Chapter 7

The Spectral Theorem

We come to one of the most important topics of the course. In simple terms, any real symmetric matrix is diagonalisable. But there is more to be said!

7.1 Orthogonal projections and orthogonal decompositions

Definition 7.1. We say that two vectors u, w in an inner product space V are *orthogonal* if $u \cdot w = 0$. We say that two subspaces U and W of V are orthogonal if $u \cdot w = 0$ for all $u \in U$ and $w \in W$.

Definition 7.2. Let V be a real inner product space, and U a subspace of V. The *orthogonal complement* of U is the set of all vectors that are orthogonal to everything in U:

$$U^{\perp} = \{w \in V : w \cdot u = 0 \text{ for all } u \in U\}.$$

Thus, the orthogonal complement of U is the largest subspace of V that is orthogonal to U.

Proposition 7.3. *If V is a real inner product space and U a subspace of V, with $\dim(V) = n$ and $\dim(U) = r$, then $U^{\perp}$ is a subspace of V, and $\dim(U^{\perp}) = n - r$. Moreover, $V = U \oplus U^{\perp}$.*

Proof. Proving that $U^{\perp}$ is a subspace is straightforward from the properties of the inner product. If $w_1, w_2 \in U^{\perp}$, then $w_1 \cdot u = w_2 \cdot u = 0$ for all $u \in U$, so $(w_1 + w_2) \cdot u = 0$ for all $u \in U$, whence $w_1 + w_2 \in U^{\perp}$. The argument for scalar multiples is similar.

Now choose a basis $(u_1, u_2, \ldots, u_r)$ for U and extend it to a basis $(u_1, u_2, \ldots, u_n)$ for V. Then apply the Gram-Schmidt process to this basis (processing the vectors in the order $u_1, u_2, \ldots, u_n$), to obtain an orthonormal basis $(v_1, \ldots, v_n)$ of V. Since the process only modifies vectors by adding multiples of earlier vectors, the first r vectors in the resulting basis will form an orthonormal basis for U. The last $n - r$ vectors will be orthogonal to U, and so lie in $U^{\perp}$. Summarising, we have $v_1, \ldots, v_r \in U$ and $v_{r+1}, \ldots, v_n \in U^{\perp}$. Since $v_1, \ldots, v_n$ is a basis for V, it follows that every vector in V can be written as the sum of a vector in U and a vector in $U^{\perp}$ or, equivalently, $V = U + U^{\perp}$.

To show that V is actually a *direct* sum of U and $U^{\perp}$ we just need to show that $U \cap U^{\perp} = \{\mathbf{0}\}$. But if $u \in U$ and $u \in U^{\perp}$ then $u \cdot u = 0$ which implies $u = \mathbf{0}$.

The claim about the dimension of subspaces follows from Lemma 1.28. $\square$

Recall the connection between direct sum decompositions and projections. If we have projections $\pi_1, \ldots, \pi_r$ whose sum is the identity and which satisfy $\pi_i \pi_j = 0$ for $i \neq j$, then the space V is the direct sum of their images. This can be refined in an inner product space as follows.

Definition 7.4. Let V be an inner product space, A linear map $\pi : V \to V$ is an *orthogonal projection* if

(a) π is a projection, that is, $\pi^2 = \pi$, and

(b) π is self-adjoint, that is, $\pi^* = \pi$.

Definition 7.5. Suppose V is an inner product space, and $U_1, \ldots, U_r$ are subspaces of V. A direct sum $V = U_1 \oplus \cdots \oplus U_r$ is an *orthogonal decomposition of* V if U_i is orthogonal to U_j for all $i \neq j$.

Proposition 7.6. *Suppose $\pi_1, \pi_2, \ldots, \pi_r$ are orthogonal projections on an inner product space V, satisfying*

(a) $\pi_1 + \pi_2 + \cdots + \pi_r = I$, *where I is the identity map, and*

(b) $\pi_i \pi_j = 0$, *for $i \neq j$.*

Let $U_i = \mathrm{Im}(\pi_i)$, for $i = 1, \ldots, r$. Then $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$ is an orthogonal decomposition of V.

