18.700 - LINEAR ALGEBRA, DAY 10
INVARIANT SUBSPACES AND EIGENVECTORS

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. Goals for lesson.

(1) Students will learn that if $\dim(V) = n$ and $\dim(W) = m$, then $\mathcal{L}(V, W) \cong \mathrm{M}_{\mathrm{m}\times\mathrm{n}}(\mathbb{F})$.
(2) Students will learn that $[T(v)]_{\mathcal{C}} = {}_{\mathcal{C}}[T]_{\mathcal{B}}[v]_{\mathcal{B}}$.
(3) Students will learn the change of basis formula.
(4) Students will learn the defintion of invariant subspace, eigenvalue, and eigenvector

I.2. Methods of assessment.

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

I.3. Materials to bring.
(1) Laptop + adapter (2) Worksheets

(0:00)

II. Lesson Plan

Announcements: • Exam 1: Wednesday in class. No pset this week; instead review packet. • TA review session: Tuesday, 7:00 - 9:00pm, 2-361

II.1. Last time.

- Defined matrix multiplication.
- Gave several interpretations of matrix multplication.
- Row rank = column rank of a matrix.
- Defined invertibility and isomorphism.
- Proved results on invertibility. (Invertible $\iff$ bijective, case of $\dim(V) = \dim(W)$.)

Definition 1. Let X and Y be sets. A function $f : X \to Y$ is *bijective* (or a *bijection*) if it is both injective and surjective.

Remark 2. Suppose $T \in \mathcal{L}(V, W)$. The result that, when $\dim(V) = \dim(W)$, then T invertible $\iff$ T one-to-one $\iff$ T onto is similar to a result about finite sets.

Lemma 3. *Suppose X and Y are finite sets such that $\#X = \#Y$, and $f : X \to Y$ is a function. TFAE.*

(i) f is bijective.
(ii) f is injective.
(iii) f is surjective.

II.1.1. *Isomorphic vector spaces, continued.*

Theorem 4 (Dimension determines isomorphism)**.** *Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic iff they have the same dimension.*

Proof. Suppose V and W are finite-dimensional vector spaces.

($\Rightarrow$): Assume V and W are isomorphic. Then there exists an isomorphism $T : V \to W$. Then T is injective and surjective so

$$\ker(T) = \{0\} \qquad \text{and} \qquad \operatorname{img}(T) = W\,.$$

By Rank-Nullity, then [ask students]

$$\dim(V) = \underbrace{\dim(\ker(T))}_{0} + \dim(\operatorname{img}(T)) = \dim(W)\,.$$

($\Leftarrow$): Assume $\dim(V) = \dim(W)$. Let $v_1, \ldots, v_n$ be a basis for V and $w_1, \ldots, w_n$ be a basis for W. By a previous result, there is a unique linear map $T : V \to W$ such that $T(v_i) = w_i$ for all $i = 1, \ldots, n$. Since $w_1, \ldots, w_n$ span W, then T is surjective. Either by Rank-Nullity, or by using the fact that $w_1, \ldots, w_n$ are linearly independent, T is injective. (Details left as exercise.) Thus T is injective and surjective, hence an isomorphism. ☐

Corollary 5. *Let V be an n-dimensional vector space. Then V is isomorphic to $\mathbb{F}^n$.*

Proof. Both have dimension n. ☐

Remark 6. We can also give an explicit isomorphism. Choose a basis $\mathcal{B} = (v_1, \ldots, v_n)$ for V and consider the coordinate vector map

$$\begin{aligned} \varphi_{\mathcal{B}} : V &\to \mathbb{F}^n \\ v &\mapsto [v]_{\mathcal{B}} \end{aligned}$$

and the linear map

$$\begin{aligned} S : \mathbb{F}^n &\to V \\ (a_1, \ldots, a_n) &\mapsto a_1 v_1 + \cdots + a_n v_n \,. \end{aligned}$$

Exercise: show these maps are mutually inverse isomorphisms.

Example 7. $\mathcal{P}_m(\mathbb{F})$ has dimension [ask students] $m+1$, hence is isomorphic to $\mathbb{F}^{m+1}$.

Proposition 8. *Suppose $\mathcal{B} := (v_1, \ldots, v_n)$ is a basis of V and $\mathcal{C} := (w_1, \ldots, w_m)$ is a basis of W (so $\dim(V) = n$ and $\dim(W) = m$). Then the map*

$$\begin{aligned} \mathcal{L}(V, W) &\to M_{m \times n}(\mathbb{F}) \\ T &\mapsto {}_{\mathcal{C}}[T]_{\mathcal{B}} \end{aligned}$$

is an isomorphism.

