ADJOINTS AND SPECTRAL THEOREMS

The main goal of these notes is to prove spectral theorems: theorems that ensure that certain linear maps are always diagonalizable, and in special ways. By “certain linear maps” we mean maps that play nicely with respect to an inner product, and by “special ways” we mean usually that the eigenbasis can be chosen to be orthonormal with respect to that inner product.

Recall definition of an adjoint: if \( V, W \) are inner product spaces over \( \mathbb{R} \) or \( \mathbb{C} \) and \( T : V \to W \) a linear map, then an adjoint \( T^* : W \to V \) is a linear map such that

\[
\langle T(x), y \rangle = \langle x, T^*(y) \rangle.
\]

At the end of last class we proved that adjoints, if they exist, are unique.

**Proposition.** In the situation above, suppose that \( V \) and \( W \) are finite-dimensional. Then every \( T : V \to W \) has an adjoint.

**Proof.** Pick orthonormal bases \( \{v_1, \ldots, v_n\} \) of \( V \) and \( \{w_1, \ldots, w_m\} \) of \( W \). Suppose that \( A \) is the matrix of \( T \) with respect to these bases. Using the construction principle, define \( T^* : W \to V \) by setting

\[
T^*(w_j) = \sum_{i=1}^n A_{ji} v_i.
\]

Then for each \( 1 \leq k \leq m \) and \( 1 \leq l \leq n \) we have

\[
\langle T(v_l), w_k \rangle = \left\langle \sum_{i=1}^m A_{il} w_i, w_k \right\rangle = \sum_{i=1}^m A_{il} \langle w_i, w_k \rangle = \sum_{i=1}^m A_{il} \delta_{ik} = A_{kl}
\]

while

\[
\langle v_l, T^*(w_k) \rangle = \left\langle v_l, \sum_{i=1}^n A_{ki} v_i \right\rangle = \sum_{i=1}^n A_{ki} \langle v_l, v_i \rangle = \sum_{i=1}^n A_{ki} \delta_{li} = A_{kl}.
\]

For arbitrary \( v \in V, w \in W \), express them in terms of the basis vectors and expand out using conjugate linearity. \( \square \)

In particular this shows that if the matrix of \( T \) w.r.t some choice of orthonormal bases is \( A \), then the matrix of \( T^* \) is given by the conjugate transpose. Call this the adjoint matrix; i.e. if \( A \in M_{m \times n} \), then define \( (A^*)_ij = \overline{A_{ji}} \). Of course if the ground field is \( \mathbb{R} \) then the adjoint is just the transpose.

**Definition.** If \( V \) is a finite-dimensional inner product space, a linear map \( T : V \to V \) is self-adjoint if \( T^* = T \) and skew-adjoint if \( T^* = -T \). In the real case this is also called symmetric resp. skew-symmetric and in the complex case Hermitian resp. skew-Hermitian.

In terms of matrices, symmetric means \( A^t = A \), skew-symmetric means \( A^t = -A \), Hermitian means \( A^* = A \), skew-Hermitian means \( A^* = -A \).

(Exercise/sanity check: Give quick \( 2 \times 2 \) examples of each.)

**Proposition.** In the above situation, if \( T \) is self-adjoint then its eigenvalues are all real.

**Proof.** If \( T(x) = \lambda x \) then

\[
\lambda ||x||^2 = \langle \lambda x, x \rangle = \langle T(x), x \rangle = \langle x, T^*(x) \rangle = \langle x, T(x) \rangle = \langle x, \lambda x \rangle = \overline{\lambda}||x||^2
\]

and \( x \neq 0 \) so \( \lambda = \overline{\lambda} \). \( \square \)

(Exercise: Show that if \( T \) is skew-adjoint then its eigenvalues are all imaginary.)

