Eigenstuff

Professor K. A. Ribet

October 13, 2025

Announcements

KR office hours:

Mondays 1:30–3 PM and Thursday
10:30 AM–noon in 732 Evans

Optional lunch at Crossroads: Thursday at 1PM

Week of October 20

- Math Monday talk on Fermat's Last Theorem, 5–6 PM, 1015 Evans
- Academic Empowerment Series event, "Office Hours Unlocked," 8–10 PM, Anchor House
- I'm teaching Math 113, TuTh, 2:10–3:30 PM, 155 Dwinelle

Invariant subspaces

A linear map $T \in \mathcal{L}(V)$ is called an *operator* on V.

If $U \subseteq V$ is a subspace and T is an operator on V, U is T-invariant (or invariant under T) if $Tu \in U$ for all $u \in U$.

The null space of an operator T is T-invariant, and so is the range of T. The subspace $\{0\}$ and the full space V are also invariant under T.

Our task will be to identify subspaces of V that are invariant under T.

Invariant subspaces

If U is T-invariant, then the restriction of T to U, a priori a linear map $U \to V$, is a map $U \to U$; in other words, it's an operator on U. We can write $T_{|U}$ for this operator.

More subtly perhaps, T induces an operator $V/U \to V/U$:

$$T_{V/U} : v + U \longmapsto Tv + U.$$

Everyone should check that this map is well defined.

In LADR, I think that $T_{V/U}$ is called $T_{/U}$, "T mod U."

Cyclic subspaces

If T is an operator on V and v is a vector in V, there is a smallest T-invariant subspace of V containing v (which may be $\{0\}$ or all of V), called the *cyclic subspace* of V generated by v.

The cyclic subspace generated by v is the span of the set $\{ T^k v \mid k \geq 1 \}$.

There'll be more about this subspace below.

Eigenvalues and eigenvectors

A 1-dimensional subspace $U = \mathbf{F} \cdot v$ of V (with $v \neq 0$) is T-invariant if and only if $Tv = \lambda v$ for some $\lambda \in \mathbf{F}$. We say that v is an *eigenvector* for T with eigenvalue λ.

Eigenvalues and eigenvectors

If λ is a scalar, λ is an eigenvalue for T if there is some nonzero $v \in V$ such that $Tv = \lambda v$. This means that the operator $T - \lambda I$ has a nonzero null space. If V is finite-dimensional, it's the same to say that $T - \lambda I$ is not invertible.

The null space of $T - \lambda I$ consists of 0, along with the (nonzero) eigenvectors whose eigenvalues are λ.

Linear independence of eigenvectors with distinct eigenvalues

Theorem (5.11)

Let T be an operator on V. Then every list of eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

Proof by induction on the length of the list: An eigenvector list of length 1 is linearly independent because eigenvectors are nonzero.

For the induction step, suppose $m \geq 2$ and that the theorem is true for lists of length $\leq m-1$. Let $v_1, \ldots, v_m$ be a list of eigenvectors with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. By induction, $v_1, \ldots, v_{m-1}$ is linearly independent. If the full list $v_1, \ldots, v_m$ is linearly *dependent*, then v_m is a linear combination of $v_1, \ldots, v_{m-1}$.

(Amazing) proof of the theorem

Write that v_m is a linear combination of $v_1, \ldots, v_{m-1}$:

$$v_m = a_1 v_1 + \cdots + a_{m-1} v_{m-1}.$$

Do two separate things to this equation: (1) apply T to both sides, and (2) multiply both sides by λ_m. Then

$$\lambda_m v_m = a_1 \lambda_1 v_1 + \cdots + a_{m-1} \lambda_{m-1} v_{m-1},$$
$$\lambda_m v_m = a_1 \lambda_m v_1 + \cdots + a_{m-1} \lambda_m v_{m-1}.$$

Subtraction yields

$$0 = a_1(\lambda_m - \lambda_1)v_1 + \cdots + (\lambda_m - \lambda_{m-1})v_{m-1}.$$

Because of the linear independence of $v_1, \ldots, v_{m-1}$, the coefficients $a_j(\lambda_m - \lambda_j)$ are 0 for $j = 1, \ldots, m-1$. Since the differences $\lambda_m - \lambda_j$ are nonzero, all a_j are 0. Therefore $v_m = 0$, which is a contradiction.

A quick corollary

Corollary (5.12)

An operator on an n-dimensional vector space has at most n distinct eigenvalues.

If an operator has m distinct eigenvalues, a list of eigenvectors for these eigenvalues is a linearly independent list of length m. By the first weeks of this course, m is at most n.

Polynomials applied to operators

IF T is an operator on V and p is a polynomial with coefficients in $\mathbf{F}$, there is a natural way to form $p(T)$.

