28 The Spectral Theorem

28.1 Review: Hermitian Spaces

Last time, we discussed Hermitian spaces which were just complex vector spaces with a positive definite Hermitian form. Given a linear operator T, we defined the adjoint T^*, which had the property that

$$\langle v, T^*w \rangle = \langle Tv, w \rangle.$$

We called a linear operator T normal if $TT^* = T^*T$. We then were able to state the Spectral Theorem.

28.2 The Spectral Theorem

The Spectral Theorem demonstrates the special properties of normal and real symmetric matrices.

Theorem 28.1

Given a Hermitian space V, for any normal linear operator T, there exists an orthonormal eigenbasis of V: $\{u_1, \cdots, u_n\}$. In matrix form, for any normal matrix $M \in GL_n(\mathbb{C})$, there exists a unitary matrix P such that $P^{-1}MP$ is diagonal.

The real version states that for a Euclidean vector space V and a *symmetric* linear operator T, there exists an orthonormal eigenbasis; equivalently, for any symmetric matrix $M \in GL_n(\mathbb{R})$, there exists an orthogonal matrix P such that $P^{-1}MP$ is diagonal. All eigenvalues of real symmetric matrices are real.

Example 28.2

Say we take the symmetric matrix $M = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$. One eigenvector is $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_1 = 2$. The other eigenvector should be orthogonal, so it is $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with eigenvalue $\lambda_2 = 4$.

In 2 dimensions, the fact that our change of basis is coming from an orthogonal basis tells us that the change in coordinates is just a rotation. This rotation makes the operator diagonalizable.

This is called the Spectral Theorem because the eigenvalues are often referred to as the spectrum of a matrix. Any theorem that talks about diagonalizing operators is often called a spectral theorem.

Now we will state some lemmas in order to prove the Spectral Theorem.

Lemma 28.3 (Lemma 1)

First, for a linear operator $T : V \longrightarrow V$ where V is Hermitian, and a subspace $W \subset V$ such that $T(W) \subset W$, then $T^*(W^\perp) \subset W^\perp$.

Proof. Take some $u \in W^\perp$. For any $w \in W$, then

$$\langle w, T^*u \rangle = \langle Tw, u \rangle = 0,$$

by the definition of the adjoint linear operator, so $T^*u \in W^\perp$. $\qquad\square$

Lemma 28.4 (Lemma 2)

If $Tv = \lambda v$, then $T^*v = \bar{\lambda}v$. This just means that T and T^* have the same eigenvectors, and eigenvalues related by complex conjugation.

Proof. We can first consider a special case.

- **Special Case:** $\lambda = 0$. We want to show that $Tv = 0$ implies that $T^*v = 0$.

$$\langle T^*v, T^*v \rangle = \langle Tv, Tv \rangle = 0$$

by normality (this was a property shown last time). Since T is positive definite, this implies that $T^*v = 0$. This also implies that if a vector $v \in \ker T$, then $v \in \ker T^*$ as well.

- **General Case:** $\lambda \in \mathbb{C}$. We want to show that $Tv = \lambda v$ implies that $T^*v = \overline{\lambda} v$. In this case, create a new linear operator $S = T - \lambda I$. Then, $Sv = 0$. Also, the adjoint is $S^* = T^* - \overline{\lambda} I$ and $SS^* = S^*S$, since $TT^* = T^*T$. By the first case, since $v \in \ker S$, we also have that $v \in \ker S^*$ so $T^* = \overline{\lambda} v$ as needed.

$\square$

Proof of the Spectral Theorem. Now we can prove the Spectral Theorem by induction on the dimension of V. The general process will be breaking V up into a direct sum and finding orthonormal eigenbases over each part. Since the field is $F = \mathbb{C}$, it is always possible to find an eigenvector $w \in V$ such that $Tw = \lambda w$. By normalizing w, assume that $\langle w, w \rangle = 1$. Then, let $W = \mathrm{Span}(w)$; then $V = W \oplus W^\perp$ since all subspaces of Hermitian spaces are non-degenerate. In addition, since w is an eigenvector, $TW \subset W$.

