Lecture 18: Normal operators, the spectral theorems, isometries, and positive operators (1)

Travis Schedler

Tue, Nov 16, 2010 (version: Tue, Nov 16, 4:00 PM)

Goals (2)

- Normal operators and the spectral theorem
  - Nice corollaries (slides (7)–(10)) which I plan to skip, but you should study!
- Isometries
- Positive operators
- Polar decomposition

Spectral theorem for self-adjoint operators (3)

From now on, all our vector spaces are finite-dimensional inner product spaces.

Theorem 1 (Theorem 7.13+)**.** *T is self-adjoint iff T admits an orthonormal eigenbasis with* real *eigenvalues.*

Proof.

- Proof for $\mathbf{F} = \mathbf{C}$: we already know that $\mathcal{M}(T)$ is upper-triangular in some orthonormal basis.
- Then, $T = T^*$ iff the matrix equals its conjugate transpose, i.e., it is upper-triangular with real values on the diagonal.
- Now let $\mathbf{F} = \mathbf{R}$. In some orthonormal basis, the matrix is block upper-triangular with 1×1 and 2×2 blocks.
- Then, the matrix equals its own transpose iff it is block diagonal with real diagonal entries and symmetric 2×2 blocks.
- However, in slide (6) we show that the 2×2 blocks are anti-symmetric. So there are none. □

Spectral theorem for complex normal operators (4)

Motivation: Which T admit an orthonormal eigenbasis but not necessarily with real eigenvalues?

Definition 2. An operator T is *normal* if $TT^* = T^*T$, i.e., T and T^* commute.

Theorem 3 (Theorem 7.9)**.** *Let $\mathbf{F} = \mathbf{C}$. Then $T \in \mathcal{L}(V)$ admits an orthonormal eigenbasis iff it is normal.*

Proof.

- Pick an orthonormal basis so that $A := \mathcal{M}(T)$ is upper-triangular. Then T is normal iff $A\overline{A}^t = \overline{A}^t A$.
- In coordinates $A = (a_{jk})$ (with $a_{jk} = 0$ for $j > k$), this means $|a_{jj}|^2 + \cdots + |a_{jn}|^2 = |a_{jj}|^2, \ \forall j. \quad (\dim V = n)$
- This is equivalent to: $a_{jk} = 0$ for $j < k$. So T is normal iff A is diagonal. □

Spectral theorem for real normal operators (5)

Motivation: What does it mean for a real operator to be normal?

Theorem 4 (Theorem 7.25)**.** *Let $\mathbf{F} = \mathbf{R}$. Then T is normal iff it admits an orthonormal basis in which $\mathcal{M}(T)$ is block-diagonal with blocks (λ_j) or $\begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}$.*

The complex eigenvalues are $\lambda_j \in \mathbf{R}$ and $a_j \pm i\, b_j$.

Proof.

- Pick an orthonormal basis so that $A = \mathcal{M}(T)$ is block upper-triangular.
- Then, T is normal iff $A\overline{A}^t = \overline{A}^t A$.
- For the rows with 1×1 blocks, this again means $|a_{jj}|^2 + \cdots + |a_{jn}|^2 = |a_{jj}|^2$, i.e., $a_{j,j+1} = \cdots = a_{jn} = 0$.
- For rows $j, j+1$ with 2×2 blocks, adding the corresponding sums for both rows, this implies $\sum_{k=j+2}^{n} |a_{j,k}|^2 + |a_{j+1,k}|^2 = 0$, i.e., $a_{j,k} = a_{j+1,k} = 0$ for $k > j+1$, so A is block diagonal.
- Finally, we apply the following proposition to the blocks. □

2×2 **case (6)**

We need just one final detail ($\mathbf{F} = \mathbf{R}$ and $\dim V = 2$):

Proposition 0.1 (Lemma 7.15, essentially)**.** *Suppose that $T \in \mathcal{L}(V)$ is normal and that T has no eigenvalues. Then, in any orthonormal basis, $\mathcal{M}(T) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, where $a \pm bi$ are the roots of the characteristic polynomial of T.*

Recall that for two-by-two matrices A, the characteristic polynomial is $x^2 - (\mathrm{tr}\, A)x + \det A$, and this does not depend on the choice of basis so makes sense for T.

Proof.

- Write $\mathcal{M}(T) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in the orthonormal basis.
- Since T is normal, $|a|^2 + |b|^2 = |a|^2 + |c|^2$, so $b = \pm c$.
- Since there are no real eigenvalues of $\mathcal{M}(T)$, $b = -c \neq 0$.
- Since T is normal, $ac + bd = ab + cd$, so $(d-a)b = (a-d)b$. So $a = d$. □

Corollaries (7)

Corollary 5 (Corollary 7.8)**.** *If T is normal, then eigenvectors u, v of distinct eigenvalues are orthogonal.*

- Proof: In an orthonormal eigenbasis (e_j) so that $\mathcal{M}(T)$ is (block) upper-triangular, an eigenvector v of eigenvalue λ is a linear combination of the e_j with the same eigenvalue.
- So u, v cannot have nonzero coefficients of the same e_j, i.e., $u \perp v$.

