LECTURE 19

The Spectral Theorem

19.1 Self-adjoint maps

There are many interesting results one can establish on adjoints; however, in this lecture we will restrict our attention to the case of endomorphisms $f \in \mathcal{L}(V)$. In particular, we wish to study the following concept.

Definition 19.1 — Self-adjoint map. An endomorphism f defined on a finite-dimensional inner product space V is called self-adjoint if $f = f^*$ with respect to the inner product on V.

Note that, by definition, there holds

$$\langle f(\mathbf{v}), \mathbf{w} \rangle_V = \langle \mathbf{v}, f(\mathbf{w}) \rangle_V .$$

This expression yields the following characterisation of a self-adjoint map.

Theorem 19.1 Let $f \in \mathcal{L}(V)$ where V is an inner product equipped with an orthonormal basis B. Then f is self-adjoint if and only if its matrix representation relative to the basis B is a symmetric matrix.

Proof. $\boxed{\Longrightarrow}$ Let f be self-adjoint and let A denote its matrix representation relative to the orthonormal basis B. We write the self-adjoint property as

$$\langle f(\mathbf{v}), \mathbf{w} \rangle_V = \langle \mathbf{v}, \mathbf{f}(\mathbf{w}) \rangle_V = \langle f(\mathbf{w}), \mathbf{v} \rangle_V .$$

Let $\varphi_V(\mathbf{v}) = \mathbf{x}, \varphi_V(\mathbf{w}) = \mathbf{y}$. By Proposition 18.2,

$$\langle f(\mathbf{v}), \mathbf{w} \rangle_V = \mathbf{y}^T A\mathbf{x}, \quad \langle f(\mathbf{w}), \mathbf{v} \rangle_V = \mathbf{x}^T A\mathbf{y}.$$

On the other hand, since $\mathbf{x}^T A\mathbf{y}$ is a scalar and the transpose of a scalar is the scalar, the above relation can be written as

$$\mathbf{y}^T A\mathbf{x} = \mathbf{x}^T A\mathbf{y} = (\mathbf{x}^T A\mathbf{y})^T = \mathbf{y}^T A^T\mathbf{x} \Longrightarrow A = A^T,$$

and the matrix is symmetric.
$\boxed{\Longleftarrow}$ Let the matrix representation A of f relative to B be symmetric. By Proposition 18.4, the matrix representation of f^* is $A^T = A$. Since the matrix representations are equal, $f = f^*$. ■

Example 19.1 Let $V = \mathcal{P}_2([-1,1])$ and let $f \in \mathcal{L}(V)$ be the map $f(p) = -\left((1-x^2)p'\right)'$. Let $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ be defined via

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

Let us show f is self-adjoint. We have

$$\langle f(p), q\rangle = -\int_{-1}^{1} \left((1-x^2)p'(x)\right)' q(x)\mathrm{d}x = \overbrace{-\left[(1-x^2)p'(x)q(x)\right]_{-1}^{1}}^{=0} + \int_{-1}^{1} (1-x^2)p'(x)q'(x)\mathrm{d}x,$$

while

$$\langle p, f(q)\rangle = -\int_{-1}^{1} p(x)\left((1-x^2)q'(x)\right)' \mathrm{d}x = \overbrace{-\left[p(x)(1-x^2)q'(x)\right]_{-1}^{1}}^{=0} + \int_{-1}^{1} (1-x^2)q'(x)p'(x)\mathrm{d}x,$$

so that $\langle f(p), q\rangle = \langle p, f(q)\rangle$ and therefore $f = f^*$. Hence $f : V \to V$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle$.

19.2 Spectral properties

Let us turn our attention to the eigenvalues of a self-adjoint map. We will continue to discuss the case of real inner product spaces. This is of particular interest, as we learned in the previous lecture that diagonalising maps over $\mathbb{R}$ is not possible. However, this changes if we restrict the set of maps to self-adjoint maps. We start with the following preliminary result.

Lemma 19.2 Let V be a real inner product space and let $f \in \mathcal{L}(V)$ be self-adjoint. Then the following spectral properties hold:

i. the eigenvalues of f are real;
ii. the eigenvectors of f can be chosen to be real.

