Chapter 4

Linear Algebra II

1 Inner product spaces

In this chapter, unless explicitly stated otherwise, K will denote either $\mathbb{R}$ or $\mathbb{C}$. We start by recalling the definition of an inner product on a vector space. Having an inner product will enable us to define geometric notions such as length.

Definition 4.1. Let V be a K-vector space. An **inner product** on V is a function $V \times V \to K$ (with the image of (u, v) being denoted $\langle u, v\rangle$) that satisfies the following conditions.

1) $\forall u, v \in V \quad \langle v, u\rangle = \overline{\langle u, v\rangle}$

2) $\forall u, v, w \in V\ \forall k, l \in K \quad \langle ku + lv, w\rangle = k\langle u, w\rangle + l\langle v, w\rangle$

3) (a) $\forall u \in V \quad \langle u, u\rangle \geqslant 0$
   (b) $\forall u \in V \quad \langle u, u\rangle = 0 \implies u = 0$

An **inner product space** is a vector space equipped with an inner product.

Remark. 1. The first condition implies that $\forall u \in V$, $\langle u, u\rangle \in \mathbb{R}$.

2. The first and second conditions imply that $\forall u, v \in V\ \forall k \in K$, $\langle u, kv\rangle = \overline{k}\langle u, v\rangle$.

Exercise 117. Show, using the above axioms, that $\forall u \in V$, $\langle 0, u\rangle = 0$.

Exercise 118. Show that the following defines an inner product on $\mathbb{C}^2$.

$$\langle (x_1, x_2), (y_1, y_2)\rangle = x_1\overline{y_1} + ix_1\overline{y_2} - ix_2\overline{y_1} + 2x_2\overline{y_2}$$

Example 4.2. 1. $V = \mathbb{R}^n$ equipped with the usual dot product.

2. The **standard inner product** on $\mathbb{C}^n$ is $\langle (u_1, \dots, u_n), (v_1, \dots, v_n)\rangle = u_1\overline{v_1} + \cdots + u_n\overline{v_n}$

3. $V = M_n(K)$, $\langle A, B\rangle = \operatorname{tr}(A(\overline{B})^t)$

4. $V = \mathcal{C}([a, b], \mathbb{C})$, $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt$

Definition 4.3. Let V be an inner product space.

1) The **length** (or norm) of a vector $u \in V$ is defined to be $\|u\| = \sqrt{\langle u, u\rangle}$

2) The **distance function** (or metric) on V is defined to be $d : V \times V \to \mathbb{R}_{\geqslant 0}$ given by $d(u, v) = \|u - v\|$

3) Two vectors $u, v \in V$ are said to be **orthogonal** if $\langle u, v\rangle = 0$

4) The **orthogonal complement** of a subspace $W \leqslant V$ is defined to be $W^\perp = \{u \in V \mid \forall w \in W\ \langle u, w\rangle = 0\}$

5) A subset $S \subseteq V$ is said to be **orthonormal** if

$$\forall u, v \in S\ \langle u, v\rangle = \begin{cases} 1 & \text{if } u = w \\ 0 & \text{if } u \neq w \end{cases}$$

Example 4.4. 1. With $V = \mathbb{C}^2$, $\langle (x_1, x_2), (y_1, y_2) \rangle = x_1\overline{y_1} + ix_1\overline{y_2} - ix_2\overline{y_1} + 2x_2\overline{y_2}$ we have $\|(1, i)\| = \sqrt{5}$

2. $V = \mathcal{C}([0, 1], \mathbb{C})$, $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt$. Let $f, g \in V$ be given by $f(t) = e^{2\pi it}$ and $g(t) = e^{4\pi it}$. Then $\|f\| = 1$ and $\langle f, g \rangle = 0$.

3. $V = \mathcal{C}(\mathbb{R}, \mathbb{R})$, $\langle f, g \rangle = \int_{-1}^1 f(t)g(t)\, dt$, $S = \{\frac{1}{\sqrt{2}}, \sin(\pi t), \cos(\pi t), \sin(2\pi t), \cos(2\pi t), \dots\}$ is an (infinite) orthonormal set.

