18.700 - LINEAR ALGEBRA, DAY 16
ORTHOGONAL COMPLEMENTS, MINIMIZATION
ADJOINT, SELF-ADJOINT, AND NORMAL OPERATORS

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. **Goals for lesson.**

(1) Students will learn how to compute the orthogonal projection of a vector onto a subspace.
(2) Students will learn properties of orthogonal complements.
(3) Students will learn the definition of the adjoint of a linear operator.
(4) Students will learn the definition of self-adjoint and normal operators.

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 (3) Chalk

(0:00)

II. Lesson Plan

Announcements: • Midterm Exam 2: Wednesday, November 13th in class

II.1. Last time.

- Stated basic results in inner product spaces, e.g., triangle inequality, Cauchy-Schwarz, Pythagorean theorem, Parallelogram Identity.
- Proved that the Gram-Schmidt procdure produces orthonormal bases.
- Defined the orthogonal complement of a subset of an inner product space.

II.2. 6C: Orthogonal complements and minimization, cont.

Definition 1. Given a subset $S \subseteq V$, the *orthogonal complement of* S is

$$S^{\perp} := \{v \in V : \langle u, v \rangle = 0 \, \forall u \in S\} = \{v \in V : v \perp u \, \forall u \in S\} .$$

Proposition 2. *If S is a subset of V, then $S \cap S^{\perp} \subseteq \{0\}$.*

The above proposition hints at the following result.

Proposition 3. *Suppose U is a finite-dimensional subspace of V. Then*

$$V = U \oplus U^{\perp} .$$

Proof. Since U and $U^{\perp}$ are subspaces, then $0 \in U$ and $0 \in U^{\perp}$, so $U \cap U^{\perp} = \{0\}$ by part (d) of the previous result. Thus $U + U^{\perp}$ is direct.

It remains to show that $V = U + U^{\perp}$. Certainly $V \supseteq U + U^{\perp}$, so it suffices to show that $V \subseteq U + U^{\perp}$. [Ask students.] Suppose $v \in V$. By a previous result, there exists an orthonormal basis $e_1, \dots, e_m$ of U. Let

$$\begin{aligned} u &:= \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m \\ w &:= v - u. \end{aligned}$$

Then $v = u + w$ and $u \in U$. Goal: $w \in U^{\perp}$. [Ask students how to show this.] For each $k \in \{1, \dots, m\}$, we have

$$\langle w, e_k \rangle = \left\langle v - \sum_{i=1}^{m} \langle v, e_i \rangle e_i, e_k \right\rangle = \langle v, e_k \rangle - \sum_i \langle v, e_i \rangle \overbrace{\langle e_i, e_k \rangle}^{=0 \text{ for } i \neq k} = \langle v, e_k \rangle - \langle v, e_k \rangle .$$

Thus w is orthogonal to $e_1, \dots, e_m$, so w is orthogonal to every vector in $\operatorname{span}(e_1, \dots, e_m) = U$. Thus $w \in U^{\perp}$. □

Corollary 4. *Suppose V is finite-dimensional and U is a subspace of V. Then*

$$\dim(U^{\perp}) = \dim(V) - \dim(U) .$$

Proposition 5. *Suppose U is a finite-dimensional subspace of V. Then*

$$(U^{\perp})^{\perp} = U .$$

Proof. ($\supseteq$): Exercise.

($\subseteq$): Suppose $v \in (U^\perp)^\perp$. By a previous result, we can write $v = u + w$ where $u \in U$ and $w \in U^\perp$. Goal: $w = 0$. From the first part, we have $u \in U \subseteq (U^\perp)^\perp$, so

$$w = v - u \in (U^\perp)^\perp .$$

But then $w \in U^\perp \cap (U^\perp)^\perp = \{0\}$, so $w = 0$ and $v = u \in U$. □

Corollary 6. *With the same hypotheses as above,*

$$U^\perp = \{0\} \iff U = V .$$

Proof. Exercise. □

Definition 7 (Orthogonal projection)**.** Suppose U is a finite-dimensional subspace of V. For each $v \in V$, we write write $v = u + w$ where $u \in U$ and $w \in U^\perp$. The *orthogonal projection of v onto U* is $\operatorname{proj}_U(v) := u$. This defines a linear map $\operatorname{proj}_U \in \mathcal{L}(V)$.

Since $V = U \oplus U^\perp$, then the expression $v = u + w$ above is unique, so the map proj_U is well-defined.