Proof. i. Let V be a real inner product space and let $f \in \mathcal{L}(V)$ be self-adjoint. Let $A \in \mathbb{R}^{n\times n}$ be the matrix representation of f in some orthonormal basis of V. Then A is symmetric. Let us consider the eigenvalue problem for A. Since in general we are not guaranteed to have eigensolutions in $\mathbb{R}$, let us consider the problem over $\mathbb{C}$: find $\lambda \in \mathbb{C}, \mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

First note that λ can be given the following representation:

$$\overline{\mathbf{x}}^T A\mathbf{x} = \lambda\overline{\mathbf{x}}^T\mathbf{x} \iff \lambda = \frac{\overline{\mathbf{x}}^T A\mathbf{x}}{\overline{\mathbf{x}}^T\mathbf{x}} =: \frac{z_1}{z_2},$$

where $\overline{\mathbf{x}}$ denotes taking the complex-conjugate entrywise. Now,

$$z_1 := \overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T(A\mathbf{x}) = \langle\overline{\mathbf{x}}, A\mathbf{x}\rangle = (A\mathbf{x})^T\overline{\mathbf{x}} = \mathbf{x}^T A^T\overline{\mathbf{x}} = \mathbf{x}^T A\overline{\mathbf{x}},$$

where we used the symmetry of A. Taking the complex-conjugate of the expression for z_1 we find

$$\overline{z_1} = \overline{(\mathbf{x}^T A\overline{\mathbf{x}})} = \overline{\mathbf{x}}^T A\mathbf{x} = z_1,$$

so that $z_1 \in \mathbb{R}$. Moreover,

$$z_2 = \overline{\mathbf{x}}^T\mathbf{x} = \sum_{i=1}^{n} \overline{x_i}x_i = \sum_{i=1}^{n} |x_i|^2 \in \mathbb{R}.$$

Hence $\lambda = z_1/z_2$ is real as the ratio of two real scalars.

ii. Let $\lambda \in \mathbb{R}$, and let $\mathbf{x} \in \mathbb{C}^n$ be the corresponding eigenvector. Let $\mathbf{x} = \mathbf{y} + i\mathbf{z}$ for some vectors $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Then both $\mathbf{y}$ and $\mathbf{z}$ must be eigenvectors associated with λ, since

$$A\mathbf{x} = \lambda\mathbf{x} \iff A(\mathbf{y} + i\mathbf{z}) = \lambda(\mathbf{y} + i\mathbf{z}) \iff \begin{cases} A\mathbf{y} = & \lambda\mathbf{y}, \\ A\mathbf{z} = & \lambda\mathbf{z}, \end{cases}$$

by comparing real and imaginary parts. Since eigenvectors are not unique, we can choose a convenient form, i.e., a real eigenvector for the real eigenvalue λ.

■

R Note that if $\mathbf{y}$ and $\mathbf{z}$ are real eigenvectors, then we must have $\mathbf{z} = c\mathbf{y}$ for some (real) scalar c. Then the complex eigenvector $\mathbf{x}$ satisfies $\mathbf{x} = (1 + ic)\mathbf{z}$, which is what we would expect: eigenvectors are unique up to multiplication by a scalar.

This result does not necessarily imply that the matrix is diagonalisable over $\mathbb{R}$; in general, this limited spectral information is insufficient to guarantee that algebraic and geometric multiplicities are equal for each eigenvalue, i.e., that the matrix is not defective. We consider this next.

The following theorem is one of the major results of Linear Algebra. The statement can be equivalently given for symmetric matrices and self-adjoint maps. We include the latter version below.

Theorem 19.3 — Real Spectral Theorem. Let V be a real inner product space of dimension n and let $f \in \mathcal{L}(V)$ be self-adjoint. Then f is diagonalisable, with real eigenvalues and real eigenvectors which form an orthonormal basis for V.

Proof. By Lemma 19.2, the eigenvalues are real, so we only have to show that the matrix is diagonalisable, with orthonormal eigenvectors. We show this by induction on the dimension of V.

