \cdot \vec{v}$ for some $\lambda \in \mathbb{F}$. The number λ is call the eigenvalue associated to $\vec{v}$.*

It's immediately clear why we would like such a thing: if $\vec{v}$ is an eigenvector for L, then we know *exactly* how to apply L to any multiple of $\vec{v}$: we just scale. That's it! These vectors are just one tiny step more complicated than elements in the kernel, but not a lot more.

In fact, these are vectors in the kernel of a different operator.

Proposition. *The set of eigenvectors associated to an eigenvalue λ is*

$$V_\lambda = \ker(L - \lambda I).$$

Thus the collection of all eigenvectors associated to an eigenvalue forms a subspace.

Proof. If $L(\vec{v}) = \lambda \cdot \vec{v} = \lambda I(\vec{v})$, then

$$(L - \lambda I)(\vec{v}) = L(\vec{v}) - \lambda I(\vec{v}) = \vec{0}.$$

The reverse inclusion follows by reversing the steps. □

Thus eigenvectors of L with eigenvalue λ are the same thing as vectors in the kernel of $L - \lambda I$. Our work with matrices tells us about how to find eigenvalues then. Using our randomly chosen basis $\mathcal{B}$, we switch to matrices and continue the story there.

Defintion. *Let A be an $n \times n$ matrix. Then $\vec{v} \in \mathbb{F}^n$ is an eigenvector of A with eigenvalue λ if*

$$A \cdot \vec{v} = \lambda \vec{v}.$$

We again know this means that $\vec{v} \in \ker(A - \lambda I)$, and coupling this with our discussion of the determinant, we learn the following.

Theorem. *A number λ is an eigenvalue with a non-zero eigenvector if and only if $\det(A - \lambda I) = 0$.*

Defintion. *The characteristic polynomial of A is the polynomial in λ given by*

$$p_A(\lambda) = \det(A - \lambda I).$$

Thus a restatement of the previous theorem is that the eigenvalues of A are precisely the roots of the polynomial $p_A(\lambda)$. Since the elements λ occur in precisely n positions along the diagonal, either our inductive result about determinants or our definition show that $p_A(\lambda)$ is a degree n polynomial.

Before continuing with this line of reasoning, we can ask how sensitive $p_A(\lambda)$ is to change-of-basis. Recall that for matrices, change of basis amounts to conjugating by an invertible element: $A \mapsto B = QAQ^{-1}$. So we check

$$\begin{aligned} p_B(\lambda) = \det(QAQ^{-1} - \lambda I) = \det(Q(A - \lambda I)Q^{-1}) = \\ \det(Q)\det(A - \lambda I)\det(Q^{-1}) = \det(Q) p_A(\lambda) \det(Q)^{-1} = p_A(\lambda). \end{aligned}$$

Thus we have an actual invariant of A that doesn't depend at all on the choice of basis. In particular, the coefficients of $p_A(\lambda)$ don't depend on the choice of basis. This gives a very nice family of invariants of matrices.

We finish today with the issue of roots of $p_A(\lambda)$. For a general field $\mathbb{F}$, we are guaranteed that there are at most n roots of $p_A(\lambda)$. What we aren't guaranteed is that there are any roots in $\mathbb{F}$.

Example: Let

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and let $\mathbb{F} = \mathbb{R}$. Then

$$p_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

This has no real roots, so we know that this has no eigenvalues over $\mathbb{R}$. It does have complex roots. If we extend scalars, passing from $\mathbb{R}^2$ to $\mathbb{C}^2$, then we get two complex eigenvalues: $\pm i$. This is true much more generally:

Theorem. *If $\mathbb{F}$ is any field, then $\mathbb{F}$ is a subfield of a field $\bar{\mathbb{F}}$ with the property that any polynomial with coefficients in $\bar{\mathbb{F}}$ has a root in $\bar{\mathbb{F}}$. Such a field is called algebraically closed.*

It's possible to do much of what we'll be talking now about over a non-algebraically closed field (for instance, over $\mathbb{R}$). The results are slightly harder to state (and significantly harder to prove). For now, for us, all of our fields can be assumed to be algebraically closed. Thus we learn

Theorem. *The polynomial $p_A(\lambda)$ has n roots, counted with multiplicity.*

The roots and the multiplicities will be increasingly important to us as we study the form of A.