

5A Invariant Subspaces

$x^2 + 1 = 0$
 $\uparrow$ cont factor over $\mathbb{R}$

5.1 definition: operator

A linear map from a vector space to itself is called an operator.

Recall that we defined the notation $\mathcal{L}(V)$ to mean $\mathcal{L}(V, V)$.

i.e. operators are maps
 $T \in \mathcal{L}(V)$

5.2 definition: invariant subspace

Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called invariant under T if $Tu \in U$ for every $u \in U$.

i.e. U is invariant under T if $T|_U$ is also an operator (i.e. $T|_U \in \mathcal{L}(U)$).

5.5 definition: eigenvalue

Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbb{F}$ is called an eigenvalue of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.

Idea The smallest nontrivial subspaces of a vector space V are $U = \text{span}(v)$ for $v \in V$ w/ $v \neq 0$.

So, if we can "break down" V into its smallest components, i.e. $V = V_1 \oplus \dots \oplus V_m$ w/ $V_i = \text{span}(v_i)$ (w/ $v_i \neq 0$)

Then, if each subspace is invariant under T , we have a good description of what T does: Since every $v \in V$ can be written uniquely as $v = v_1 + \dots + v_m$ w/ $v_i \in \text{span}(v_i)$,

then $v_i = a_i v_i$ for some $a_i \in \mathbb{F}$. By linearity, $Tv = a_1 Tv_1 + \dots + a_m Tv_m$.

So, if V_i is invariant under T , this says $T(v_i) \in V_i$, i.e.

$T(v_i) = b_i v_i$. So, $Tv_i = \frac{b_i}{a_i} v_i$ if $a_i \neq 0$. Hence, T is just like a

1-dimensional linear map on V_i & for $v \in V$, $Tv = a_1 v_1 + \dots + a_m v_m$

That is, we can tell exactly what T does in terms of the most simple linear map by finding 1-dimensional subspaces that are invariant under T .

5.7 equivalent conditions to be an eigenvalue

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Then the following are equivalent.

- (a) λ is an eigenvalue of T .
- (b) $T - \lambda I$ is not injective.
- (c) $T - \lambda I$ is not surjective.
- (d) $T - \lambda I$ is not invertible.

Reminder: $I \in \mathcal{L}(V)$ is the identity operator. Thus $Iv = v$ for all $v \in V$.

Proof (Sketch) Notice: $Tv = \lambda v$ for some $0 \neq v \in V$ if & only if $Tv - \lambda v = 0$ which says $(T - \lambda I)v = 0$. $\Leftrightarrow T - \lambda I$ has a nontrivial nullspace (i.e. $\text{null } T - \lambda I \neq \{0\}$)
 $\Leftrightarrow T - \lambda I$ is not injective
 $\Leftrightarrow T - \lambda I$ is not surjective
 $\Leftrightarrow T - \lambda I$ is not invertible
 where the last two $\Leftrightarrow$'s are due to the fact that $T - \lambda I$ is a map between V.S.'s of the same dimension.

5.8 definition: eigenvector

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of T . A vector $v \in V$ is called an **eigenvector** of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.

5.11 linearly independent eigenvectors

Suppose $T \in \mathcal{L}(V)$. Then every list of eigenvectors of T corresponding to distinct eigenvalues of T is linearly independent.

Proof $\S$ there is a linearly dependent list of eigenvectors of $T \in \mathcal{L}(V)$ corresponding to distinct eigenvalues of T . Let $v_1, \dots, v_m$ be this list & further that m is minimal. Necessarily, $m \geq 2$, since a list w/ one nonzero vector is lin. indep. By our assumed linear dependence, $\exists a_1, \dots, a_m \in \mathbb{F}$ st. $a_1 v_1 + \dots + a_m v_m = 0$ & not all of the a_i 's are zero. By minimality of m , notice that none of the a_i are zero, as if this were the case, we could remove the corresponding v_i from our list. Now, let us apply $T - \lambda_m I$ where λ_m is the eigenvalue corresponding to v_m , to both sides:

$$0 = (T - \lambda_m I)(a_1 v_1 + \dots + a_m v_m) = (T - \lambda_m I)a_1 v_1 + \dots + (T - \lambda_m I)a_m v_m = a_1 T v_1 - \lambda_m a_1 v_1 + \dots + a_{m-1} T v_{m-1} - \lambda_m a_{m-1} v_{m-1} = a_1 (\lambda_1 - \lambda_m) v_1 + \dots + a_{m-1} (\lambda_{m-1} - \lambda_m) v_{m-1}$$
 But notice, since each eigenvalue is distinct, $\lambda_i - \lambda_m \neq 0$ for $i = 1, \dots, m-1$. Hence, we have a linear combo of $v_1, \dots, v_{m-1}$ that is equal to zero & has no coefficient equal to 0. This contradicts the minimality of m !