Proof. Exercise. (Similar to previous result.) □

Corollary 9. *Suppose V and W are finite-dimensional. Then* $\dim(\mathcal{L}(V, W)) = \dim(V)\dim(W)$.

II.1.2. *Linear maps as matrices.*

Proposition 10 (Multiplication by a matrix is linear)**.** *Let $A \in \mathrm{M_{m \times n}}(\mathbb{F})$. The left multiplication map*

$$\begin{aligned} L_A : \mathbb{F}^n &\to \mathbb{F}^m \\ v &\mapsto Av \end{aligned}$$

is linear.

Proof. Considering v as an $n \times 1$ matrix, this follows by properties of matrix multiplication. □

Let V and W be vector spaces with bases $\mathcal{B} := (v_1, \ldots, v_n)$ and $\mathcal{C} := (w_1, \ldots, w_m)$, respectively. Recall, for $T : V \to W$ linear, the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ is given by

$${}_{\mathcal{C}}[T]_{\mathcal{B}} = \begin{pmatrix} | & & | \\ [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix}$$

Proposition 11. *With notation as above,*

$$[T(v)]_{\mathcal{C}} = {}_{\mathcal{C}}[T]_{\mathcal{B}} [v]_{\mathcal{B}}$$

for all $v \in V$.

Proof. Given $v \in V$, there exist unique scalars $a_1, \ldots, a_n \in \mathbb{F}$ such that $v = a_1 v_1 + \cdots + a_n v_n$. Since T is linear, then

$$T(v) = T(a_1 v_1 + \cdots + a_n v_n) = a_1 T(v_1) + \cdots + a_n T(v_n)\,.$$

Since the coordinate vector map is linear, then

$$\begin{aligned} [T(v)]_{\mathcal{C}} &= [a_1 T(v_1) + \cdots + a_n T(v_n)]_{\mathcal{C}} = a_1 [T(v_1)]_{\mathcal{C}} + \cdots + a_n [T(v_n)]_{\mathcal{C}} \\ &= \begin{pmatrix} | & & | \\ [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \\ | & & | \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = {}_{\mathcal{C}}[T]_{\mathcal{B}}\, [v]_{\mathcal{B}}\,. \end{aligned}$$

□

The equality $[T(v)]_{\mathcal{C}} = {}_{\mathcal{C}}[T]_{\mathcal{B}}\, [v]_{\mathcal{B}}$ can be stated by saying the following diagram "commutes."

$$\begin{array}{ccc} V & \xrightarrow{\quad T \quad} & W \\ {\scriptstyle \varphi_{\mathcal{B}}}\big\downarrow & & \big\downarrow{\scriptstyle \varphi_{\mathcal{C}}} \\ \mathbb{F}^n & \xrightarrow[\quad L_{{}_{\mathcal{C}}[T]_{\mathcal{B}}} \quad]{} & \mathbb{F}^m \end{array}$$

[Draw image of v traveling both directions.]

Proposition 12. *Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then the rank of T (i.e., $\dim(\operatorname{img}(T))$) is equal to the (column) rank of $[T]$.*

Proof. Exercise. □

II.1.3. *Change of basis.* Q: How does the matrix ${}_{\mathcal{C}}[T]_{\mathcal{B}}$ change if we change the bases $\mathcal{B}$ and $\mathcal{C}$?

Definition 13. Let $n \in \mathbb{Z}_{\geq 0}$. The *$n \times n$ identity matrix I* is the $n \times n$ matrix with 1s on the diagonal and 0s elsewhere:

$$I = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Remark 14. We use I for both the identity operator and the identity matrix. With respect to *any* basis, the matrix of the identity operator I_V is I.

Definition 15. An $n \times n$ matrix A is *invertible* if there is a $n \times n$ matrix B such that $AB = BA = I$. We call B the *inverse* of A and denote it A^{-1}.

Lemma 16. *The inverse of a matrix is unique.*

Proof. Same as for linear maps. □

Theorem 17. *Let U, V, and W be vector spaces with bases $\mathcal{B}, \mathcal{C}$, and $\mathcal{D}$, respectively. Given $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then*

$${}_{\mathcal{D}}[ST]_{\mathcal{B}} = {}_{\mathcal{D}}[S]_{\mathcal{C}}\ {}_{\mathcal{C}}[T]_{\mathcal{B}}\,.$$

Proof. Follows by the definition of matrix multiplication. □

Corollary 18 (Change of basis matrix)**.** *Suppose $\mathcal{B}$ and $\mathcal{C}$ are both bases for V. Then*

$$_{\mathcal{B}}[I]_{\mathcal{C}} = {}_{\mathcal{C}}[I]_{\mathcal{B}}^{-1}\,.$$

Proof.