**Proposition.** If \( T \) is self-adjoint, then its eigenvectors with distinct eigenvalues are orthogonal.
Proof. Suppose $T(x) = \lambda x$ and $T(y) = \mu y$, $\lambda \neq \mu$. Then

\[ \langle \lambda x, y \rangle = \langle T(x), y \rangle = \langle x, T^*(y) \rangle = \langle x, T(y) \rangle = \langle x, \mu y \rangle = \mu \langle x, y \rangle \]

where we used previous proposition at the last step. Therefore

\[ (\lambda - \mu) \langle x, y \rangle = 0 \implies \langle x, y \rangle = 0. \]

\[ \square \]

(Exercise: Show that the same is true for $T$ skew-adjoint.)

Yet another kind of matrix:

**Definition.** Let $V$ be finite-dimensional inner product space over $\mathbb{R}$ (resp. $\mathbb{C}$) and $T : V \to V$ linear. Say $T$ is orthogonal (resp. unitary) if for all $x, y \in V$,

\[ \langle x, y \rangle = \langle T(x), T(y) \rangle. \]

Call a square real (resp. complex) matrix orthogonal (resp. unitary) if the corresponding linear map $\mathbb{R}^n \to \mathbb{R}^n$ (resp. $\mathbb{C}^n \to \mathbb{C}^n$) is.

Idea: orthogonal means “preserves distances and angles” i.e. these are like rotations or reflections.

**Proposition.** $T$ is orthogonal (or unitary) if and only if it is an isomorphism and $T^{-1} = T^*$.

**Proof.** Suppose $T$ is orthogonal or unitary. $T(x) = 0$ implies that $\langle x, x \rangle = \langle T(x), T(x) \rangle = \langle 0, 0 \rangle = 0$ so $x = 0$. Thus $\ker T = \{0\}$, and by rank-nullity $T$ is an isomorphism. Also for $x, y \in V$ we have

\[ \langle x, T^{-1}(y) \rangle = \langle T(x), T(T^{-1}(y)) \rangle = \langle T(x), y \rangle = \langle x, T^*y \rangle \]

so $\langle x, T^{-1}(y) - T^*(y) \rangle = 0$. For any given $y$ we can pick $x = T^{-1}(y) - T^*(y)$ so $||T^{-1}(y) - T^*(y)||^2 = 0$ implying $T^{-1}(y) = T^*(y)$. Therefore $T^{-1} = T^*$.

In the other direction, if $T^{-1} = T^*$ then for $x, y \in V$

\[ \langle T(x), T(y) \rangle = \langle x, T^*(T(y)) \rangle = \langle x, T^{-1}(T(y)) \rangle = \langle x, y \rangle. \]

\[ \square \]

**Corollary.** A matrix is orthogonal (resp. unitary) if and only if $A^{-1} = A^t$ (reps. $A^{-1} = A^*$).

(Exercises: Show that a matrix is orthogonal/unitary if and only if its rows form an orthonormal basis if and only if its columns form an orthonormal basis. Show that a change-of-basis between two orthonormal bases is orthogonal/unitary. Show that if $A$ and $B$ are orthogonal/unitary then so are $A^{-1}$ and $AB$.)

**Theorem** (Spectral theorem for symmetric matrices). Let $V$ be a finite-dimensional inner product space over $\mathbb{R}$ and $T : V \to V$ linear. Then $T$ is symmetric if and only if $V$ has an orthonormal basis of eigenvectors of $T$ (over the real numbers!).

**Proof.** Assume $T$ symmetric. Induct on $n = \dim V$. If $n = 1$ result is clear. Assume for $(n-1)$-dimensional spaces. By the Fundamental Theorem of Algebra, the characteristic polynomial of $T$ has a complex root, so $T$ has a complex eigenvalue $\lambda$, which we’ve already shown has to be real. Therefore the $T$ has a real eigenvector $x \in V$ (not just a complex one). Let $W = (\text{Span}(x))^\perp$. By a problem on the current homework, $\dim W = n - 1$.