Lemma 4.5

Let V be an inner product space and $S \subseteq V$ a subset. If S is orthonormal, then S is linearly independent.

Proof. Let $u_1, \dots, u_n \in S$ and $k_i, \dots, k_n \in K$ be such that $\sum_{i=1}^n \alpha_i u_i = 0$. Then for all j we have

$$0 = \langle 0, u_j \rangle = \langle \sum_{i=1}^n k_i u_i, u_j \rangle = \sum_{i=1}^n k_i \langle u_i, u_j \rangle = k_j$$

□

Exercise 119. Find the length of

(a) $(2 + i, 3 - 2i, -1)$ in the standard inner product on $\mathbb{C}^3$.

(b) $x^2 - 3x + 1 \in \mathcal{P}_2(R)$ using inner product $\langle p(x), q(x) \rangle = \int_0^1 p(x)q(x)\, dx$.

(c) $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \in M_2(\mathbb{C})$ using inner product $\langle A, B \rangle = \mathrm{tr}(A(\overline{B})^t)$.

Exercise 120. An exercise (from an anonymous textbook) claims that, for all elements u, v of an inner product space, $\|u + v\| + \|u - v\| = 2\|u\| + 2\|v\|$. Prove that this is false. Can you guess what was intended?

2 Gram-Schmidt

Bases that are orthonormal are convenient to work with (see Proposition 4.8 below, for example). Although not all vector spaces admit an orthonormal basis, all finite dimensional vector spaces do.

Theorem 4.6: Gram-Schmidt

Let V be a finite dimensional inner product space. Any orthonormal set $S \subset V$ can be extended to a basis.

Remark. It follows that every finite dimensional inner product space has an orthonormal basis.

Proof. Let $S \subset V$ be an orthonormal set. Then S is linearly independent and therefore $|S| \leqslant \dim(V)$. Say $S = \{u_i, \dots, u_k\}$. We want to show that there is a basis $\mathcal{B}$ with $\mathcal{B} \supseteq S$. If S is a spanning set, we take $\mathcal{B} = S$. Otherwise, let $w \in V \setminus \mathrm{span}(S)$ and let $v = w - \sum_{i=1}^k \langle w, u_i \rangle u_i$. Note that $v \neq 0$ since $w \notin \mathrm{span}(S)$. Also, $\forall j \in \{1, \dots, k\}$ we have

$$\langle v, u_j \rangle = \langle w, u_j \rangle - \sum_{i=1}^k \langle w, u_i \rangle \langle u_i, u_j \rangle = \langle w, u_j \rangle - \langle w, u_j \rangle = 0$$

Defining $u_{k+1} = v/\|v\|$, the set $\{u_1, \dots, u_k, u_{k+1}\}$ is orthonormal. If S' is a spanning set for V, then it is a basis and we are done. Otherwise we repeat the above with S' in place of S. □

Example 4.7. Consider $\mathcal{P}_2(\mathbb{R})$ quipped with the inner product $\langle p(x), q(x) \rangle = \int_0^1 p(x)q(x)\, dx$. The set $S = \{1\}$ is an orthonormal set. We extend to an orthonormal basis in the way described in the above proof. Note that

$x \notin \operatorname{span}\{1\}$ and $x^2 \notin \operatorname{span}\{1, x\}$. We have

$$\begin{aligned}
v_1 &= x - \langle x, 1\rangle 1 = x - \frac{1}{2} \\
\|v_1\|^2 = \langle v_1, v_1\rangle &= \int_0^1 (x - \frac{1}{2})^2\, dx = \frac{1}{12} \\
u_1 &= v_1/\|v_1\| = \sqrt{3}(2x-1) \\
v_2 &= x^2 - \langle x^2, 1\rangle 1 - \langle x^2, u_1\rangle u_1 = x^2 - x + \frac{1}{6} \\
\|v_2\|^2 &= \int_0^1 (x^2 - x + \frac{1}{6})^2\, dx = \frac{1}{180} \\
u_2 &= v_2/\|v_2\| = \sqrt{5}(6x^2 - 6x + 1)
\end{aligned}$$

The set $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$ is an orthonormal basis for $\mathcal{P}_2(x)$.