Proof. The fact that V is the direct sum of the images of the π_i follows from Proposition 5.4. We only have to prove that U_i and U_j are orthogonal for all $i \neq j$. Recall that if π is a projection, then $v \in \mathrm{Im}(\pi)$ if and only if $\pi(v) = v$. So take $u_i \in U_i$ and $u_j \in U_j$ with $i \neq j$. Then $\pi_i(u_i) = u_i$ and $\pi_j(u_j) = u_j$ and hence

$$u_i \cdot u_j = \pi_i(u_i) \cdot \pi_j(u_j) = u_i \cdot \pi_i^*(\pi_j(u_j)) = u_i \cdot \pi_i(\pi_j(u_j)) = 0,$$

where the second equality is the definition of the adjoint, and the third holds because π_i is self-adjoint. □

As with Proposition 5.4, there is a converse.

Proposition 7.7. *Suppose $V = U_1 \oplus \cdots \oplus U_r$ is an orthogonal decomposition of an inner product space V. Then there exist orthogonal projections $\pi_1, \pi_2, \ldots, \pi_r$ on V satisfying*

(a) $\pi_1 + \pi_2 + \cdots + \pi_r = I$,

(b) $\pi_i \pi_j = 0$, *for $i \neq j$, and*

(c) $\mathrm{Im}(\pi_i) = U_i$, *for all i.*

Proof. From Proposition 5.5 we know that there are projections π_i, for $1 \leq i \leq r$, satisfying conditions (a)–(c). Only one extra thing needs to be checked, namely that these projections are orthogonal, i.e., that the π_i are self adjoint. To see this, suppose v, w are arbitrary vectors in V. Then

$$\begin{aligned}
\pi_i(v) \cdot w &= \pi_i(v) \cdot (\pi_1(w) + \cdots + \pi_r(w)) \\
&= \pi_i(v) \cdot \pi_i(w) \\
&= (\pi_1(v) + \cdots + \pi_r(v)) \cdot \pi_i(w) \\
&= v \cdot \pi_i(w),
\end{aligned}$$

where the first and fourth equalities follow from $\pi_1 + \cdots + \pi_r = I$ and the second and third from the fact that $\mathrm{Im}(\pi_i) = U_i$ is orthogonal to $\mathrm{Im}(\pi_j) = U_j$ when $i \neq j$. It follows that π_i is self-adjoint. $\square$

7.2 The Spectral Theorem

The main theorem can be stated in different ways. We list three alternatives here.

Theorem 7.8. *If α is a self-adjoint linear map on a real inner product space V, then there is an orthonormal basis of V consisting of eigenvectors of α. Thus, the eigenspaces of α form an orthogonal decomposition of V.*

Equivalently, we can state the result as follows.

Corollary 7.9. *Suppose α and V are as in the previous theorem, and $\lambda_1, \ldots, \lambda_r$ are the distinct eigenvalues of α. Then there exist orthogonal projections $\pi_1, \ldots, \pi_r$ satisfying*

(a) $\pi_1 + \cdots + \pi_r = I$,

(b) $\pi_i\pi_j = 0$, *whenever* $i \neq j$, *and*

(c) $\alpha = \lambda_1\pi_1 + \cdots + \lambda_r\pi_r$.

Proof of Corollary 7.9. By Theorem 7.8 we known that V has an orthogonal decomposition $V = E(\lambda_1, \alpha) \oplus \cdots \oplus E(\lambda_r, \alpha)$, where $E(\lambda_i, \alpha)$ is the eigenspace correponding to the eigenvector λ_i. Then, by Proposition 7.7, there exist orthogonal projections, satisfying (a) and (b), such that $\mathrm{Im}(\pi_i) = E(\lambda_i, \alpha)$ for $1 \leq i \leq r$. Condition (c) then follows from the following chain of equalities:

$$\alpha(v) = \alpha(\pi_1(v) + \cdots + \pi_r(v)) = \lambda_1\pi_1(v) + \cdots + \lambda_r\pi_r(v) = (\lambda_1\pi_1 + \cdots + \lambda_r\pi_r)(v).$$

$\square$

Yet another statement of the spectral theorem is in terms of matrices. Since a symmetric matrix represents a self-adjoint linear map with respect to some orthonormal basis, e.g., the standard basis of $\mathbb{R}^n$:

Corollary 7.10. *Let A be a real symmetric matrix. Then there exists an orthogonal matrix P such that $P^{-1}AP$ is diagonal. In other words, any real symmetric matrix is orthogonally similar to a diagonal matrix.*

Proof of Theorem 7.8. The proof will be by induction on $n = \dim(V)$. There is nothing to do if $n = 1$. So we assume that the theorem holds for $(n-1)$-dimensional spaces. The first job is to show that α has an eigenvector.

