18.700 - LINEAR ALGEBRA, DAY 18
THE SPECTRAL THEOREM
POSITIVE OPERATORS, ISOMETRIES, AND MATRIX FACTORIZATIONS

SAM SCHIAVONE

CONTENTS

I. Pre-class Planning 1
I.1. Goals for lesson 1
I.2. Methods of assessment 1
I.3. Materials to bring 1
II. Lesson Plan 2
II.1. Last time 2
II.2. 7B The Spectral Theorem, cont. 2
II.3. 7C: Positive Operators 4
II.4. 7D: Isometries, Unitary Operators, and Matrix Factorizations 6

I. PRE-CLASS PLANNING

I.1. Goals for lesson.

(1) Students will learn the Spectral Theorem over $\mathbb{R}$.
(2) Students will learn the defintion of positive linear operators.
(3) Students will learn characterization and properties of positive linear operators.
(4) Students will learn the defintion of isometries and unitary operators.
(5) Students will learn QR Decomposition.

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 grades posted tonight

II.1. **Last time.**

- Proved properties of the adjoint of a linear operator.
- Defined self-adjoint ($T^* = T$) and normal ($TT^* = T^*T$) operators.
- Proved properties of self-adjoint and normal operators.
- Stated the Spectral Theorem over $\mathbb{C}$.

II.2. **7B The Spectral Theorem, cont.** Throughout today, let V be a finite-dimensional inner product space over $\mathbb{F}$.

Theorem 1 (Spectral Theorem over $\mathbb{C}$)**.** *Suppose $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. TFAE.*

(i) T is normal.
(ii) There is an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is diagonal.
(iii) There is an orthonormal basis of V consisting of eigenvectors of T.

Proof. (i) $\implies$ (ii): Last time.

(ii) $\implies$ (i): Assume there is an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is diagonal. Then $[T^*]_{\mathcal{E}} = ([T]_{\mathcal{E}})^*$, which is also diagonal. Since diagonal matrices commute, then

$$[TT^*]_{\mathcal{E}} = [T]_{\mathcal{E}}[T^*]_{\mathcal{E}} = [T^*]_{\mathcal{E}}[T]_{\mathcal{E}} = [T^*T]_{\mathcal{E}},$$

so $TT^* = T^*T$. (Recall that $[\cdot]_{\mathcal{E}} : \mathcal{L}(V) \to M_{n\times n}(\mathbb{F})$ is an isomorphism.) Thus T is normal. □

Theorem 2 (Spectral Theorem over $\mathbb{R}$)**.** *Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. TFAE.*

(i) T is self-adjoint.
(ii) There is an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is diagonal.
(iii) There is an orthonormal basis of V consisting of eigenvectors of T.

We'll need some preliminary results to prove the theorem.

Lemma 3. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint, and $b, c \in \mathbb{R}$ with $b^2 - 4c < 0$. Then*

$$T^2 + bT + cI$$

is an invertible operator.

Proof. [Skip if necessary.] Suppose $0 \neq v \in V$. By Cauchy-Schwarz, we have

$$\begin{aligned}|\langle bT(v), v\rangle| &\leq \|bT(v)\|\|v\| = |b|\|T(v)\|\|v\| \\ &\iff -|b|\|T(v)\|\|v\| \leq \langle bT(v), v\rangle \leq |b|\|T(v)\|\|v\|.\end{aligned}$$

Since T is self-adjoint, then

$$\begin{aligned}\langle (T^2 + bT + cI)(v), v\rangle &= \langle T^2(v), v\rangle + b\langle T(v), v\rangle + c\langle v, v\rangle \\ &= \langle T(v), T(v)\rangle + b\langle T(v), v\rangle + c\|v\|^2 \\ &\geq \|T(v)\|^2 - |b|\langle T(v), v\rangle + c\|v\|^2 \\ &= \left(\|T(v)\| - \frac{|b|\|v\|}{2}\right)^2 - \frac{|b|^2\|v\|^2}{4} + c\|v\|^2 \\ &= \underbrace{\left(\|T(v)\| - \frac{|b|\|v\|}{2}\right)^2}_{\geq 0} + \underbrace{\frac{4c - b^2}{4}}_{>0}\|v\| > 0\,.\end{aligned}$$

Thus $(T^2 + bT + cI)(v) \neq 0$. This shows that $\ker(T^2 + bT + cI) = \{0\}$, so this operator is injective. Since the domain and codomain are both V, then this implies that it is invertible. □

Proposition 4. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Then*

$$\operatorname{minpoly}(T) = (z - \lambda_1)\cdots(z - \lambda_m)$$

for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

Proof. Case 1: $\mathbb{F} = \mathbb{C}$. Then $\operatorname{minpoly}(T)$ splits into degree 1 factors. Recall that the roots of $\operatorname{minpoly}(T)$ are the eigenvalues of T. Since T is self-adjoint, by a previous result, all its eigenvalues are real. Thus minpoly has the desired form.