Proposition 4.8

Let V be an inner product space and $S = \{u_1, \dots, u_n\}$ an orthonormal set. Let $v \in V$.

1) $\sum_{i=1}^n |\langle v, u_i\rangle| \leqslant \|v\|^2$
2) If S is a basis, then $v = \sum_{i=1}^n \langle v, u_i\rangle u_i$

Proof. We have

$$\begin{aligned}
\|v - \sum_{i=1}^n \langle v, u_i\rangle u_i\|^2 &= \langle v - \sum_{i=1}^n \langle v, u_i\rangle u_i, v - \sum_{i=1}^n \langle v, u_i\rangle u_i\rangle \\
&= \langle v, v\rangle - \sum_{i=1}^n \overline{\langle v, u_i\rangle}\langle v, u_i\rangle - \sum_{i=1}^n \langle v, u_i\rangle\langle u_i, v\rangle + \sum_{i=1}^n\sum_{j=1}^n \langle v, u_i\rangle\overline{\langle v, u_j\rangle}\langle u_i, u_j\rangle \\
&= \langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2 - \sum_{i=1}^n |\langle v, u_i\rangle|^2 + \sum_{i=1}^n |\langle v, u_i\rangle|^2 \\
&= \langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2
\end{aligned}$$

Therefore $\langle v, v\rangle - \sum_{i=1}^n |\langle v, u_i\rangle|^2 \geqslant 0$. The final statement is left as an exercise. □

Corollary 4.9: Cauchy-Schwartz

Let V be an inner product space. Then $\forall\, u, v \in V, \quad |\langle u, v\rangle| \leqslant \|u\|\|v\|$

Proof. If $u = 0$, then the inequality holds since both sides are zero. So we can assume that $u \neq 0$. Apply Proposition 4.8 with $S = \{u/\|u\|\}$. □

Example 4.10.

1. If we take $V = \mathbb{R}^n$ and the dot product, this becomes
$$|\sum_{i=1}^n a_i b_i| \leqslant (\sum_{i=1}^n a_i^2)^{\frac{1}{2}} (\sum_{i=1}^n b_i^2)^{\frac{1}{2}}$$
for any real numbers a_i, b_i.
2. If we take the inner product space of Example 4.4.2 above, then we have
$$\left|\int_0^1 f(t)\overline{g(t)}\, dt\right| \leqslant \left(\int_0^1 f(t)^2\, dt\right)^{\frac{1}{2}} \left(\int_0^1 g(t)^2\, dt\right)^{\frac{1}{2}}$$
for any $f, g \in \mathcal{C}([0,1], \mathbb{C})$.

Exercise 121. Let V be an inner product space. Show that the distance function $d : V \times V \to \mathbb{R}$ (defined by $d(u, u) = \sqrt{\langle u, u\rangle}$) satisfies the following properties:

(a) $d(u,v) = 0 \iff u = v$ (b) $d(u,v) = d(v,u)$ (c) $d(u,v) \leqslant d(u,w) + d(w,v)$

3 Orthogonal complements

Definition 4.11. Let V be an inner product space and let $W \leqslant V$ be a subspace. The **orthogonal complement** of W in V is denoted $W^\perp$ and defined to be

$$W^\perp = \{u \in V \mid \forall w \in W,\ \langle u, w\rangle = 0\}$$

Exercise 122. Show that

(a) $W^\perp$ is a subspace of V (b) $W \cap W^\perp = \{0\}$ (c) $W \subseteq (W^\perp)^\perp$

Proposition 4.12

Let V be a finite dimensional inner product space and let $W \leqslant V$ be a subspace. Then $V = W \oplus W^\perp$.

Proof. We know that $W \cap W^\perp = \{0\}$ from Exercise 122. It remains to show that $V = W + W^\perp$. From Theorem 4.6 we know that W has an orthonormal basis, say $\{w_1, \dots, w_k\}$. Given $u \in V$ define $w = \sum_{i=1}^k \langle u, w_i\rangle w_i$. Then $w \in W$ and $\langle u - w, w_i\rangle = 0$ for all i. Therefore $u - w \in W^\perp$ and we have $u = w + (u - w) \in W + W^\perp$. □

Remark. It follows from the proposition that $\dim(V) = \dim(W) + \dim(W^\perp)$.