5.12 operator cannot have more eigenvalues than dimension of vector space

Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Polynomials Applied to Operators

5.13 notation: $\underline{T}^m$

Suppose $\underline{T} \in \mathcal{L}(V)$ and m is a positive integer.

- $\underline{T}^m \in \mathcal{L}(V)$ is defined by $\underline{T}^m = \underbrace{\underline{T} \cdots \underline{T}}_{m \text{ times}}$.
- $\underline{T}^0$ is defined to be the identity operator I on V .
- If T is invertible with inverse $\underline{T}^{-1}$, then $\underline{T}^{-m} \in \mathcal{L}(V)$ is defined by

$$\underline{T}^{-m} = \underbrace{(\underline{T}^{-1})^m}.$$

5.14 notation: $\underline{p}(T)$

Suppose $\underline{T} \in \mathcal{L}(V)$ and $\underline{p} \in \mathcal{P}(\mathbf{F})$ is a polynomial given by

$$\underline{p}(z) = \underline{a_0} + \underline{a_1}z + \underline{a_2}z^2 + \cdots + \underline{a_m}z^m$$

for all $z \in \mathbf{F}$. Then $\underline{p}(\underline{T})$ is the operator on V defined by

$$\underline{p}(\underline{T}) = \underline{a_0}I + \underline{a_1}\underline{T} + \underline{a_2}\underline{T}^2 + \cdots + \underline{a_m}\underline{T}^m.$$

5.16 definition: product of polynomials

If $\underline{p}, \underline{q} \in \mathcal{P}(\mathbf{F})$, then $\underline{pq} \in \mathcal{P}(\mathbf{F})$ is the polynomial defined by

$$(\underline{pq})(z) = \underline{p(z)q(z)}$$

for all $z \in \mathbf{F}$.

5.17 multiplicative properties

Suppose $p, q \in \mathcal{P}(\mathbb{F})$ and $T \in \mathcal{L}(V)$.
Then

- (a) $(pq)(T) = p(T)q(T)$;
- (b) $p(T)q(T) = q(T)p(T)$.

Informal proof: When a product of polynomials is expanded using the distributive property, it does not matter whether the symbol is z or T .

Proof The second part is easy: T commutes w/ itself & both $p(T)$ & $q(T)$ only involve scaled sums of powers of T .

For a longer answer: (a) $(pq)(z) = p(z)q(z)$ by definition. So, plug in T to get $(pq)(T) = p(T)q(T)$ (this is not entirely correct to say, but is basically the idea). For full clarity, $(pq)(z) = \sum_{j=0}^m \sum_{k=0}^n a_j b_k z^{j+k}$

where $p(z) = \sum_{j=0}^m a_j z^j$ & $q(z) = \sum_{k=0}^n b_k z^k$. Thus, $(pq)(T) = \sum_{j=0}^m \sum_{k=0}^n a_j b_k T^{j+k} = \left(\sum_{j=0}^m a_j T^j \right) \left(\sum_{k=0}^n b_k T^k \right) = p(T)q(T)$.

For (b) just use (a) twice: $(pq)(T) = p(T)q(T)$
& $(qp)(T) = q(T)p(T)$.

5.18 null space and range of $p(T)$ are invariant under T

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Then $\text{null } p(T)$ and $\text{range } p(T)$ are invariant under T .

Proof $\forall u \in \text{null } p(T)$. Then, $p(T)u = 0$. Thus,

$$p(T)(Tu) = T(p(T)u) = T0 = 0. \text{ Hence, } Tu \in \text{null } p(T).$$

Since T is just $q(T)$ for $q(z) = z$

$\forall v \in \text{range } p(T)$. Then, $\exists u \in V$ s.t. $p(T)u = v$.

$$Tv = T(p(T)u) = p(T)(Tu). \text{ So, } Tv \in \text{range } p(T).$$

So, in either case, T does not move things in the null space (resp. range) out of the null space (resp. range) of $p(T)$. Hence, both $\text{null } p(T)$ & $\text{range } p(T)$ are invariant under T .