LECTURE 28: ADJOINTS AND NORMAL OPERATORS

Today's lecture will tie linear operators into our study of Hilbert spaces and discuss an important family of linear operators.

Adjoints

We start with a generalization of the Riesz representation theorem.

Theorem 1 (Riesz Representation Theorem)**.** *Let V be a finite dimensional inner product space. Then the map $r\colon V \to V^*$ defined by*

$$r(\vec{v}) = \langle -, \vec{v} \rangle$$

is a conjugate-linear isomorphism. The inverse will be denoted R

A word about notation: "conjugate linear" means that $r(a\vec{v}) = \bar{a}r(\vec{v})$. This arises because of the sesquilinearity: the second factor is only conjugate linear.

Proof. Our argument for bilinear forms can be repeated *mutatis mutandis*, so we need not check injectivity. Surjectivity is a little easier in this context. If $f \in V^*$, then we can choose a $\vec{u} \in \ker(f)^\perp$. Since $V = \langle \vec{u} \rangle \oplus \ker(f)$, and since this sum is an orthogonal direct sum, we can easily check that

$$R(f) = \frac{\overline{f(\vec{u})}}{||\vec{u}||^2}\vec{u}.$$

□

This will give us a very important construction which provides a beautiful symmetry for maps between inner product spaces.

Theorem 2. *Let $\tau \in \mathcal{L}(V, W)$ where V and W are finite dimensional inner product spaces. Then there is a unique $\tau^* \in L(W, V)$ such that*

$$\langle \tau\vec{v}, \vec{w} \rangle = \langle \vec{v}, \tau^*\vec{w} \rangle.$$

Definition 1. The linear transformation τ^* is the adjoint of τ.

Proof. For each $\vec{w} \in W$, we consider the linear functional on V given by

$$\vec{v} \mapsto \langle \tau\vec{v}, \vec{w} \rangle.$$

This gives us a conjugate linear map $t\colon W \to V^*$. By the Riesz representation theorem, we have a conjugate linear map $V^* \to V$ that associates to each linear functional its Riesz vector. Thus we let $\tau^* = R \circ t$. Since both R and t are conjugate linear, the composite is linear.

We check that τ^* has the required property:

$$\langle \vec{v}, \tau^*\vec{w} \rangle = \langle \vec{v}, R(t(\vec{w})) \rangle = t(\vec{w})(\vec{v}) = \langle \tau\vec{v}, \vec{w} \rangle,$$

where the middle equality is the definition of R.

For uniqueness, we assume that S is another linear transformation that has the same properties as τ^*. Then for all $\vec{v}$ and $\vec{w}$

$$\langle \vec{v}, \tau^*\vec{w} \rangle = \langle \vec{v}, S\vec{w} \rangle,$$

and so $S = \tau^*$. $\square$

How can we make this more concrete? Add in having a basis.

Proposition 1. *If A is an $n \times m$-matrix representing a linear transformation $T\colon V \to W$, then T^* is represented by $\overline{A}^t$.*

Why would this be true? Assume that our bases are orthonormal bases. Then the inner product of $\vec{v}$ and $\vec{w}$ is $\vec{v}^t \cdot \overline{\vec{w}}$. Thus

$$\langle A\vec{v}, \vec{w}\rangle = (A\vec{v})^t \cdot \overline{\vec{w}} = \vec{v}^t \cdot A^t \cdot \overline{\vec{w}} = \vec{v}^t \cdot \overline{(\bar{A}^t\vec{w})} = \langle \vec{v}, \bar{A}^t\vec{w}\rangle.$$

The operation of "conjugate transpose" is clearly compatible with conjugation by an invertible matrix, so this also tells us the general case.

Passage to adjoints is a very nice operation. The map that sends τ to τ^* is conjugate linear, and moreover, the conjugate symmetry of the inner products shows that

$$\tau^{**} = \tau$$

for any linear operator.

There is a very close connection between adjunctions, orthogonality, and kernels and images.

Proposition 2. *Let $T \in \mathcal{L}(V, W)$ where everything in sight is a finite dimensional inner product space.*

1. $\ker(T^*) = Im(T)^\perp$
2. $\ker(T^*T) = \ker(T)$

Proof. Let $\vec{u} \in \ker(T^*)$. This is the same statement as $T^*\vec{u} = 0$, and so in particular for all $\vec{v} \in V$, we see

$$0 = \langle T^*\vec{u}, \vec{v}\rangle = \langle \vec{u}, T\vec{v}\rangle.$$

In particular, we see that $\vec{u}$ is orthogonal to $T\vec{v}$ for all $\vec{v} \in V$, and therefore $\vec{u} \in Im(T)^\perp$.