Proposition 8. *Suppose U is a finite-dimensional subpsace of V. Then*

(i) $\operatorname{proj}_U \in \mathcal{L}(V)$;
(ii) $\operatorname{proj}_U|_U = I_U$, *i.e.*, $\operatorname{proj}_U(u) = u$ *for all* $u \in U$;
(iii) $\operatorname{proj}_U|_{U^\perp} = 0$, *i.e.*, $\operatorname{proj}_U(w) = 0$ *for all* $w \in U^\perp$;
(iv) *[Ask students]* $\operatorname{img}(\operatorname{proj}_U) = U$;
(v) *[Ask students]* $\ker(\operatorname{proj}_U) = U^\perp$;
(vi) $v - \operatorname{proj}_U(v) \in U^\perp$ *for all* $v \in V$;
(vii) $\operatorname{proj}_U^2 = \operatorname{proj}_U$;
(viii) $\|\operatorname{proj}_U(v)\| \leq \|v\|$ *for all* $v \in V$;
(ix) *if* $e_1, \ldots, e_m$ *is an orthonormal basis of* U, *then*

$$\operatorname{proj}_U(v) = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m .$$

Proof. Exercise. □

Remark 9. Property (ix) gives us a formula to compute an orthogonal projection, given an orthonormal basis for the subspace.

Proposition 10 (Minimizing distance to a subspace)**.** *Suppose U is a finite-dimensional subpsace of V and $v \in V$. Then*

$$\|v - \operatorname{proj}_U(v)\| \leq \|v - u\|$$

for all $u \in U$, with equality iff $u = \operatorname{proj}_U(v)$.

Proof. Given $u \in U$, then $\operatorname{proj}_U(v) - u \in U$. By orthogonal decomposition, $v - \operatorname{proj}_U(v) \in U^\perp$. Since

$$v - u = (v - \operatorname{proj}_U(v)) + (\operatorname{proj}_U(v) - u)$$

and these last two are orthogonal, then

$$\|v - u\|^2 = \|v - \operatorname{proj}_U(v)\|^2 + \overbrace{\|\operatorname{proj}_U(v) - u\|^2}^{\geq 0} \geq \|v - \operatorname{proj}_U(v)\|^2 .$$

Taking square roots yields the result. □

In calculus, you were sometimes faced with the following problem. Suppose L is a line through the origin in $\mathbb{R}^2$ and P is a point not lying on the line L. What is the distance from P to L, i.e., what is the point on L closest to P? [Draw picture.]

The answer uses the ideas of orthogonal projection and orthogonal decomposition. Let u be the vector from the origin to P, and let v be a vector in the direction of L. [Continue picture.] Then $L = \operatorname{span}(v)$ and $\frac{1}{\|v\|}v$ is an orthonormal basis for L. By the proposition, then

$$\operatorname{proj}_L(u) = \left\langle u, \frac{1}{\|v\|}v \right\rangle \frac{1}{\|v\|}v = \frac{1}{\|v\|^2}\langle u, v\rangle v$$

is the point on L that is closest to P.

II.3. **Worksheet.**

II.4. **7A: Adjoint, Self-Adjoint, and Normal Operators.** For today, let V and W be nonzero finite-dimensional inner product spaces over $\mathbb{F}$.

A brief addendum to section 6B:

Definition 11. Let V be an $\mathbb{F}$-vector space.

- A *linear functional* on V is a linear map $\varphi : V \to \mathbb{F}$.
- The *dual space of* V, denoted $V^\vee$ or V^* or V', is

$$V^\vee := \mathcal{L}(V, \mathbb{F}).$$

In other words, the vector space of all linear functionals on V.

Theorem 12 (Riesz Representation Theorem)**.** *Suppose V is a finite-dimensional inner product space, and $\varphi \in \mathcal{L}(V, \mathbb{F})$. Then there is a unique vector $v \in V$ such that*

$$\varphi(u) = \langle u, v\rangle$$

for all $u \in V$.

Proof. Existence: Let $e_1, \dots, e_n$ be an orthonormal basis for V. Let

$$v := \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n .$$

Given $u \in V$, then

$$\begin{aligned}\langle u, v\rangle &= \left\langle u, \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n \right\rangle = \left\langle u, \overline{\varphi(e_1)}e_1 \right\rangle + \cdots + \left\langle u, \overline{\varphi(e_n)}e_n \right\rangle \\ &= \varphi(e_1)\langle u, e_1\rangle + \cdots + \varphi(e_n)\langle u, e_n\rangle = \varphi(\langle u, e_1\rangle e_1 + \cdots + \langle u, e_n\rangle e_n) = \varphi(u) .\end{aligned}$$

[Ask students if v depends on u.]

