Math 108b: Notes on the Spectral Theorem

From section 6.3, we know that every linear operator T on a finite dimensional inner product space V has an adjoint. (T^* is defined as the *unique* linear operator on V such that $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ for every $x, y \in V$ – see Theroems 6.8 and 6.9.) When V is infinite dimensional, the adjoint T^* may or may not exist.

One useful fact (Theorem 6.10) is that if β is an *orthonormal* basis for a finite dimensional inner product space V, then $[T^*]_\beta = [T]^*_\beta$. That is, the matrix representation of the operator T^* is equal to the complex conjugate of the matrix representation for T.

For a general vector space V, and a linear operator T, we have already asked the question "when is there a basis of V consisting only of eigenvectors of T?" – this is exactly when T is diagonalizable. Now, for an *inner product* space V, we know how to check whether vectors are orthogonal, and we know how to define the norms of vectors, so we can ask "when is there an *orthonormal* basis of V consisting only of eigenvectors of T?" Clearly, if there is such a basis, T is diagonalizable – and moreover, eigenvectors with distinct eigenvalues must be orthogonal.

DEFINITIONS Let V be an inner product space. Let $T \in \mathcal{L}(V)$.

(a) T is **normal** if $\underline{T^*T = TT^*}$

(b) T is **self-adjoint** if $\underline{T^* = T}$

For the next two definitions, assume V is finite-dimensional: Then,
(c) T is **unitary** if $F = \mathbb{C}$ and $\|T(x)\| = \|x\|$ for every $x \in V$

(d) T is **orthogonal** if $F = \mathbb{R}$ and $\|T(x)\| = \|x\|$ for every $x \in V$

NOTES
1. Self-adjoint $\Rightarrow$ Normal (*Clearly, if $T^* = T$, then T^* and T commute!*)

2. Unitary or Orthogonal $\Rightarrow$ Invertible (*Show that* $N(T) = \{0\}$.)

3. Unitary or Orthogonal $\Rightarrow$ Normal (*This is because, when T is unitary (or orthogonal), we must have $T^*T = TT^* = I$, the identity operator on V – see Theorem 6.18.*)

EXAMPLE A simple example of an orthogonal operator is the rotation operator T_θ on $\mathbb{R}^2$. In the standard basis,

$$A_\theta = [T_\theta]_\beta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Notice that A is orthogonal. Also, $A^* = A^T \neq A$, so A is not self-adjoint. However, it is easy to check that $AA^* = A^*A = I$, so A is normal. If the matrix A were over the complex field, it would have a orthonormal basis of eigenvectors by the spectral theorem (see below)! However, over the reals, we see that the characteristic polynomial of A does not split (unless $\sin\theta = 0$) – therefore, there are no real eigenvalues and therefore no eigenvectors at all for the operator T_θ on $\mathbb{R}^2$.

SPECTRAL THEOREM 1 Let T be a linear operator on a finite dimensional complex inner product space V. Then, T is normal if and only if there exists an orthonormal basis (for V) consisting of eigenvectors of T.

SPECTRAL THEOREM 2 Let T be a linear operator on a finite dimensional real inner product space V. Then, T is self-adjoint if and only if there exists an orthonormal basis (for V) consisting of eigenvectors of T.

We'll prove the simple direction first: Assume that there is an orthonormal basis β consisting of eigenvectors of T. Then, we know that $[T]_\beta$ is diagonal, and also that $[T^*]_\beta = [T]_\beta^*$ is diagonal. Since diagonal matrices commute, we have that

$$[T^*T]_\beta = [T^*]_\beta[T]_\beta = [T]_\beta[T^*]_\beta = [TT^*]_\beta.$$

Hence, $T^*T = TT^*$, so T is normal. This holds whether $F = \mathbb{C}$ or $F = \mathbb{R}$.

We'll see in the lemmas below that T normal implies that every eigenvector of T with eigenvalue λ is also an eigenvector of T with eigenvalue $\bar{\lambda}$. Since β is a basis of eigenvectors of either T or T^*, we have that both $[T]_\beta$ and $[T^*]_\beta$ are diagonal matrices with their eigenvalues on the diagonal. In the case $F = \mathbb{R}$, every eigenvalue of T is real, so $\lambda = \bar{\lambda}$, and we must have that $[T^*]_\beta = [T]_\beta$. Therefore, in the case of a real inner product space, we know that T is not only normal, but also self-adjoint!

LEMMA (pg. 269) Let V be a finite-dimensional inner product space and $T \in \mathcal{L}(V)$. If T has an eigenvector with eigenvalue λ, then T^* has an eigenvector with eigenvalue $\bar{\lambda}$

See the book for the proof. Note that the above lemma is true for any linear operator T – but we do not know that the eigenvectors for T and T^* are related! The next theorem states that when T is normal, we know the eigenvectors for T and T^* are the same – this is the fact we used in the proof of the spectral theorem above.