Exercise 123. Show that if V is a finite dimensional inner product space and $W \leqslant V$ is a subspace of V, then $(W^\perp)^\perp = W$.

Example 4.13. This is an example in which $W \neq (W^\perp)^\perp$. Denote by ℓ^2 the vector space of all square-summable real-valued sequences, that is

$$\ell^2 = \{(x_1, x_2, \dots) \mid x_i \in \mathbb{R} \quad \text{and} \quad \sum_{i=1}^{\infty} |x_i|^2 \text{ converges}\}$$

The following is an inner product on ℓ^2

$$\langle (x_1, x_2, \dots), (y_1, y_2, \dots)\rangle = \sum_{i=1}^{\infty} x_i y_i$$

Let W be the subspace of ℓ^2 consisting of all sequences that are eventually zero, that is,

$$W = \{(x_1, x_2, \dots) \mid x_i \in \mathbb{R}, \quad \exists N \in \mathbb{N} \text{ such that } i \geqslant N \implies x_i = 0\}$$

Now define $v \in \ell^2$ to be the sequence $v = (1/i)_{i\in\mathbb{N}}$. Clearly, $v \notin W$, however $v \in (W^\perp)^\perp$ because for any $(\xi_i) \in W^\perp$ we have

$$\langle v, \xi\rangle = \sum_{i=1}^{\infty} \xi_i v_i = \lim_{N\to\infty} \sum_{i=1}^{N} \xi_i v_i = \lim_{N\to\infty} \langle \xi, u_i\rangle = \lim_{N\to\infty} 0 = 0$$

where $u_i \in W$ is the sequence given by $(u_i)_j = \begin{cases} v_j & j \leqslant i \\ 0 & j > i \end{cases}$

Therefore $W \subsetneqq (W^\perp)^\perp$.

4 Adjoint transformations

Definition 4.14. Let V be an inner product space and $f : V \to V$ a linear transformation. An **adjoint** of f is a linear transformation $f^* : V \to V$ satisfying

$$\forall u, v \in V \quad \langle f(u), v\rangle = \langle u, f^*(v)\rangle$$

For a matrix $A \in M_n(K)$ the notation A^* is used to denote the matrix $A^* = (\overline{A})^t$.

Lemma 4.15

Let V be a finite inner product space and $f : V \to V$ a linear transformation.

1. If an adjoint of f exists, it is unique. (This justifies the notation f^*.)
2. If V is finite dimensional, then an adjoint of f exists.

Proof. For the first part, suppose that $g, h : V \to V$ are such that

$$\forall u, v \in V \quad \langle f(u), v\rangle = \langle u, g(v)\rangle = \langle u, h(v)\rangle$$

Let $v \in V$ and define $u = g(v) - h(v)$. We have

$$\begin{aligned}\langle u, u\rangle &= \langle u, g(v)\rangle - \langle u, h(v)\rangle \\ &= \langle f(u), v\rangle - \langle f(u), v\rangle \\ &= 0\end{aligned}$$

From which it follows that $g(v) = h(v)$. Since this holds for all $v \in V$, we have that $g = h$.
We now establish the second part. Since V is finite dimensional, there is an orthonormal basis. Let $\mathcal{B} = \{b_1, \dots, b_n\}$ be an orthonormal basis for V and let $A = [f]_\mathcal{B}$. Let $g : V \to V$ be the linear transformation determined by the condition that $[g]_\mathcal{B} = A^*$. We will now show that g is an adjoint for f. Denote the entries in the matrix A by A_{ij}.