Case 2: $\mathbb{F} = \mathbb{R}$. Then $\operatorname{minpoly}(T)$ factors into a product of degree 1 and degree 2 factors:

$$\operatorname{minpoly}(T) = (z - \lambda_1)\cdots(z - \lambda_m)(z^2 + b_1 z + c_1)\cdots(z^2 + b_N z + c_N)$$

for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ and some $b_1, \ldots, b_N, c_1, \ldots, c_N \in \mathbb{R}$ with $b_k^2 - 4c_k < 0$ for all $k = 1, \ldots, N$. Goal: $N = 0$. Since this is $\operatorname{minpoly}(T)$, then

$$(T - \lambda_1)\cdots(T - \lambda_m)(T^2 + b_1 T + c_1)\cdots(T^2 + b_N T + c_N) = 0\,. \qquad (*)$$

For contradiction, suppose $N > 0$. Then $(T^2 + b_N T + c_N)$ is invertible by the previous result, so multiplying both sides of $(*)$ by its inverse, we have

$$(T - \lambda_1)\cdots(T - \lambda_m)(T^2 + b_1 T + c_1)\cdots(T^2 + b_{N-1} T + c_{N-1}) = 0\,.$$

But this has degree strictly smaller than $\operatorname{minpoly}(T)$, contradiction. Thus $N = 0$. □

Proof of Spectral Theorem over $\mathbb{R}$. We have already seen (b) $\iff$ (c).

(a) $\implies$ (b): Assume T is self-adjoint. Since $\operatorname{minpoly}(T)$ splits into degree 1 factors by the previous result, then there exists an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is upper triangular. Since T is self-adjoint, then

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

Now $([T]_{\mathcal{E}})^t$ is lower triangular, so we must have that $[T]_{\mathcal{E}}$ is diagonal.

(b) $\implies$ (a): Assume there exists an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is diagonal. Then

$$[T]_{\mathcal{E}} = ([T]_{\mathcal{E}})^t = ([T]_{\mathcal{E}})^*$$

so $T = T^*$. Thus T is self-adjoint. □

II.3. **7C: Positive Operators.** Throughout today, let V be an inner product space.

Definition 5. An operator $T \in \mathcal{L}(V)$ is *positive* if

(1) T is self-adjoint; and
(2) $\langle T(v), v\rangle \geq 0$ for all $v \in V$.

Remark 6. These should really be called nonnegative operators. Blame the French!

Definition 7. Let $T \in \mathcal{L}(V)$. A *square root of* T is an operator $R \in \mathcal{L}(V)$ such that $R^2 = T$.

Example 8. Let $T \in \mathcal{L}(\mathbb{F}^3)$ be the operator whose matrix is

$$[T]_{\mathcal{E}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with respect to the standard basis $\mathcal{E}$. Then the operator $R \in \mathcal{L}(\mathbb{F}^3)$ with matrix

$$[R]_{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

is a square root of T. (Exercise.)

Theorem 9. *Let $T \in \mathcal{L}(V)$. TFAE.*

(a) T is a positive operator.
(b) T is self-adjoint and all the eigenvalues of T are nonnegative.
(c) There exists an orthonormal basis $\mathcal{E}$ of V such that $[T]_{\mathcal{E}}$ is diagonal and its diagonal entries are nonnegative.
(d) T has a positive square root.
(e) T has a self-adjoint square root.
(f) $T = R^*R$ for some $R \in \mathcal{L}(V)$.

Proof. (a) $\implies$ (b): Assume T is positive. By definition, then T is self-adjoint. Suppose that λ is an eigenvalue of T and v is a corresponding eigenvector. Then

$$0 \leq \langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2 .$$

Since $\|v\|^2 \geq 0$, then $\lambda \geq 0$.

(b) $\implies$ (c): Assume T is self-adjoint and all its eigenvalues are nonnegative. By [ask students] the Spectral Theorem, then there is an orthonormal basis $\mathcal{E}$ of V consisting of eigenvecors of T. Letting $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues, then

$$[T]_{\mathcal{E}} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} .$$

(c) $\implies$ (d): Assume (c) holds, so there exists an orthonormal basis $\mathcal{E} := (e_1, \ldots, e_n)$ of V of eigenvectors of T such that

$$[T]_{\mathcal{E}} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

and $\lambda_i \geq 0$ for all i. Define $R \in \mathcal{L}(V)$ by

$$R(e_i) = \sqrt{\lambda_i} e_i$$

for all $i = 1, \ldots, n$. Then (exercise) R is positive and $R^2 = T$.

