Lecture 17: The spectral theorem for normal operators; isometries; positive operators; preview of polar decomposition (1)

Travis Schedler
Thurs, Nov 10, 2011 (version: Thurs, Nov 10, 1:00 PM)

Goals (2)

- Unitary matrices, isometries, and the spectral theorem
- Normal operators and the spectral theorem
- Positive operators
- Preview: Polar decomposition and singular value decomposition

Read Chapter 7 and do PS 8.

Symmetric and hermitian matrices (3)

- Definition: A symmetric matrix $A$ is one such that $A = A^t$.
- A Hermitian matrix $A$ is one such that $A = \bar{A}^t$.
- If $A$ is real, then Hermitian and symmetric are the same.
- By yesterday: $T_A$ is self-adjoint if and only if $A$ is Hermitian.

Warm-up exercise (4)

Goal: find the orthonormal eigenbasis and eigenvalues of the Hermitian matrix $A = \begin{pmatrix} -23 & 36 \\ 36 & -2 \end{pmatrix}$.

(a) First, verify that $A$ is real symmetric, and hence Hermitian. Conclude that $T_A$ is self-adjoint.

(b) Compute the eigenvalues $\lambda_1$ and $\lambda_2$, using the characteristic polynomial $x^2 - \text{tr}(A)x + \det(A)$ (they are always the roots of this polynomial). Hint: Factor the polynomial into integer-coefficient factors!
(c) Using the answer to (b) (I will give you this), compute the orthonormal eigenbasis \((v_1, v_2)\) of eigenvalues \((\lambda_1, \lambda_2)\). Hint: You only need to compute one norm-one eigenvector; the other one is the unique norm-one vector orthogonal to this up to scaling.

Conclusion: \(A = (v_1v_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} (v_1v_2)^{-1}\). We don’t have to multiply (but you can to double-check!)

Solution to warm-up exercise (5)

(a) All coefficients of \(A\) are real and \(A^t = A\). (The two off-diagonal entries are equal.)

(b) \(\text{tr}(A) = -23 - 2 = -25\); \(\det(A) = -23 \cdot -2 - 36 \cdot 36 = 46 - 1296 = -1250\); char. poly is \(x^2 + 25x - 1250\). To factor as \((x - a)(x - b)\), we want \(ab = -1250\) and \(a + b = 25\). Note 1250 = 50 * 25, so we can take \(a = 25, b = -50\). So \(x^2 + 25x - 1250 = (x - 25)(x + 50)\). Eigenvalues: 25 and -50. (Note: the quadratic formula would give \(\frac{-25 \pm \sqrt{25^2 + 5000}}{2} = \frac{-25 \pm 75}{2}\), i.e., 25 and -50.)

Solution of warm-up continued (6)

(c) Now we need to find the null space of \(A - \lambda_1 I\) or \(A - \lambda_2 I\), for \(\lambda_1 = 25\) and \(\lambda_2 = -50\). Take the first: we get \(A - 25I = \begin{pmatrix} -48 & 36 \\ 36 & -27 \end{pmatrix}\). Since the nullspace is nonzero it has rank one (or note that the two rows are multiples of each other, or equivalently the two columns are multiples of each other). The nullspace is evidently spanned by \(\begin{pmatrix} 36 \\ 48 \end{pmatrix}\) or equivalently\(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\) if we divide by 12. The norm of this is \(\sqrt{3^2 + 4^2} = 5\) so a norm-one eigenvector is \(v_1 := \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}\).

The other eigenvector is orthogonal to this one, so it must be, up to scaling,\(v_2 := \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix}\) (which has norm one). We get the orthonormal eigenbasis \((v_1, v_2) = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix}\).

Unitary matrices (7)

• Take the matrix \(B = \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix}\) from before. We have \(A = BDB^{-1}\), for \(D = \begin{pmatrix} 25 & 0 \\ 0 & -50 \end{pmatrix}\).
• What is the inverse of $B$? Answer: $B^{-1} = \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}$. Why?

• In general, if $U = (e_1 \cdots e_n) \in \text{Mat}(n, 1, \mathbb{F})$ with $(e_1, \ldots, e_n)$ orthonormal, then $U^*U = I$, because these entries are the dot products. So $U^t = U^{-1}$!

Definition 1. A unitary matrix is a square matrix whose columns are orthonormal.

Equivalently, this is $U$ such that $U^{-1} = U^t$. (Equivalently, the rows are orthonormal.) A real unitary matrix is also called real orthogonal. (Orthogonal means $A^t = A^{-1}$.)