$$I = {}_{\mathcal{B}}[I]_{\mathcal{B}} = {}_{\mathcal{B}}[I]_{\mathcal{C}}\ {}_{\mathcal{C}}[I]_{\mathcal{B}}\,.$$

□

Proposition 19 (Change of basis formula)**.** *Suppose $\mathcal{B}$ and $\mathcal{C}$ are both bases of V. Given $T \in \mathcal{L}(V)$, let $A := [T]_{\mathcal{B}}$, $B := [T]_{\mathcal{C}}$, and $C = {}_{\mathcal{B}}[I]_{\mathcal{C}}$. Then*

$$A = CBC^{-1}\,.$$

Proof.

$$_{\mathcal{B}}[T]_{\mathcal{B}} = {}_{\mathcal{B}}[I]_{\mathcal{C}}\ {}_{\mathcal{C}}[T]_{\mathcal{C}}\ {}_{\mathcal{C}}[I]_{\mathcal{B}} = ({}_{\mathcal{C}}[I]_{\mathcal{B}})^{-1}{}_{\mathcal{C}}[T]_{\mathcal{C}}\ {}_{\mathcal{C}}[I]_{\mathcal{B}}\,.$$

□

Definition 20. Two $n \times n$ matrices A and B are *similar* or *conjugate* if there is an invertible matrix P such that $B = PAP^{-1}$.

II.2. **Summary of Ch. 4 of Axler.** Let $p \in \mathcal{P}(\mathbb{F})$ be a polynomial.

- There is a division algorithm for polynomials.
- If $p(r) = 0$ for some $r \in \mathbb{F}$, then there exists $q \in \mathcal{P}(\mathbb{F})$ such that $p(z) = (z-r)q(z)$.
- A degree m polynomial has at most m roots in $\mathbb{F}$.
- The fundamental theorem of algebra: A degree m polynomial in $\mathcal{P}(\mathbb{C})$ has exactly m roots in $\mathbb{C}$. Equivalently, every polynomial in $\mathcal{P}(\mathbb{C})$ splits into linear factors:
  $$p(z) = c(z-r_1)\cdots(z-r_m)$$
  for some $c, r_1, \ldots, r_m \in \mathbb{C}$.
- Every polynomial in $\mathcal{P}(\mathbb{R})$ splits into factors of degree at most 2.

II.3. **5A: Invariant subspaces and Eigenvectors.** Throughout this section, let V be a vector space over $\mathbb{F}$. Recall that a *linear operator* is a linear map $T : V \to V$, i.e., from a vector space to itself.

Definition 21. Let $T \in \mathcal{L}(V)$. A subspace U of V is *stable* or *invariant under T* (or T-stable or T-invariant) if $T(u) \in U$ for all $u \in U$.

Remark 22. If U is T-invariant, then the restriction $T|_U : U \to U$ is well-defined, and is a linear operator on U.

Example 23. Let $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ be the differentiation operator $T(f) = f'$. Then $\mathcal{P}_4(\mathbb{R})$ is T-stable: if $\deg(f) \leq 4$, then $\deg(f') = \deg(f) - 1 \leq 4$. Similarly, $\mathcal{P}_m(\mathbb{R})$ is T-stable for every $m \in \mathbb{Z}_{\geq 0}$.

Lemma 24. *Let $T \in \mathcal{L}(V)$. Then $\{0\}$, V, $\ker(T)$, and $\operatorname{img}(T)$ are all T-invariant.*

Remark 25. These are not necessarily all distinct!

Proof. Exercise. □

Q: Does every linear operator have an invariant subspace other than $\{0\}$ and V?

We'll see later that the answer is yes for $\mathbb{F} = \mathbb{C}$ if $\dim(V) \geq 2$, and yes for $\mathbb{F} = \mathbb{R}$ if $\dim(V) \geq 3$.

Let's first consider 1-dimensional invariant subspaces. Given $v \in V$ with $v \neq 0$, let [ask students]

$$U := \operatorname{span}(v) = \{\lambda v : \lambda \in \mathbb{F}\}.$$

If U is T-invariant, then in particular, $T(v) \in U$, so $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. Conversely, if $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$, then $\operatorname{span}(v)$ is T-invariant.

Definition 26. Let $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if there exists $v \in V$ with $v \neq 0$ such that $T(v) = \lambda v$. Such a v is called an *eigenvector* corresponding to λ.