If we can show that $TW^\perp$ preserves $W^\perp$, then we can induct. By Lemma 2, w is also an eigenvector of T^*, so $T^*(W) \subset W$, and by Lemma 1, $(T^*)^*(W^\perp) = T(W^\perp) \subset W^\perp$. Then $V = W \oplus W^\perp$, and T maps each part of the splitting to itself, so by induction there exists $u_2, \cdots, u_n$ an orthonormal eigenvasis for $W^\perp$, and $\{w, u_2, \cdots, u_n\}$ is an orthonormal eigenbasis for V. $\square$

Student Question. *Why is the equivalent of the theorem not true over the real numbers?*

Answer. *Most of this argument works, except in the very first step, where we found an eigenvector and eigenvalue. We cannot guarantee this will happen with normal linear operators over the real numbers. However, as we found last week, for symmetric (and Hermitian) matrices, the eigenvalues are all real, and in particular it is always possible to find one eigenvector $w \in V$ with real eigenvalue to allow the induction to occur.*

As an application, suppose we have a quadratic function $f(x, y) = ax^2 + bxy + cy^2$. Let this function be represented by a matrix $\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$; then

$$f(x, y) = (x, y) M \begin{pmatrix} x \\ y \end{pmatrix} = \langle v, Tv \rangle$$

for a linear operator T given by M. By the Spectral Theorem, there exists an orthogonal change of coordinates $P^T M P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, where P is an orthogonal matrix. It takes $\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} x' \\ y' \end{pmatrix}$. Then

$$f(x, y) = (x, y) M \begin{pmatrix} x \\ y \end{pmatrix} = (x', y') \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \lambda_1 (x')^2 + \lambda_2 (y')^2.$$

Example 28.5

If $f(x, y) = 3x^2 - 2xy + 3y^2$, the matrix would be $M = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$. Then

$$3x^2 - 2xy + 3y^2 = 2(x')^2 + 4(y')^2,$$

for some change of coordinates taking x and y to x' and y', since we saw earlier that the eigenvalues were 2 and 4.

In n variables, let $f(x_1, \cdots, x_n) = a_{11} x_1^2 + \cdots + a_{nn} x_n^2 + \sum_{i<j} 2 a_{ij} x_i x_j$ be represented by the matrix $M = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{ij} \\ \vdots & \ddots & \\ a_{ji} & & \ddots \\ & & & a_{nn} \end{pmatrix}$.

Applying the spectral theorem gives P and λ_i, where P is a change of coordinates taking $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$.

In this new coordinate system,

$$f(x_1, \cdots, x_n) = \lambda_1 (x'_1)^2 + \cdots + \lambda_n (x'_n)^2.$$

In two and three dimensions, this has a geometric interpretation. Consider a curve in $\mathbb{R}^2$ with an equation of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

There are several options; the first three are called conics, and the rest are degenerate conics of some form.

- **Ellipse.** $ax^2 + by^2 = 1$
- **Hyperbola.** $ax^2 - by^2 = 1$
- **Parabola.** $ax^2 - y = 0$
- **Two crossing lines.** $(a_1x + b_1y)(a_2x + b_2y) = 0$
- **Two parallel lines.** $x^2 = a, a > 0$
- **One line.** $x^2 = 0$
- **One point.** $x^2 + y^2 = 0$
- **The empty set $\varnothing$.** $x^2 + y^2 = -1$

> **Theorem 28.6**
> After an isometry, all curves of the form $ax^2 + bxy + cy^2 + dx + ey + f = 0$ look like one of these options.

Proof. The curve be rewritten as $\vec{v}^T A \vec{v} + B \vec{v} + f = 0$. After an orthogonal change of basis, we can assume that $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2. \end{pmatrix}$ Then $\lambda_1 x^2 + \lambda_2 y^2 + b_1 x + b_2 y + f = 0$.

- **Nonzero eigenvalues.** If $\lambda_1, \lambda_2 \neq 0$, then we can complete the square, taking $x \rightsquigarrow x + \frac{b_1}{2\lambda_1}$ and $y \rightsquigarrow y + \frac{b_2}{2\lambda_2}$. Then

$$\lambda_1 x^2 + \lambda_2 y^2 = c.$$

  If $c = 0$, it is a single point or two intersecting lines, and if $c \neq 0$, it is an ellipse or a hyperbola.

- **Zero eigenvalues.** The other cases.

The details of this aren't that important, and we glossed over much of it at the end. The main idea to take away is just the power of the Spectral Theorem to take a complicated quadratic form and simplify into something more readily analyzed. $\qquad\square$

One idea where the Spectral Theorem is at play is in multivariable calculus with the second derivative test. Hidden behind the test is that we have a symmetric matrix and we are trying to figure out what the signs of the eigenvalues are, as that will tell us whether our critical point is a minimum, maximum, or saddle point.