Claim: $y \in W \implies T(y) \in W$. Proof of claim: $y \in W$ implies $0 = \langle x, y \rangle$, so

\[ 0 = \lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle T(x), y \rangle = \langle x, T^*(y) \rangle = \langle x, T(y) \rangle \]

so $T(y) \in W$. Therefore $T$ restricts to a linear map $T|_W : W \to W$. Certainly $\langle T(y), z \rangle = \langle y, T(z) \rangle$ for all $y, z \in W$, so by uniqueness of the adjoint $(T|_W)^* = T_W$. So $T_W$ is still symmetric. By inductive hypothesis, $W$ has an orthonormal basis of $T$-eigenvectors. Adding $x$ to this basis yields the desired orthonormal eigenbasis of $V$. 
In the other direction, let \( \{v_1, \ldots, v_n\} \) be an orthonormal basis with \( T(v_j) = \lambda_j v_j \) for some \( v_j \in \mathbb{R} \). For \( x, y \in V \) let \( x = \sum_{j=1}^{n} x_j v_j \) and \( y = \sum_{k=1}^{n} y_k v_k \). We compute

\[
\langle T(x), y \rangle = \left\langle T \left( \sum_{j=1}^{n} x_j v_j \right), \sum_{k=1}^{n} y_k v_k \right\rangle \\
= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j y_k \langle T(v_j), v_k \rangle \\
= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j y_k \langle \lambda_j v_j, v_k \rangle \\
= \sum_{j=1}^{n} \sum_{k=1}^{n} x_j y_k \lambda_j \delta_{jk} \\
= \sum_{j=1}^{n} \lambda_j x_j y_j.
\]

A similar computation yields

\[
\langle x, T(y) \rangle = \sum_{j=1}^{n} \lambda_j x_j y_j.
\]

For both we are really using that \( \lambda_j \) is real! \( \square \)

**Corollary (Matrix version).** For \( A \in M_{n \times n}(\mathbb{R}) \), \( A^t = A \) if and only if there exists an orthogonal matrix \( B \) and a real diagonal matrix \( D \) such that \( A = BDB^{-1} \).

**Proof.** Put together previous results and exercise that says that change of basis matrices between orthonormal bases are orthogonal. \( \square \)

There is a corresponding result in the complex case, but it is not as strong (only one direction, not if and only if):

**Theorem (Spectral theorem for Hermitian matrices).** Let \( V \) be a finite-dimensional inner product space over \( \mathbb{C} \) and \( T : V \to V \) Hermitian. Then \( V \) has an orthonormal basis of eigenvectors of \( T \).

The proof is exactly as in (one direction of) the real case. As a corollary, if \( A = A^* \) is a complex matrix, then there exists a unitary matrix \( B \) and a diagonal matrix with real entries (!) such that \( A = BDB^{-1} \).

If we want an if and only if statement we need to include maps with complex eigenvalues too. The general result, a proof of which is outlined on the optional homework assignment, is the following:

**Theorem (Spectral theorem for normal matrices).** Let \( V \) be a finite-dimensional inner product space over \( \mathbb{C} \) and \( T : V \to V \) linear. Then \( T \) is normal, which means that \( T \circ T^* = T^* \circ T \) (\( T \) commutes with its adjoint), if and only if \( V \) has an orthonormal basis of \( T \)-eigenvectors.

(Exercise: All of the special types of linear map we have introduced today are normal.)

Corresponding matrix version: if \( A \) is a complex square matrix, then \( AA^* = A^* A \) if and only if there exists a unitary \( B \) and diagonal \( D \) (which can be complex this time!) such that \( A = BDB^{-1} \).

It is certainly possible for a real matrix/linear map to be normal but not symmetric. In this case there will be an orthonormal basis over \( \mathbb{C} \) but not \( \mathbb{R} \). For example, this will be the case for a rotation matrix in \( \mathbb{R}^2 \):

\[
\begin{bmatrix}
0 & 1 \\
-1 & 0
\end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix}
1 & 1 \\
i & -i
\end{bmatrix} \cdot \begin{bmatrix}
i & 0 \\
0 & -i
\end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix}
1 & -i \\
1 & i
\end{bmatrix}
\]