Corollary 6 (Corollary 7.7)**.** *Let T be normal. If v is an eigenvector of T of eigenvalue λ, then it is also an eigenvector of T^* of eigenvalue $\overline{\lambda}$.*

- Proof: Again, v must be a linear combination of the e_j that are eigenvectors of eigenvalue λ.
- Since $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)}^t$, these e_j are eigenvectors of T^* of eigenvalue $\overline{\lambda}$.

A characterization of normal operators (8)

Proposition 0.2 (Proposition 7.4)**.** *If T is self-adjoint, then $\langle Tv, v\rangle = 0$ for all v iff $T = 0$.*

- Assume $\langle Tv, v\rangle = 0$ for all v. Then $\langle T(v+u), v+u\rangle - \langle T(v), v\rangle - \langle T(u), u\rangle = 0$ for all v, u.
- Thus, $\langle T(v), u\rangle + \langle T(u), v\rangle = 0$ for all u, v.
- When T is self-adjoint, this says $2\Re\langle T(v), u\rangle = 0$ for all u, v.
- Plugging in iu for u, also $2\Im\langle T(v), u\rangle = 0$ for all u, v. Thus $T(v) = 0$ for all v, i.e., $T = 0$.

Corollary 7 (Proposition 7.6)**.** *An operator T is normal iff $\|Tv\| = \|T^*v\|$ for all v.*

- $\|Tv\| = \|T^*v\| \Leftrightarrow \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \Leftrightarrow \langle (T^*T - TT^*)v, v\rangle = 0$.
- Since $T^*T - TT^*$ is self-adjoint, by the corollary, the last condition is satisfied for all v iff T is normal.

$\langle Tv, v\rangle = 0$ and anti-self-adjoint operators on R (9)

Proposition 0.3 (Proposition 7.2+)**.** *$\langle Tv, v\rangle = 0$ for all v iff either $T = 0$, or $\mathbf{F} = \mathbf{R}$ and $T = -T^*$ (T is anti-self-adjoint).*

Note: anti-self-adjoint ($T = -T^*$) implies normal.

Proof.

- Take an orthonormal basis (e_j) in which $A = (a_{jk}) = \mathcal{M}(T)$ is (block) upper-triangular.
- We claim $a_{jj} = 0$ for all j. Indeed, $\langle Te_j, e_j\rangle = a_{jj} = 0$.
- It remains only to show that A is block diagonal (since then the blocks are antisymmetric).
- Otherwise, if $a_{jk} \neq 0$ above the block diagonal, then $\langle T(e_j + \lambda e_k), e_j + \lambda e_k\rangle = \langle (a_{jj} + \lambda a_{jk})e_j, e_j\rangle = a_{jj} + \lambda a_{jk}$. For $\lambda \neq -a_{jj}a_{jk}^{-1}$, this is nonzero. Contradiction. □

Anti-self-adjoint operators for $\mathbf{F} = \mathbf{C}$ (10)

Proposition 0.4. *For $\mathbf{F} = \mathbf{C}$, $T = -T^*$ if and only if T has an orthonormal eigenbasis with purely imaginary eigenvalues (i.e., eigenvalues in $i \cdot \mathbf{R}$).*

Proof.

- In an orthonormal basis in which $\mathcal{M}(T)$ is upper-triangular, $T = -T^*$ means the matrix equals its negative conjugate transpose.
- This means it is diagonal with purely imaginary diagonal entries. □

Alternatively: anti-self-adjoint operators are normal, so admit an orthonormal eigenbasis; then $\langle Tv, v\rangle = \lambda\langle v, v\rangle = \langle v, T^*v\rangle = -\overline{\lambda}\langle v, v\rangle$ implies that $\lambda = -\overline{\lambda}$ for all eigenvalues λ. So they are purely imaginary.

Complex eigenvalues of real operators (11)

We have often spoken about complex eigenvalues of $\mathcal{M}(T)$ when $\mathbf{F} = \mathbf{R}$. Let's formalize it:

Definition 8. Let $\mathbf{F} = \mathbf{R}$. Then the *complex eigenvalues* of T are the complex eigenvalues of $\mathcal{M}(T)$ in any basis.

Why do these not depend on the basis? The change of basis formula! We know that conjugate matrices A and SAS^{-1} have the same (complex) eigenvalues (directly, or by using complex operators).