Isometries (8)

Definition 2. Let $V$ and $W$ be i.p.s. Then $T \in \mathcal{L}(V, W)$ is an isometry if $\langle v_1, v_2 \rangle = \langle Tv_1, Tv_2 \rangle$ for all $v_1, v_2 \in V$.

Proposition 0.1. Assume $T^*$ exists. $T$ is an isometry if and only if $T^*T = I$.

Proof: Using on. bases of f.d. vs., $M(T)$ will be unitary!

Proposition 0.2. Let $V, W$ be f.d. i.p.s. Then $T$ is an isometry if and only if $T$ takes an orthonormal basis of $V$ to an orthonormal basis of $W$.

Proof: $\Rightarrow$ is clear; the other direction is because, for every orthonormal $(e_i)$, $\langle \sum_i a_i e_i, \sum_i b_i e_i \rangle = \sum_i a_i \bar{b}_i$.

Orthogonal complements of invariant subspaces (9)

Why did the spectral theorem for self-adjoint operators work?

Proposition 0.3. (a) If $U \subseteq V$ is $T$-invariant, then $U^\perp$ is $T^*$-invariant.

(b) If $T$ is moreover self-adjoint or an isometry, then $U^\perp$ is $T$-invariant.

Proof: (a) If $u \in U$ and $v \in U^\perp$, then $0 = \langle Tu, v \rangle = \langle u, T^*v \rangle$. So $T^*v \in U^\perp$.

(b) In the self adjoint case, this is immediate from (a). In the isometry case, $T^{-1}(U^\perp) \subseteq U^\perp$ implies $T^{-1}(U^\perp) = U^\perp$, since $T^{-1}$ is an isomorphism. Then $T(U^\perp) = U^\perp$ as well.

Alternative proof of (b) in isometry case: since $U$ and $U^\perp$ are orthogonal complements, so also are $T(U)$ and $T(U^\perp)$. So $T(U) = U$ iff $T(U^\perp) = U^\perp$. 

3
Spectral theorem for isometries and self-adjoint operators over $\mathbb{C}$

**Corollary 3.** Let $\mathbf{F} = \mathbb{C}$ and $V$ be f.d. Then if $T$ is self-adjoint or an isometry, $V$ admits an orthonormal eigenbasis.

Proof: By induction on $\dim V$; the case $\dim V = 1$ is trivial.

- There exists a nonzero eigenvector $v \in V$; rescale it to have norm one.
- Since $\text{Span}(v)$ is $T$-invariant, so is $\text{Span}(v)^\perp$.
- Apply the inductive hypothesis to $\text{Span}(v)^\perp$, getting an orthonormal eigenbasis.
- Put it together with $v$. Get: orthonormal eigenbasis of $V$.

Note: we proved this already in the self-adjoint case, but quite differently.

**Biconditional for isometries (11)**

Let $V$ be f.d. over $\mathbb{C}$.

**Theorem 4.** $T \in \mathcal{L}(V)$ is an isometry if and only if $V$ has an orthonormal eigenbasis with eigenvalues of absolute value one.

Proof: Suppose $T$ is an isometry. We need: eigenvalues have absolute value one. If $Tv = \lambda v$ with $v \neq 0$, then $\|v\| = \|Tv\| = |\lambda|\|v\|$, so $|\lambda| = 1$.

Conversely, if $T$ has an orthonormal eigenbasis $(e_1, \ldots, e_n)$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, then $T^*e_i = \lambda_i e_i$, since $\langle e_i, T^*e_j \rangle = \langle \lambda_i e_i, e_j \rangle = \lambda_i \delta_{ij}$. So $T^*T = I$ if and only if $|\lambda|^2 = 1$.

**Spectral theorem for complex normal operators (12)**

Motivation: Common generalization of isometry and self-adjoint: Which $T$ admit an orthonormal eigenbasis but not necessarily with real or absolute value one eigenvalues?

**Definition 5.** An operator $T$ is normal if $TT^* = T^*T$, i.e., $T$ and $T^*$ commute.


**Theorem 6** (Theorem 7.9). Let $\mathbf{F} = \mathbb{C}$. Then $T \in \mathcal{L}(V)$ admits an orthonormal eigenbasis iff it is normal.

**Proof of Theorem 7.9 (13)**

**Theorem 7** (Theorem 7.9). Let $\mathbf{F} = \mathbb{C}$. 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 $A^T = A$. 
• In coordinates $A = (a_{jk})$ (with $a_{jk} = 0$ for $j > k$), this means $|a_{jj}|^2 + \cdots + |a_{jn}|^2 = |a_{1j}|^2 + \cdots + |a_{jj}|^2, \forall j.$ (dim $V = n$)

• By induction on $j$, this implies $a_{jk} = 0$ for all $k > j$, i.e., $A$ is diagonal.