(d) $\implies$ (e): A positive operator is self-adjoint by definition.

(e) $\implies$ (f): Assume T has a self-adjoint square root R. Since R is self-adjoint, then

$$T = R^2 = RR = R^* R\,.$$

(f) $\implies$ (a): Assume $T = R^* R$ for some $R \in \mathcal{L}(V)$. Then

$$T^* = (R^* R)^* = R^* (R^*)^* = R^* R = T$$

so T is self-adjoint. Moreover, given $v \in V$, we have

$$\begin{aligned} \langle T(v), v \rangle &= \langle R^* R(v), v \rangle = \overline{\langle v, R^* R(v) \rangle} = \overline{\langle R(v), R(v) \rangle} = \langle R(v), R(v) \rangle \\ &= \|R(v)\|^2 \geq 0\,. \end{aligned}$$

Thus T is positive. □

Proposition 10. *Let $T \in \mathcal{L}(V)$ be positive. Then T has a unique* positive *square root.*

Remark 11. T can have infinitely many (necessarily not positive) square roots! But only one is positive.

Proof of Proposition. Omitted; see 7.39 of textbook.

□

This allows us to specify a unique square root, namely the only positive one.

Definition 12. Let $T \in \mathcal{L}(V)$ be a positive operator. Then $\sqrt{T}$ denotes the unique positive square root of T.

Remark 13. The proof of part (c) of the theorem shows how to define the positive square root of a linear operator.

Proposition 14. *Suppose that $T \in \mathcal{L}(V)$ is positive and $v \in V$ with $\langle T(v), v \rangle = 0$. Then $T(v) = 0$.*

Proof.

$$0 = \langle T(v), v \rangle = \langle \sqrt{T}\sqrt{T}(v), v \rangle = \langle \sqrt{T}(v), \sqrt{T}(v) \rangle = \|\sqrt{T}(v)\|^2\,,$$

so $\sqrt{T}(v) = 0$. Then

$$T(v) = \sqrt{T}\sqrt{T}(v) = \sqrt{T}(0) = 0\,.$$

□

II.4. **7D: Isometries, Unitary Operators, and Matrix Factorizations.** An isometry is a norm-preserving map.

Definition 15. A linear map $S \in \mathcal{L}(V, W)$ is an *isometry* if

$$\|S(v)\| = \|v\|$$

for all $v \in V$.

Lemma 16. *Isometries are injective.*

Proof. Exercise. $\square$

Theorem 17. *Let $S \in \mathcal{L}(V, W)$, and let $\mathcal{E} := (e_1, \dots, e_n)$ and $\mathcal{F} := (f_1, \dots, f_m)$ be orthonormal bases for V and W, respectively. TFAE.*

(a) *S is an isometry.*
(b) *$S^*S = I$.*
(c) *S preserves inner products, i.e.,*

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

for all $u, v \in V$.
(d) *$S(e_1), \dots, S(e_n)$ is an orthonormal list in W.*
(e) *The columns of ${}_{\mathcal{F}}[S]_{\mathcal{E}}$ form an orthonormal list in $\mathbb{F}^m$ with respect to the usual inner product.*

Lemma 18. *Let $T \in \mathcal{L}(V)$ be self-adjoint. If $\langle T(v), v \rangle = 0$ for all $v \in V$, then $T = 0$.*

Proof. Given $v \in V$, let $u = v + T(v)$. Then

$$\begin{aligned} 0 = \langle T(u), u \rangle &= \langle T(v + T(v)), v + T(v) \rangle = \langle T(v) + T^2(v), v + T(v) \rangle \\ &= \cancelto{0}{\langle T(v), v \rangle} + \langle T(v), T(v) \rangle + \langle T^2(v), v \rangle + \langle T^2(v), T(v) \rangle \\ &= \langle T(v), T(v) \rangle + \langle T(v), T(v) \rangle + \cancelto{0}{\langle T(T(v)), T(v) \rangle} = 2\|T(v)\|^2 . \end{aligned}$$

Thus $T(v) = 0$. Since v was arbitrary, then $T = 0$. $\square$

Proof of Theorem. (a) $\implies$ (b): Assume S is an isometry. Given $v \in V$, then

$$\langle (I - S^*S)(v), v \rangle = \langle v, v \rangle - \langle S^*S(v), v \rangle = \|v\|^2 - \langle S(v), S(v) \rangle = \|v\|^2 - \|S(v)\|^2 = \|v\|^2 - \|v\|^2 = 0 .$$

By the lemma, then $I - S^*S = 0$, so $S^*S = I$.