Uniqueness: Suppose $v_1, v_2 \in V$ both satisfy

$$\langle u, v_1\rangle = \varphi(u) = \langle u, v_2\rangle$$

for all $u \in V$. Then

$$\langle u, v_1 - v_2\rangle = \langle u, v_1\rangle - \langle u, v_2\rangle = 0$$

for all $u \in V$, so $v_1 - v_2 = 0$ by a previous result. Thus $v_1 = v_2$. □

Proposition 13. *Given $T \in \mathcal{L}(V, W)$, then there exists a unique linear map $T^* \in \mathcal{L}(W, V)$ such that*

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle \tag{*}$$

for all $v \in V$ and all $w \in W$.

Definition 14. The linear map T^* above is called the *adjoint of T*. That is, it is the unique map satisfying (*).

Proof of proposition. Fix $w \in W$ and consider the linear functional

$$\begin{aligned} \varphi : V &\to \mathbb{F} \\ v &\mapsto \langle T(v), w\rangle . \end{aligned}$$

By the Riesz Representation Theorem, there exists a unique $u \in V$ such that

$$\langle T(v), w\rangle = \varphi(v) = \langle v, u\rangle$$

for all $v \in V$. Define $T^*(w) := u$; then

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle$$

for all $v \in V$.

It remains to show that $T^* : W \to V$ is linear: exercise. □

Remark 15. In the above equation, the LHS is the inner product on W, while the righthand side is the inner product on V.

Proposition 16. *Suppose $T \in \mathcal{L}(V, W)$.*

(i) $(S + T)^* = S^* + T^*$ for all $S \in \mathcal{L}(V, W)$.
(ii) $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbb{F}$.
(iii) $(T^*)^* = T$.
(iv) Let U be a finite-dimensional inner product space. Then $(ST)^* = T^*S^*$ for all $S \in \mathcal{L}(W, U)$.
(v) $I^* = I$.
(vi) If T is invertible, then T^* is also invertible, and $(T^*)^{-1} = (T^{-1})^*$.

Proof. Suppose $v \in v$, $w \in W$, and $\lambda \in \mathbb{F}$.

(i) By definition,

$$\langle (S+T)(v), w\rangle = \langle v, (S+T)^*(w)\rangle .$$

Now

$$\begin{aligned} \langle (S+T)(v), w\rangle &= \langle S(v), w\rangle + \langle T(v), w\rangle = \langle v, S^*(w)\rangle + \langle v, T^*(w)\rangle \\ &= \langle v, S^*(w) + T^*(w)\rangle = \langle v, (S^* + T^*)(w)\rangle . \end{aligned}$$

(ii) Similar.

(iii)

$$\langle T^*(w), v\rangle = \overline{\langle v, T^*(w)\rangle} = \overline{\langle T(v), w\rangle} = \langle w, T(v)\rangle .$$

(iv) Given $S \in \mathcal{L}(W, U)$ and $u \in U$, then

$$\langle (ST)(v), u\rangle = \langle S(T(v)), u\rangle = \langle T(v), S^*(u)\rangle = \langle v, T^*(S^*(u))\rangle .$$

(v) Exercise.

(vi) Apply * to the equations $T^{-1}T = I$ and $TT^{-1} = I$ and then apply the two previous parts.

□

Proposition 17. *Suppose $T \in \mathcal{L}(V, W)$. Then*

(i) $\ker(T^*) = (\operatorname{img}(T))^\perp$;
(ii) $\operatorname{img}(T^*) = \ker(T)^\perp$;
(iii) $\ker(T) = (\operatorname{img}(T^*))^\perp$;
(iv) $\operatorname{img}(T) = (\ker(T^*))^\perp$.

Proof. (i) ($\subseteq$): Given $w \in \ker(T^*)$, then $0 = T^*(w)$. Given $x \in \operatorname{img}(T)$, then $x = T(v)$ for some $v \in V$. Then

$$\langle x, w\rangle = \langle T(v), w\rangle = \langle v, T^*(w)\rangle = \langle v, 0\rangle = 0\,.$$

Thus $w \in (\operatorname{img}(T))^\perp$.

($\supseteq$): Similar.

(ii) Replace T by T^* in the previous part.
(iii) Take the orthogonal complement of (i).
(iv) Take the orthogonal complement of (ii).

□

Q: After having chosen bases, how does the matrix of T^* relate to the matrix of T?