We define $T^0 = I$ (the identity operator), $T^1 = T$, $T^2 = T \circ T$, $T^{k+1} = T \circ T^k$. Then

$$p(z) = a_0 + a_1 z + \cdots + a_d z^d \quad \rightsquigarrow \quad p(T) = a_0 I + a_1 T + \cdots + a_d T^d.$$

Polynomials applied to operators

You can verify properties like

$$(pq)(T) = p(T)q(T), \quad (p+q)(T) = p(T)q(T).$$

If I were teaching Math 113, I would say that the map

$$\mathcal{P}(\mathbf{F}) \to \mathcal{L}(V), \quad p \mapsto p(T)$$

is a *homomorphism*. That just means that it respects the additive and multiplicative structures on the two sides (source and target).

More invariant subspaces

Proposition

If p is a polynomial over $\mathbf{F}$, the null space and range of $p(T)$ are T-invariant subspaces of V.

These properties follow directly from definitions. For example, the range of $p(T)$ consists of all vectors $p(T)v$ with $v \in V$. Apply T: $T(p(T)v) = p(T)(Tv)$ because T commutes with all powers of T and thus with all polynomials in T. The vector $p(T)(Tv)$ is in the range of $p(T)$.

Remark: The system $\mathcal{P}(\mathbf{F})$ is commutative, whereas $\mathcal{L}(V)$ is highly noncommutative (unless $\dim V \leq 1$). Thus the image of $p \mapsto p(T)$ is a small part of $\mathcal{L}(V)$. If $V = \mathbf{F}^n$, so that T is a matrix A, the image consists of all polynomials in A.

Existence of eigenvalues over **C**

Theorem (5.19)

Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.

Proof: Let $n = \dim V$. Take $v \in V$, $v \neq 0$ and consider the list $v, Tv, T^2v, \ldots, T^nv$, which has length $n+1$. It is linearly dependent because $n+1 > n$. The linear dependence means that there is a nonzero polynomial p of degree $\leq n$ such that $p(T)v = 0$. Since v is nonzero, the polynomial is nonconstant. After division of p by its top coefficient, p becomes monic.

Since we're over **C**, p is a product $(z-\lambda_1)\cdots(z-\lambda_d)$, where the λ_j are complex numbers. Then

$$0 = p(T)v = (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_d I)v.$$

Because the product on the right has a nonzero null space, at least one of its factors $T-\lambda_j I$ is not 1-1. Then the number λ_j is an eigenvalue of T.

Polynomials applied to operators

If I were teaching Math 113, I would say that the linear map

$$\alpha : \mathcal{P}(\mathbf{F}) \to \mathcal{L}(V), \quad p \mapsto p(T)$$

is a *homomorphism*.

You've seen this before; but now I've written α for the linear map.

Lemma

The null space of α is nonzero.

This lemma follows from the fact that $\mathcal{P}(\mathbf{F})$ is infinite-dimensional, whereas $\mathcal{L}(V)$ has dimension $(\dim V)^2 = n^2$ (let's say). The list $I, T, T^2, \ldots, T^{n^2}$ has length $n^2 + 1$ and thus is linearly dependent. Hence there's a nonzero polynomial of degree $\leq n^2$ in the null space of α.

Minimal polynomial

What can we say about the null space of α, which is the set of polynomials p such that $p(T) = 0$? So far we know that it's nonzero. A key fact is that if $p(T) = 0$, then $(pq)(T) = 0$ for all polynomials q. (Math 113 talk: the null space is an *ideal*.)
Let m be a nonzero polynomial of minimal degree in the null space of α. If p is a polynomial, divide p by m to get $p = mq + r$, where q (quotient) is some polynomial and $\deg r < \deg m$. If $p(T) = 0$, then $r(T) = 0$. Since m is the nonzero polynomial of minimal degree that sends T to 0, r is the zero polynomial. Thus $p = mq$ is a multiple of m. We've proved something!

Proposition

The null space of α consists of all multiples of the polynomial m.

Minimal polynomial

Proposition

The null space of α consists of all multiples of the polynomial m.

The polynomial m is unique up to multiplication by constants. If m' is another polynomial of minimal degree in the null space of α, then m' is a *constant* multiple of m because $\deg m' = \deg m$. To rigidify things further, we take m to be monic — then it's literally unique.

We say that m is the *minimal polynomial* of T.

So far, we know that $\deg m$ is at most the square of $\dim V$. This is a very coarse estimate.

Examples

Let $V = \mathbf{F}^2$, so that T is given by a 2×2 matrix A.

If $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ then $m = z^2$.

If $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (the identity matrix), $m = z - 1$.

If $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then $m = (z - 1)(z + 1)$.