(b) $\implies$ (c): Assume $S^*S = I$. Given $u, v \in V$, then

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

(c) $\implies$ (d): Assume $\langle S(u), S(v) \rangle = \langle u, v \rangle$ for all $u, v \in V$. Then

$$\langle S(e_j), S(e_k) \rangle = \langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{cases}$$

for each $j, k \in \{1, \dots, n\}$.

(d) $\implies$ (e): Assume $S(e_1), \dots, S(e_n)$ is an orthonormal list. Let $A = {}_{\mathcal{F}}[T]_{\mathcal{E}}$. Then

$$\langle A_{\cdot,j}, A_{\cdot,k}\rangle = \sum_{i=1}^{m} A_{i,j}\overline{A_{i,k}} = \left\langle \sum_{i=1}^{m} A_{i,j} f_i, \sum_{i=1}^{m} A_{i,k} f_i \right\rangle = \langle S(e_j), S(e_k)\rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where the second equality follows from the Pythagorean Theorem.

(e) $\implies$ (a): Assume the columns of ${}_{\mathcal{F}}[S]_{\mathcal{E}}$ form an orthonormal list. Given $v \in V$, then

$$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n\,,$$

so

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$

by the Pythagorean Theorem. By a similar calculation to (19), then $S(e_1), \dots, S(e_n)$ is an orthonormal list. Then

$$S(v) = S(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n) = \langle v, e_1\rangle S(e_1) + \cdots + \langle v, e_n\rangle S(e_n)$$

so

$$\|S(v)\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 = \|v\|^2\,.$$

Thus $\|S(v)\| = \|v\|$, so S is an isometry. □

Definition 20. An operator $S \in \mathcal{L}(V)$ is *unitary* if S is an invertible isometry.

Theorem 21. *Let $S \in \mathcal{L}(V)$, and let $\mathcal{E} := (e_1, \dots, e_n)$ be an orthonormal basis of V. TFAE.*

(a) S is a unitary operator.
(b) $S^*S = SS^* = I$.
(c) S is invertible and $S^{-1} = S^*$.
(d) $S(e_1), \dots, S(e_n)$ is an orthonormal basis of V.
(e) The rows of $[S]_{\mathcal{E}}$ form an orthonormal basis of $\mathbb{F}^n$.
(f) S^* is a unitary operator.

Proof. Similar to the previous theorem. See text for details. □

Proposition 22. *Let $S \in \mathcal{L}(V)$ be a unitary operator and suppose λ is an eigenvalue of S. Then $|\lambda| = 1$.*

Proof. Let $0 \neq v \in V$ be a corresponding eigenvector. Then

$$|\lambda|\|v\| = \|\lambda v\| = \|S(v)\| = \|v\|\,.$$

Since $v \neq 0$, then $\|v\| \neq 0$. Dividing, then $|\lambda| = 1$. □

Definition 23. A matrix $Q \in M_{n\times n}(\mathbb{F})$ is *unitary* if the associated linear operator

$$\begin{aligned} L_Q : \mathbb{F}^n &\to \mathbb{F}^n \\ v &\mapsto Qv \end{aligned}$$

is unitary. Equivalently, if the columns of Q form an orthonormal basis of $\mathbb{F}^n$.

Theorem 24 (QR Factorization)**.** *Suppose $A \in M_{n\times n}(\mathbb{F})$ is a square matrix with linearly independent columns. Then there exist unique matrices $Q, R \in M_{n\times n}(\mathbb{F})$ such that*

(i) Q is unitary;
(ii) R is upper triangular with positive diagonal entries; and

(iii) $A = QR$.

Proof. This follows from a matrix interpretation of the Gram-Schmidt procedure. Let $v_1, \ldots, v_n$ be the columns of A. Let $e_1, \ldots, e_n$ be the orthonormal list resulting from the Gram-Schmidt procedure, and let Q be the matrix whose columns are $e_1, \ldots, e_n$. The equations

$$\begin{aligned} f_k &:= v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, f_j \rangle}{\|f_j\|^2} f_j \\ e_k &:= \frac{1}{\|f_k\|} f_k \end{aligned}$$

give the entries for the upper triangular matrix R^{-1} such that $Q = AR^{-1}$. Details left as an exercise. □