Definition 18. Let $A \in M_{m\times n}(\mathbb{F})$. The *conjugate transpose of A*, denoted A^*, is defined by

$$(A^*)_{ij} = \overline{(A^t)_{ij}} = \overline{A_{ji}}\,.$$

Proposition 19. *Let $T \in \mathcal{L}(V, W)$, let $\mathcal{E} := (e_1, \ldots, e_n)$, and $\mathcal{F} := (f_1, \ldots, f_m)$ be orthonormal bases for V and W, respectively. Then*

$$_{\mathcal{E}}[T^*]_{\mathcal{F}} = (_{\mathcal{F}}[T]_{\mathcal{E}})^*\,.$$

Proof. Recall that the k^{th} column of $_{\mathcal{F}}[T]_{\mathcal{E}}$ is $[T(e_k)]_{\mathcal{F}}$. Since $\mathcal{F}$ is orthonormal, we have

$$T(e_k) = \langle T(e_k), f_1\rangle f_1 + \cdots \langle T(e_k), f_m\rangle f_m\,,$$

so

$$[T(e_k)]_{\mathcal{F}} = \begin{pmatrix} \langle T(e_k), f_1\rangle \\ \vdots \\ \langle T(e_k), f_1\rangle \end{pmatrix}\,.$$

Thus

$$(_{\mathcal{F}}[T]_{\mathcal{E}})_{jk} = \langle T(e_k), f_j\rangle\,.$$

Similarly

$$(_{\mathcal{E}}[T^*]_{\mathcal{F}})_{jk} = \langle T^*(f_k), e_j\rangle = \langle f_k, T(e_j)\rangle = \overline{\langle T(e_j), f_k\rangle}\,.$$

Thus

$$(_{\mathcal{E}}[T^*]_{\mathcal{F}})_{jk} = \overline{(_{\mathcal{F}}[T]_{\mathcal{E}})_{kj}} = \overline{(_{\mathcal{F}}[T]_{\mathcal{E}})^t_{jk}}\,.$$

□

II.4.1. *Self-adjoint operators.*

Definition 20. An operator $T \in \mathcal{L}(V)$ is *self-adjoint* if $T = T^*$.

Lemma 21. *If $\mathcal{E}$ is an orthonormal basis for V, then T is self-adjoint iff $[T^*]_{\mathcal{E}} = ([T]_{\mathcal{E}})^*$.*

Example 22. Let $T \in \mathcal{L}(\mathbb{F}^2)$ be the linear operator such that

$$[T]_{\mathcal{E}} = \begin{pmatrix} 2 & i \\ -i & 7 \end{pmatrix}.$$

T is self-adjoint.

Remark 23. The adjoint of a linear operator is similar to the complex conjugate of a complex number.

Proposition 24. *Let $T \in \mathcal{L}(V)$ be self-adjoint. Then every eigenvalue of T is real.*

Proof. Suppose $\lambda \in \mathbb{F}$ is an eigenvalue of T, so $T(v) = \lambda v$ for some $0 \neq v \in V$. Then

$$\langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda < v, v\rangle = \lambda \|v\|^2$$
$$\langle T(v), v\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\|v\|^2.$$

Since $v \neq 0$, then $\|v\|^2 \neq 0$, so $\overline{\lambda} = \lambda$. Thus $\lambda \in \mathbb{R}$. □

Proposition 25. *Suppose V is an inner product space over $\mathbb{C}$ and $T \in \mathcal{L}(V)$. Then T is self-adjoint iff $\langle T(v), v\rangle \in \mathbb{R}$ for all $v \in V$.*

Proof. ($\Rightarrow$): Assume T is self-adjoint, so $T = T^*$. Given $v \in V$, then

$$\langle T^*(v), v\rangle = \overline{\langle v, T^*(v)\rangle} = \overline{\langle T(v), v\rangle}.$$

Then

$$0 = \langle 0(v), v\rangle = \langle (T - T^*)(v), v\rangle = \langle T(v), v\rangle - \langle T^*(v), v\rangle = \langle T(v), v\rangle - \overline{\langle T(v), v\rangle}.$$

Thus $\langle T(v), v\rangle$ is real.

($\Leftarrow$): Exercise. (Similar.) □

Definition 26. An operator $T \in \mathcal{L}(V)$ is *normal* if T commutes with its adjoint, i.e.,

$$TT^* = T^*T.$$

Example 27. Let $T \in \mathcal{L}(F^2)$ whose matrix with respect to the standard basis is

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

Then T is not self-adjoint (matrix is not symmetric), but is normal. [Compute TT^* and T^*T.]