THEOREM 6.15 (pg. 371) Let V be an inner product space and $T \in \mathcal{L}(V)$ be a normal operator. Then,

(c) If $T(x) = \lambda x$ (for some $x \in V$ and some $\lambda \in F$), then $T^*(x) = \bar{\lambda} x$.

(d) If λ_1 and λ_2 are distinct eigenvalues of T with eigenvectors x_1 and x_2, then $\langle x_1, x_2 \rangle = 0$.

Sketch of proof: First prove that, since T is normal, $\|T(x)\| = \|T^*(x)\|$ for every $x \in V$. Then, since $U = T - \lambda I$ is normal, and $U^* = T^* - \bar{\lambda} I$, we know that $U(x) = 0$ if and only if $U^*(x) = 0$. In other words, $T(x) = \lambda x$ if and only if $T^*(x) = \bar{\lambda} I$.

For (d), Simply compute $\lambda_1 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$. Since $\lambda_1 \neq \lambda_2$, we must have $\langle x_1, x_2 \rangle = 0$.

We're almost ready to finish the proof of the first spectral theorem, but first, we'll need the following theorem:

SCHUR'S THEOREM Let T be a linear operator on a finite-dimensional inner product space V. Suppose the characteristic polynomial of T splits. Then, there exists an orthonormal basis β for V such that the matrix $[T]_\beta$ is upper triangular.

Proof: Induct on $n = \dim V$. Of course, if $n = 1$, take any basis $\beta = \{x\}$ with $\|x\| = 1$. Then, $[T]_\beta = (a)$ (where $a = T(x)$) is upper-triangular. Now, assume the result for $n-1$: that is, suppose that if V is any finite-dimensional inner product space with dimension $n-1$ and if T is any linear operator such that the characteristic polynomial splits, then there exists an orthonormal basis for V such that $[T]_\beta$ is upper triangular. We wish to show this result for n.

Fix an $n-$dimensional inner product space V and a $T \in \mathcal{L}(V)$ such that the characteristic polynomial of T splits. This implies that T has an eigenvalue λ. Therefore, T^* has an eigenvalue $\bar{\lambda}$ (by Lemma, pg. 369). Let z be a unit eigenvector of T^*, so $T^*z = \bar{\lambda} z$. Consider the space $W = \text{span}(\{z\})$. Clearly, $W^\perp$ is $n-1$ dimensional, and $T_{W^\perp}$ is a linear operator on $W^\perp$, so by the induction hypothesis, there exists a basis γ such that $[T_{W^\perp}]_\gamma$ is upper-triangular. (Notice that we know that $T_{W^\perp}$ splits since $W^\perp$ is $T-$invariant! Check this fact below:)

Let $y \in W^\perp$. We want to show that $T(y)$ is in $W^\perp$. For every ________, compute

Finally $\beta = \gamma \cup \{z\}$ is orthonormal, and $[T]_\beta$ is upper-triangular: Letting $[Tz]_\beta = a \in F^n$,

$$[T]_\beta = \begin{pmatrix} [T_{W^\perp}]_\gamma & | \\ & a \\ 0 & | \end{pmatrix}. \qquad \square$$

Proof of Spectral Theorem 1 Assume T is normal. Since $F = \mathbb{C}$, we know (from the fundamental theorem of algebra) that the characteristic polynomial of T splits. By Schur's theorem, we can find an orthonormal basis $\beta = \{v_1, v_2, ..., v_n\}$ such that $A = [T]_\beta$ is upper-triangular. Clearly v_1 is an eigenvector of T (since $T(v_1) = A_{11}v_1$.) The fact that T is normal will imply that v_2, ..., v_n are also eigenvectors!

Assume $v_1, v_2, ..., v_{k-1}$ are all eigenvectors Let $\lambda_1, \lambda_2, ..., \lambda_{k-1}$ be the corresponding eigenvalues. Then, since A is upper-triangular, look at the k^{th} column of A and note $A_{mk} = 0$ if $m > k$: This lets us compute

$$T(v_k) = A_{1k}v_1 + A_{2k}v_2 + ...A_{kk}v_k$$

From Theorem 6.15, the fact that T is normal implies that $T^*(v_j) = \bar{\lambda}_j v_j$. Since the basis is orthonormal, we have the formula $A_{jk} = \langle T(v_k), v_j\rangle$ for the coefficients in the equation above. Computing,

$$A_{jk} = \langle T(v_k), v_j\rangle = \langle v_k, T^*(v_j)\rangle = \left\langle v_k, \bar{\lambda}_j v_j\right\rangle = \lambda_j \langle v_k, v_j\rangle = 0.$$

Therefore, we have simply that $T(v_k) = A_{kk}v_k$, so v_k is an eigenvector for T.