An amazing fact

Proposition

The minimal polynomial of an operator $T \in \mathcal{L}(V)$ has degree $\leq \dim V$.

The proof is by induction on $\dim V$. The base case is $\dim V = 0$, when $T = 0$ and we can take $m = 1$. (The identity map on the space $\{0\}$ is also the zero map.)

For fun, consider the case $\dim V = 1$. Take a nonzero v in V. Then $Tv = \lambda v$ for some $\lambda \in \mathbf{F}$ and the minimal polynomial of T is $z - \lambda$.

An amazing fact

Proposition

The minimal polynomial of an operator $T \in \mathcal{L}(V)$ has degree $\leq \dim V$.

The proof is by induction on $\dim V$. We can and will assume that $\dim V \geq 2$ and that the proposition is true for vector spaces of dimension $< \dim V$.

Imagine for a moment that we can find a nonzero T-invariant U subspace of V for which the proposition is true: the minimal polynomial p of $T_{|U}$ has degree $\leq \dim U$. By induction, the minimal polynomial q of $T_{V/U}$ will have degree $\leq \dim V/U$. The product pq then will have degree $\leq \dim U + \dim V/U = \dim V$, and will be such that $(pq)(T) = 0$.

The reason for the last statement is that $q(T)(V) \subseteq U$ by the definition of q, and then $p(q(T)(V)) = 0$ by the definition of p. Thus $(pq)(T) = 0$, so that T satisfies a nonzero polynomial of degree $\leq \dim V$.

Imagination runs wild

> *Imagine for a moment that we can find a nonzero T-invariant U subspace of V for which the proposition is true: the minimal polynomial p of $T_{|U}$ has degree $\leq \dim U$.*

We will take U to be the *cyclic subspace* generated by a nonzero vector $v \in V$ and the action of T. The concept of cyclic subspaces is a very important one. It's explained on the next few slides.

Back to cyclic subspaces

In proving the proposition, we now can and will stick to the case $V \neq \{0\}$. Take v nonzero in V and let U be the cyclic subspace of V generated by v. This space was introduced before as the smallest T-invariant subspace of V containing v.

Let $n = \dim V$. The list

$$v, Tv, T^2v, \ldots, T^nv$$

has length $n+1$ and is thus linearly dependent. By early in the course, there is some $d \geq 0$ so that T^dv is a linear combination of $v, Tv, \ldots, T^{d-1}v$. We take d as small as possible; then $v, Tv, \ldots, T^{d-1}v$ is linearly *independent*.

By the way, the case $d = 0$ corresponds to the situation where v is 0, which we excluded. Thus d is at least 1.

The claim is that $\text{span}(v, Tv, \ldots, T^{d-1}v)$ contains T^kv for all $k \geq 0$. This makes it T-invariant, and we see that it's U, which is the span of all the T^kv.

Cyclic subspaces

The claim is that $\operatorname{span}(v, Tv, \ldots, T^{d-1}v)$ contains $T^k v$ for all $k \geq 0$.

The proof of the claim is by a little induction. The span of $v, Tv, \ldots, T^{d-1}v$ contains $T^d v$ because $T^d v$ is a linear combination of $v, Tv, \ldots, T^{d-1}v$. Hence $\operatorname{span}(v, Tv, \ldots, T^{d-1}v) = \operatorname{span}(v, Tv, \ldots, T^{d-1}v, T^d v)$. Because T^d is a linear combination of $v, Tv, \ldots, T^{d-1}v$, $T^{d+1}v$ is a linear combination of $Tv, \ldots, T^{d-1}, T^d v$. Thus $T^{d+1}v \in \operatorname{span}(Tv, \ldots, T^{d-1}, T^d v) = \operatorname{span}(v, Tv, \ldots, T^{d-1}v)$. We continue in this way, or use a formal induction.

Cyclic subspaces

Note now that the cyclic subspace of V generated by v is $U = \text{span}(v, Tv, \ldots, T^{d-1}v)$ and has dimension d because the spanning list is linearly independent.

Write $T^d v = a_0 v + a_1 Tv + \cdots + a_{d-1} T^{d-1} v$ and let $p(z) = z^d - (a_{d-1} z^{d-1} + \cdots + a_1 z + a_0)$. Then $p(T)v = 0$. Further, $p(T) = 0$ on U because $p(T)(T^k v) = T^k(p(T)v) = 0$. Hence there is a polynomial p of degree $d = \dim U$ so that $p(T) = 0$ on U.

> *Imagine for a moment that we can find a nonzero T-invariant U subspace of V for which the proposition is true: the minimal polynomial of $T_{|U}$ has degree $\leq \dim U$.*

Mission accomplished! (Even if p is not the minimal polynomial, the minimal polynomial will be a divisor of p and will therefore also have degree $\leq \dim U$.)