Choose an orthonormal basis; then α is represented by a real symmetric matrix A. Its characteristic polynomial has a root λ over the complex numbers. (The so-called "Fundamental Theorem of Algebra" asserts that any polynomial over $\mathbb{C}$ has a root.) We temporarily enlarge the field from $\mathbb{R}$ to $\mathbb{C}$. Now we can find a column vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Taking the complex conjugate, remembering that A is real, we have $A\overline{v} = \overline{\lambda}\overline{v}$. Then we have

$$\overline{\lambda}\,\overline{v}^\top v = (A\overline{v})^\top v = (\overline{v}^\top A)v = \overline{v}^\top(Av) = \overline{v}^\top(\lambda v) = \lambda\overline{v}^\top v.$$

Since v is not the zero vector, $\overline{\lambda} = \lambda$, that is, λ is real. Now since α has a real eigenvalue, we can choose a real eigenvector v, and (multiplying by a scalar if necessary) we can assume that $|v| = 1$.

Let U be the subspace $U = \{u \in V : v \cdot u = 0\}$. This is a subspace of V of dimension $n-1$, by Proposition 7.3. We claim that $\alpha : U \to U$. For take $u \in U$. Then

$$\alpha(u) \cdot v = u \cdot \alpha^*(v) = u \cdot \alpha(v) = \lambda(u \cdot v) = 0,$$

where we use the fact that α is self-adjoint. So $\alpha(u) \in U$.

So α restricted to U is a self-adjoint linear map on the $(n-1)$-dimensional inner product space. By the inductive hypothesis, U has an orthonormal basis consisting of eigenvectors of α. They are all orthogonal to the unit vector v; so, adding v to the basis, we get an orthonormal basis for V, as required.

The fact that V is a direct sum of eigenspaces comes from Theorem 5.14, so for the final part of the theorem we just need to show that these eigenspaces are orthogonal. We could use the orthonormal basis just constructed to prove this but it is easier to go directly. Suppose $v \in E(\lambda, \alpha)$ and $w \in E(\mu, \alpha)$ are vectors in distinct eigenspaces. Then $\alpha(v) = \lambda v$ and $\alpha(w) = \mu w$, and

$$\lambda(v \cdot w) = \lambda v \cdot w = \alpha(v) \cdot w = v \cdot \alpha^*(w) = v \cdot \alpha(w) = v \cdot \mu w = \mu(v \cdot w),$$

so, since $\lambda \neq \mu$, we see that $v \cdot w = 0$. □

Remark 7.11. The theorem is almost a canonical form for real symmetric relations under the relation of orthogonal congruence. If we require that the eigenvalues occur in decreasing order down the diagonal, then the result is a true canonical form: each matrix is orthogonally similar to a unique diagonal matrix with this property.

Example 7.12. Let

$$A = \begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\begin{vmatrix} x-10 & -2 & -2 \\ -2 & x-13 & -4 \\ -2 & -4 & x-13 \end{vmatrix} = (x-9)^2(x-18),$$

so the eigenvalues are 9 and 18.

For eigenvalue 18 the eigenvectors satisfy

$$\begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 18x \\ 18y \\ 18z \end{bmatrix},$$

so the eigenvectors are multiples of $\begin{bmatrix} 1 & 2 & 2 \end{bmatrix}^\top$. Normalising, we can choose a unit eigenvector $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}^\top$.

For the eigenvalue 9, the eigenvectors satisfy

$$\begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9x \\ 9y \\ 9z \end{bmatrix},$$

that is, $x+2y+2z=0$. (This condition says precisely that the eigenvectors are orthogonal to the eigenvector for $\lambda=18$, as we know.) Thus the eigenspace is 2-dimensional. We need to choose an orthonormal basis for it. This can be done in many different ways: for example, we could choose $\begin{bmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}^\top$ and $\begin{bmatrix}-4/3\sqrt{2} & 1/3\sqrt{2} & 1/3\sqrt{2}\end{bmatrix}^\top$. Then we have an orthonormal basis of eigenvectors. We conclude that, if

$$P=\begin{bmatrix}1/3 & 0 & -4/3\sqrt{2}\\ 2/3 & 1/\sqrt{2} & 1/3\sqrt{2}\\ 2/3 & -1/\sqrt{2} & 1/3\sqrt{2}\end{bmatrix},$$

then P is orthogonal, and

$$P^\top AP=\begin{bmatrix}18 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9\end{bmatrix}.$$

You might like to check that the orthogonal matrix in the example in the last chapter of the notes also diagonalises A.