Example 9. The complex eigenvalues of any T such that $\mathcal{M}(T) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ are $a + bi$ and $a - bi$.

- Note: $|a + bi| = 1 = |a - bi|$ iff it is a rotation matrix, i.e., $a = \cos\theta$ and $b = \sin\theta$ for some angle θ.

Isometries (12)

Let $T \in \mathcal{L}(V, W)$, with V and W inner product spaces.

Definition 10. An isometry is an operator such that $\langle u, v\rangle = \langle Tu, Tv\rangle$ for all $u, v \in V$.

That is, isometries are operators that preserve the inner product.

Proposition 0.5. *T is an isometry iff it preserves merely the norm: $\|Tv\| = \|v\|$ for all v.*

Proof. The inner product is given by a formula from the norm (on the homework), so preservation of the norm implies preservation of inner product. The converse is obvious. □

Characterization of isometries (13)

Let $V = W$ be finite-dimensional. **Useful characterization:** $T \in \mathcal{L}(V)$ is an isometry iff $T^*T = I = TT^*$. (In particular isometries of V are invertible!)

Theorem 11 (Theorem 7.37)**.** *Isometries are the same as normal operators whose complex eigenvalues all have absolute value one.*

Proof.

- First, isometries are normal by the characterization.
- Given a normal operator, pick an orthonormal basis as in the spectral theorem. Then $\mathcal{M}(T)\mathcal{M}(T^*) = I$ iff $|\lambda_i|^2 = 1$ and $|a_i|^2 + |b_i|^2 = 1$ for all i. □

That is, in our usual orthonormal basis, an isometry has blocks which are either numbers of absolute value one, or rotation matrices.

Positive operators (14)

Definition 12. A positive operator T is a self-adjoint operator such that $\langle Tv, v\rangle \geq 0$ for all v.

In view of the spectral theorem, a self-adjoint operator is positive iff its eigenvalues are nonnegative (part of Theorem 7.27).

Theorem 13 (Remainder of Theorem 7.27)**.**

(i) Every operator of the form $T = S^*S$ is positive.

(ii) Every positive operator admits a positive square root.

Proof.

- (i) First, $T^* = (S^*S) = S^*S$ is self-adjoint.
- Next, $\langle Tv, v\rangle = \langle Sv, Sv\rangle \geq 0$ for all v.
- (ii) For any orthonormal eigenbasis of T, let $\sqrt{T}$ be the operator with the same orthonormal eigenbasis, but with the nonnegative square root of the eigenvalues. □

Polar decomposition (15)

After the spectral theorem, the second-most important theorem of Chapters 6 and 7 is:

Theorem 14 (Polar decomposition: Theorem 7.41)**.** *Every $T \in \mathcal{L}(V)$ equals $S\sqrt{T^*T}$ for some isometry S.*

Main difficulty: T need not be invertible!

Lemma 15. *For all $v \in V$, $\|Tv\| = \|\sqrt{T^*T}v\|$.*

Proof. $\|\sqrt{T^*T}v\|^2 = \langle\sqrt{T^*T}v, \sqrt{T^*T}v\rangle = \langle(\sqrt{T^*T})^*\sqrt{T^*T}v, v\rangle = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$. □

Corollary: $\text{null}(T) = \text{null}(\sqrt{T^*T})$. We may thus define $S_1 : \text{range}(\sqrt{T^*T}) \xrightarrow{\sim} \text{range}(T)$ by $S_1(\sqrt{T^*T}v) = Tv$. Thus, for all $v \in V$, $S_1\sqrt{T^*T}v = Tv$. Also, $\|S_1u\| = \|u\|$ for all $u (= \sqrt{T^*T}v)$ by the lemma. So S_1 is an isometry.

Completion of proof (16)

- We only have to extend S_1 to an isometry on all of V.
- Note that $\text{range}(\sqrt{T^*T}) \oplus \text{range}(\sqrt{T^*T})^\perp = V = \text{range}(T) \oplus \text{range}(T)^\perp$.
- Thus, the extensions of $S_1 : \text{range}(\sqrt{T^*T}) \xrightarrow{\sim} \text{range}(T)$ to an isometry $S : V \xrightarrow{\sim} V$ are exactly $S = S_1 \oplus S_2$, where $S_2 : \text{range}(\sqrt{T^*T})^\perp \xrightarrow{\sim} \text{range}(T)^\perp$ is an isometry.
- Since these are inner product spaces of the same dimension, there always exists an isometry, by taking an orthonormal basis to an orthonormal basis.

Recall here that $T_1 \oplus T_2$ on $U_1 \oplus U_2$ means $(T_1 \oplus T_2)(u_1 + u_2) = T_1(u_1) + T_2(u_2)$, $\forall u_1 \in U_1, u_2 \in U_2$.