By induction, we have shown that $v_1, v_2, ..., v_n$ are all eigenvectors of T, so β is an orthonormal basis consisting of only eigenvectors of T, and the spectral theorem is proven. □

Before we can prove the second version of the spectral theorem, for $F = \mathbb{R}$, we need the following lemma:

LEMMA (pg. 373) Let T be a self-adjoint operator on a finite-dimensional inner product space V. Then the following two facts hold (*whether we have $F = \mathbb{R}$ or $F = \mathbb{C}$*)

(a) Every eigenvalue of T is real.

(b) The characteristic polynomial of T splits.

Proof of (a): From Theorem 6.15, if x is an eigenvalue of T, we have both $T(x) = \lambda x$ for some $\lambda \in F$ and $T^*(x) = \bar{\lambda}x$. However, since T is self-adjoint, $T(x) = T^*(x)$. Therefore, $\lambda x = \bar{\lambda}x$. Since $x \neq 0$, we have that $\lambda = \bar{\lambda}$; i.e., λ is real.

Proof of (b): When $F = \mathbb{C}$, we already know the characteristic polynomial splits. Let β be any orthonormal basis for V; we know that $A = [T]_\beta$ satisfies $A^* = A$. Define, for every $x \in \mathbb{C}^n$,

$$S(x) = Ax.$$

S has the same characteristic polynomial as T, but since S is a linear operator on a complex inner product space, we know the characteristic polynomial of S splits. Moreover, since S is self-adjoint, each eigenvalue $\lambda_1, ..., \lambda_n$ is real, and the polynomial $p_T(t) = p_S(t) = (\lambda_1 - t)(\lambda_2 - t)...(\lambda_n - t)$ splits over the field $F = \mathbb{R}$. □

Proof of Spectral Theorem 2: Assume $F = \mathbb{R}$ and T is a self-adjoint linear operator on V. Then, the characteristic polynomial of T splits, and Schur's theorem implies that there exists an orthonormal basis β for V such that $A = [T]_\beta$ is upper-triangular. The same proof as for the first spectral theorem now works since T is normal, but it is easier to note that since $T^* = T$, we know that both A and $A^T = A^* = A$ are upper triangular. Therefore, A is diagonal.

Finally, we will show the next theorem, which includes the fact that unitary (and orthogonal) operators must be normal!

THEOREM 6.18 Let V be a finite dimensional inner product space and $T \in \mathcal{L}(V)$. The following are equivalent:

(a) $T^*T = TT^* = I$

(b) $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for every $x, y \in V$

(e) $\|T(x)\| = \|x\|$ for every $x \in V$

Proof:

(a) $\Rightarrow$ (b) Assume (a) $T^*T = TT^* = I$.

Given $x, y \in V$, since $T^*T(x) = I(x) = x$, we have $\langle x, y\rangle = \langle T^*T(x), y\rangle = \langle T(x), T(y)\rangle$.

(b) $\Rightarrow$ (e) Assume (b) $\langle x, y\rangle = \langle T(x), T(y)\rangle$ for every $x, y \in V$

Let $\beta = \{v_1, v_2, ..., v_n\}$ be an orthnormal basis for V. Then $T(\beta) = \{T(v_1), T(v_2), ..., T(v_n)\}$ is an orthonormal basis for V (since, from (b), $\langle T(v_i), T(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$).

Given any $x \in V$, write $x = \sum_{i=1}^n a_i v_i$, we can easily compute $\|x\|^2 = \sum_{i=1}^n |a_i|^2$.

Since T is linear, $T(x) = \sum_{i=1}^n a_i T(v_i)$; the same computation shows $\|T(x)\|^2 = \sum_{i=1}^n |a_i|^2$

(e) $\Rightarrow$ (a) Assume (e) $\|T(x)\|^2 = \|x\|^2$ for every $x \in V$.

For all $x \in V$,

$$\langle x, x\rangle = \|x\|^2 = \|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*T(x)\rangle$$

This implies, for all $x \in V$, that

$$(\star) \qquad \langle x, (I - T^*T)(x)\rangle = 0.$$

Let $U = I - T^*T$. Check that U is self-adjoint: $U^* = I^* - (T^*T)^* = I - T^*T$. Then, by the spectral theorem there exists an orthonormal basis β consisting of eigenvectors of U.

For all $x \in \beta$, $U(x) = \lambda x$. By $(\star)$, $0 = \langle x, U(x)\rangle = \bar{\lambda}\langle x, x\rangle$. Since $x \neq 0$, we must have $\lambda = 0$. This implies that $[U]_\beta$ is the zero matrix! Hence, $U = I - T^*T = T_o$ (the zero transformation on V), so $T^*T = I$. Since V is finite dimensional, this proves that T^* must be the inverse of T and therefore $TT^* = I$. $\square$