Remark 27.
- "eigen-" means "self" or "own".
- We require that $v \neq 0$ because $T(0) = \lambda 0$ for all $\lambda \in \mathbb{F}$.

[Show gif depicting eigenvectors in $\mathbb{R}^2$.]

Theorem 28. *Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. TFAE.*

(i) λ is an eigenvalue of T.
(ii) $T - \lambda I$ is not injective.
(iii) $T - \lambda I$ is not surjective.
(iv) $T - \lambda I$ is not invertible.

Proof. (a) $\implies$ (b): Assume λ is an eigenvalue of T with corresponding eigenvector $v \neq 0$, so $T(v) = \lambda v$. Then

$$0 = T(v) - \lambda v = (T - \lambda I)(v)$$

so $0 \neq v \in \ker(T)$. Thus T is not one-to-one.

(b) $\implies$ (a): Assume $T - \lambda I$ is not injective. Then $\ker(T - \lambda I) \neq \{0\}$ so there exists $0 \neq v \in \ker(T - \lambda I)$. Then

$$0 = (T - \lambda I)(v) = T(v) - \lambda v \implies T(v) = \lambda v$$

so v is an eigenvector with eigenvalue λ.

We previously showed the equivalence of (b), (c), and (d). □

Proposition 29. *Let $T \in \mathcal{L}(V)$. Suppose that $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_k$. Then $v_1, \ldots, v_k$ are linearly independent.*

Proof. We proceed by induction on k, the number of eigenvalues.

Base case: $k = 1$. An eigenvector is nonzero by definition, so the list v_1 is linearly independent by a previous homework problem.

Inductive step: Assume the result holds for $k - 1$ and assume T has k distinct eigenvalues. Suppose that

$$a_1 v_1 + \cdots + a_k v_k = 0 \tag{30}$$

for some $a_1, \ldots, a_k \in \mathbb{F}$. Goal: $a_i = 0$ for all i. Note that

$$(T - \lambda_k I)(v_i) = T(v_i) - \lambda_k v_i = \lambda_i v_i - \lambda_k v_i = (\lambda_i - \lambda_k) v_k$$

for all $i = 1, \ldots, k$. Applying $T - \lambda_k I$ to (30), we find

$$\begin{aligned} 0 &= (T - \lambda_k I)(a_1 v_1 + \cdots + a_m v_m) \\ &= a_1(\lambda_1 - \lambda_k)v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} + \cancelto{0}{a_k(\lambda_k - \lambda_k)v_k} \,. \end{aligned}$$

Since $v_1, \ldots, v_{k-1}$ are linearly independent by the inductive hypothesis, then $a_i(\lambda_i - \lambda_k) = 0$ for all $i = 1, \ldots, k-1$. Since the λ_i are distinct, then $a_i = 0$ for all $i = 1, \ldots, k-1$. Then (30) becomes

$$a_k v_k = 0 \,.$$

But v_k is an eigenvector, hence is nonzero, so $a_k = 0$ by the base case. □

Corollary 31. *If V is finite-dimensional, then every operator $T \in \mathcal{L}(V)$ has at most $\dim(V)$ distinct eigenvalues.*

Proof. Apply the previous result and LI $\leq$ span. □

II.3.1. *Polynomials applied to linear operators.* Given a linear operator $T : V \to V$, then we can compose T with itself: $T \circ T = T^2$. We similarly define

$$T^m = \begin{cases} \overbrace{T \cdots T}^{m \text{ times}} & \text{if } m > 0; \\ I & \text{if } m = 0; \\ (T^{-1})^{|m|} & \text{if } m < 0 \text{ and } T \text{ is invertible.} \end{cases}$$

Lemma 32.

- $T^m T^n = T^{m+n}$
- $(T^m)^n = T^{mn}$

Proof. Exercise. □

Definition 33. Given $T \in \mathcal{L}(V)$, and $p \in \mathcal{P}(\mathbb{F})$ with

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m \,,$$

define the operator $p(T) \in \mathcal{L}(V)$ by

$$p(T) := a_0 I + a_1 T + \cdots + a_m T^m \,.$$

Definition 34. Let $p, q \in \mathcal{P}(\mathbb{F})$. Their product pq is defined pointwise:

$$(pq)(z) := p(z)q(z)$$

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

Note that multiplication of polynomials is commutative. The same is true when we apply polynomials to linear operators.

Lemma 35.

(i) $(pq)(T) = p(T)q(T)$;
(ii) $p(T)q(T) = q(T)p(T)$.

Proof. Exercise. □

Lemma 36. *Let $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Then $\ker(T)$ and $\operatorname{img}(T)$ are T-invariant.*

Proof. Exercise. □