For the second part, it is obvious that $\ker(T) \subset \ker(T^*T)$. Now let $\vec{u} \in \ker(T^*T)$. Then

$$0 = \langle T^*T\vec{u}, \vec{u}\rangle = \langle T\vec{u}, T\vec{u}\rangle,$$

and so $T\vec{u} = 0$. $\square$

Combining these, we also see

$$Im(T^*T) = \ker(T^*T)^\perp = \ker(T)^\perp = Im(T^*).$$

We learn a number of things on top of this. We see that T is surjective iff T^* is injective, and T is injective iff T^* is surjective. This should remind you of the dual of short exact sequences being short exact. Here we've replaced the dual of a linear operator with the adjoint.

Now we can focus on a few specific kinds of special linear transformations.

Definition 2. A linear operator $T\colon V \to V$ is

1. <u>Normal</u> if $T^*T = TT^*$
2. <u>self-adjoint</u> if $T^* = T$ (<u>Hermitian</u> if $\mathbb{F} = \mathbb{C}$ and <u>symmetric</u> if $\mathbb{F} = \mathbb{R}$)
3. <u>skew-self-adjoint</u> if $T^* = -T$
4. <u>unitary</u> if $T^* = T^{-1}$

Proposition 3. *If T is a normal operator and $p(x)$ is any polynomial, then $p(T)$ is a normal operator. In particular, $T - \lambda I$ is normal.*

Normal Operators

We finish the class tying together this theory of normal operators with Jordan form.

Theorem 3. *If T is a normal operator, then T is diagonalizable. Moreover, the eigenspaces are orthogonal.*

We'll quickly see that T is diagonalizable. Let $\lambda_1, \dots, \lambda_k$ be the eigenvalues of T. We know that we can put T into Jordan form. The algorithm starts with $B = T - \lambda_i I$, and then it considers the ascending chain

$$\ker(B) \subset \ker(B^2) \subset \cdots \subset \ker(B^r) = \dots.$$

The smallest value of r such that $\ker(B^r) = \ker(B^{r+1})$ is the size of the largest Jordan block. Thus the diagonalizablity is equivalent to $r = 1$, since then all of the Jordan blocks are 1×1 blocks. We will have proved the theorem if we can see $\ker(B) = \ker(B^r)$ for all $r \geq 1$.

Lemma 1. *If T is normal, then $\ker(T) = \ker(T^r)$ for all $r \geq 1$.*

Proof. Let $S = T^*T$. This is a self-adjoint linear operator. We can see that $\ker(S) = \ker(S^r)$ for all $r \geq 1$. Assume $\vec{u} \in \ker(S^r)$. Then we have

$$0 = \langle S^r(\vec{u}), S^{r-2}(\vec{u})\rangle = \langle S^{r-1}(\vec{u}), S^{r-1}(\vec{u})\rangle,$$

and we therefore know that $S^{r-1}(\vec{u}) = 0$. By downward induction on r, we see that $\vec{u} \in \ker(S)$.

Now let $\vec{v} \in \ker(T^r)$. Then since $S^r = (T^*)^r T^r$,

$$\vec{v} \in \ker(T^r) \subset \ker(S^r) = \ker(S) = \ker(T),$$

and so $\ker(T) = \ker(T^r)$ for all $r \geq 1$. □

Now we can check that the eigenspaces are orthogonal.

Lemma 2. *If T is normal, then*

$$\langle T\vec{v}, T\vec{w}\rangle = \langle T^*\vec{v}, T^*\vec{w}\rangle,$$

so $\ker(T) = \ker(T^*)$.

Proof. We use normality to see this:

$$\langle T\vec{v}, T\vec{w}\rangle = \langle T^*T\vec{v}, \vec{w}\rangle = \langle TT^*\vec{v}, \vec{w}\rangle = \langle T^*\vec{v}, T^*\vec{w}\rangle.$$

The second part follows from the first by considering

$$\langle T\vec{v}, T\vec{v}\rangle = \langle T^*\vec{v}, T^*\vec{v}\rangle.$$

□

In particular, this shows us that the eigenvalues of T^* are the conjugates of those of T, and the eigenvectors are the same.

Proposition 4. *If $\vec{v}$ and $\vec{u}$ are eigenvectors for T for distinct eigenvalues, then $\vec{v} \perp \vec{u}$.*

Proof. Let λ be the eigenvalue for $\vec{v}$ and let μ be the eigenvalue for $\vec{u}$. Then

$$\lambda\langle \vec{v}, \vec{u}\rangle = \langle T\vec{v}, \vec{u}\rangle = \langle \vec{v}, T^*\vec{u}\rangle = \langle \vec{v}, \bar{\mu}\vec{u}\rangle = \mu\langle \vec{v}, \vec{u}\rangle.$$

Since $\lambda \neq \mu$, we only have such an equality if $\langle \vec{v}, \vec{u}\rangle = 0$, and so $\vec{v} \perp \vec{u}$. □