• Conversely, if $A$ is diagonal, clearly $AA^t = A^t A$. \hfill \Box

**Normal real operators: $2 \times 2$ case (14)**

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

**Proposition 0.4** (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 $x^2 - (\text{tr } A)x + \det A$, 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 $a = d$. \hfill \Box

**Spectral theorem for real normal operators (15)**

**Theorem 8** (Theorem 7.25). Let $F = \mathbb{R}$. Then $T$ is normal iff it admits an orthonormal basis in which $\mathcal{M}(T)$ is block-diagonal with blocks $(\lambda_j)$ or $(a_j, -b_j; b_j, a_j)$.

As we saw before, the complex eigenvalues are the eigenvalues of the blocks, which are $\lambda_j \in \mathbb{R}$ and $a_j \pm ib_j$.

**Proof of Theorem 7.25 (16)**

• Pick an orthonormal basis so that $A = \mathcal{M}(T)$ is block upper-triangular.

• Then, $T$ is normal iff $AA^t = A^t A$.

• We claim that this implies $A$ is block diagonal with blocks as described. The converse is immediate.

• We prove that entries to the right of each diagonal block are zero by induction. Assume true for all blocks in the first $j - 1$ rows.
• If the $j$-th diagonal entry is a $1 \times 1$ block, comparing this entry of $AA^t = A^t A$, we get $|a_{jj}|^2 + \cdots + |a_{jn}|^2 = |a_{jj}|^2$. So $a_{jk} = 0$ for $k > j$.

• If the $j$-th entry is the upper-left corner of a $2 \times 2$ block, summing the $j$-th and $j+1$-th entries of $AA^t = A^t A$, we get $\sum_{k=j+2}^n |a_{jk}|^2 + |a_{j+1,k}|^2 = 0$, i.e., $a_{j,k} = a_{j+1,k} = 0$ for $k > j + 1$.

• So $A$ is block diagonal. Then we apply the preceding proposition to the blocks to see they are as claimed.

Real isometries (17)

**Theorem 9.** Let $V$ be a f.d. real i.p.s. Then $T \in \mathcal{L}(V)$ is an isometry iff it has an orthonormal basis in which $M(T)$ is block diagonal with all blocks either $(\pm 1)$, or $2 \times 2$-rotation matrices.

**Proof.**

- $T$ is normal iff it has an orthonormal basis in which $M(T)$ is block diagonal with blocks $(\lambda)$ for $\lambda$ a real eigenvalue, or $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ for $a \pm bi$ complex eigenvalues.

- $T$ is an isometry iff it is normal with eigenvalues of absolute value one, i.e., $\lambda = \pm 1$ and $a^2 + b^2 = 1$ above. \hfill $\square$

Positive operators (18)

Idea: complete (physical) analogy: Real numbers::self-adjoint operators = Nonnegative numbers::positive operators.

**Definition 10.** A positive operator is a self-adjoint operator such that $\langle Tv, v \rangle \geq 0$ for all $v$.

For f.d. case, equivalently: self-adjoint with all nonnegative eigenvalues; equivalently, has on. eigenbasis with nonnegative eigenvalues.

[Physics: Observables are self-adjoint ops; pure states of observable $\lambda$ are eigenvectors of eigenvalue $\lambda$; positive operators are nonnegative observables.]

Preview: polar decomposition and SVD (19)

Proposition: every complex number $z$ is expressible as $z = r \cdot e^{i\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. (Unique if $z$ nonzero). Equivalently: $z = s \cdot r$, for $|s| = 1$ and $r = |z| = \sqrt{-z \cdot \bar{z}} \geq 0$.

**Theorem 11.** Let $V$ be a f.d. i.p.s. and $T \in \mathcal{L}(V)$. Then there is an expression $T = SP$, for $S$ an isometry and $P$ positive. Moreover, $P = \sqrt{T^* T}$ (i.e., $P^2 = T^* T$). This expression is unique if $T$ is invertible.

**Corollary 12** (Singular Value Decomposition (SVD)). There exists orthonormal bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ of $V$ such that $Te_i = s_i f_i$, for $s_i \geq 0$ the singular values. Moreover, $(e_1, \ldots, e_n)$ is an orthonormal eigenbasis of $T^* T$ with eigenvalues $s_i^2$. 

6
Proof: Let \((e_1, \ldots, e_n)\) be an orthonormal eigenbasis of \(T^*T\) and \(s_1, \ldots, s_n\) the square roots of the eigenvalues. When \(s_i \neq 0\), set \(f_i := s_i^{-1}T e_i\). Then extend the resulting \(f_i\) to an orthonormal eigenbasis.