$P(n = 1)$: Let $n = 1$. Then the statement holds trivially.
$P(n = k-1)$: Assume that the statement holds for $n = k-1$: all self-adjoint maps on $(k-1)$-dimensional real inner product spaces are diagonalisable with real eigenvalues and orthonormal corresponding eigenvectors.
$P(n = k-1) \Rightarrow P(n = k)$: Let $n = k$ and consider a real eigenpair $(\mu, \mathbf{u})$ of f, where we assume that $\mathbf{u}$ is a unit vector. Let $U = \operatorname{span}\{\mathbf{u}\}$ and consider its orthogonal complement in V: $U^\perp = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u}\rangle = 0\}$. By Proposition 6.9, $V = U \oplus U^\perp$, so that dim $U^\perp = k-1$. Now, $U^\perp$ is an f-invariant subspace of V since for all $\mathbf{v} \in U^\perp$, there holds $f(\mathbf{v}) \in U^\perp$ as

$$\langle f(\mathbf{v}), \mathbf{u}\rangle = \langle \mathbf{v}, f(\mathbf{u})\rangle = \langle \mathbf{v}, \lambda\mathbf{u}\rangle = \mu \langle \mathbf{v}, \mathbf{u}\rangle = 0.$$

By Proposition 15.1, the restriction $\tilde{f}$ of f to $U^\perp$ is an endomorphism on $U^\perp$, i.e., $\tilde{f} \in \mathcal{L}(U^\perp)$. Moreover, $\tilde{f}$ is self-adjoint on $U^\perp$, since for all $\mathbf{v}_1, \mathbf{v}_2 \in U^\perp$

$$\left\langle \tilde{f}(\mathbf{v}_1), \mathbf{v}_2 \right\rangle = \langle f(\mathbf{v}_1), \mathbf{v}_2\rangle = \langle \mathbf{v}_1, f(\mathbf{v}_2)\rangle = \left\langle \mathbf{v}_1, \tilde{f}(\mathbf{v}_2) \right\rangle.$$

We conclude that $\tilde{f}$ is a self-adjoint endomorphism on a space of dimension $k-1$; by the inductive hypotesis it is diagonalisable, with $k-1$ real eigenvalues and $k-1$ real eigenvectors which form an orthonormal basis of $U^\perp$, say B_{k-1}. However, any eigenpair of $\tilde{f}$ is also an eigenpair of f, since $\lambda\mathbf{v} = \tilde{f}(\mathbf{v}) = f(\mathbf{v})$. Therefore the spectrum of f is $\operatorname{sp} f = \{\mu\} \cup \operatorname{sp}\tilde{f}$ and we note that these are all the eigenvalues of f. The corresponding eigenvectors form a set with n elements $B_k = \{\mathbf{u}\} \cup B_{k-1}$, which is linearly independent, as $\mathbf{u} \perp B_{k-1}$ where B_{k-1} is an orthonormal basis of $U^\perp$, by the inductive hypothesis. Hence B_k is an orthonormal basis of V. ■

We immediately obtain the following result, by using the results of the Spectral Theorem and of Theorem 19.1.

Corollary 19.4 Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then A is diagonalisable, with real eigenvalues and real eigenvectors which form an orthonormal basis for $\mathbb{R}^n$.

We can immediately derive the canonical form for the matrix representation of a self-adjoint endomorphism:

$$A = QDQ^T,$$

where

- D is a diagonal matrix containing the real eigenvalues of A;
- Q is a matrix with orthonormal vectors; this is known as an orthogonal matrix; we note here that orthonormality implies $Q^TQ = QQ^T = I_n$ so that $Q^{-1} = Q^T$.

This representation of A is also known as the **spectral decomposition of** A. This can also be written in the following form

$$A = QDQ^T = \lambda_1 \mathbf{q}_1\mathbf{q}_1^T + \lambda_2 \mathbf{q}_2\mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n\mathbf{q}_n^T,$$

where $\mathbf{q}_i$ are the columns of Q.

R A similar description can be derived for the complex case. For details, see the references provided.