$$\begin{aligned}\langle f(b_i), b_j\rangle &= \langle \sum_{k=1}^{n} A_{ki} b_k, b_j\rangle = \sum_{k=1}^{n} A_{ki}\langle b_k, b_j\rangle = A_{ji} \\ \langle b_i, g(b_j)\rangle &= \langle b_i, \sum_{i=1}^{k} (A^*)_{kj} b_k\rangle = \sum_{i=1}^{k} \overline{(A^*)_{kj}}\langle b_i, b_k\rangle = \overline{(A^*)_{ij}} = A_{ji}\end{aligned}$$

Therefore, for all $i, j \in \{1, \dots, n\}$ we have $\langle f(b_i), b_j\rangle = \langle b_i, g(b_j)\rangle$. It follows that for all $u, v \in V$ we have $\langle f(u), v\rangle = \langle u, g(v)\rangle$. $\square$

Remark. As part of the above proof we showed that $[f^*]_\mathcal{B} = ([f]_\mathcal{B})^*$ for any orthonormal basis $\mathcal{B}$ of a finite dimensional V.

Example 4.16.

1. $f : \mathbb{C}^2 \to \mathbb{C}^2$, $f(x, y) = (x, 0)$ has adjoint $f^* = f$.
2. $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by a rotation has adjoint $f^* = f^{-1}$
3. Let W be as in Example 4.13. The linear transformation $f : W \to W$ given by $f(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$ has adjoint given by $f^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$.
4. Let $V = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is infinitely differentiable and } \forall n \in \mathbb{Z}\ f(x + n) = f(x)\}$ with inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Let $\Delta : V \to V$ be given by $\Delta(f) = \frac{d^2 f}{dt^2}$. Then $\Delta^* = \Delta$.

Lemma 4.17: Properties of the adjoint

Let V be an inner product space and let $f, g : V \to V$ be two linear transformations and $k \in K$. Then

1. $(f + g)^* = f^* + g^*$
2. $(kf)^* = \overline{k} f^*$
3. $(f \circ g)^* = g^* \circ f^*$
4. $(f^*)^* = f$

Exercise 124. Write out a proof of the above lemma. Note that there is no assumption that V be finite dimensional, merely that f^* and g^* exist.

Definition 4.18. Let $f : V \to V$ be a linear transformation on an inner product space. We say that f is:

1. **self-adjoint** if $f^* = f$ (also called **symmetric** if $K = \mathbb{R}$ or **hermitian** if $K = \mathbb{C}$)
2. **isometric** if $f^* \circ f = \mathrm{Id}_V$ (also called **orthogonal** if $K = \mathbb{R}$ or **unitary** if $K = \mathbb{C}$)
3. **normal** if $f^* \circ f = f \circ f^*$

Remark. It follows from the definitions that

1. If f is self-adjoint, then it is normal.
2. If V is finite dimensional and f is an isometry, then $f^* = f^{-1}$ and f is normal.

Example 4.19. Considering the linear transformations in Example 4.16 we see that:

1. f is self-adjoint and therefore normal,
2. f is an isometry and therefore normal (since $\mathbb{R}^2$ is finite dimensional),
3. f is an isometry since $f^* \circ f = \mathrm{Id}$, but f is not invertible and not normal,
4. Δ is self-adjoint and therefore normal.

Lemma 4.20

Let $f : V \to V$ be a linear transformation on an inner product space. The following are equivalent:

1. $f^* \circ f = \mathrm{Id}_V$ (i.e., f is an isometry as defined above)
2. $\forall u, v \in V$, $\langle f(u), f(v)\rangle = \langle u, v\rangle$
3. $\forall v \in V$, $\|f(v)\| = \|v\|$

Proof. If the first holds, then we have $\langle f(u), f(v)\rangle = \langle u, f^* \circ f(v)\rangle = \langle u, \mathrm{Id}_V(v)\rangle = \langle u, v\rangle$, so the second holds.
If the second holds, then we have $\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = \|v\|^2$, so the third holds.
Now suppose that the third condition holds and define $g = f^* \circ f - \mathrm{Id}_V$. We will show that $g = 0$. From Lemma 4.17 we have that g is self-adjoint: $g^* = (f^* \circ f) - \mathrm{Id}_V^* = f^* \circ (f^*)^* - \mathrm{Id}_V = f^* \circ f - \mathrm{Id}_V = g$. For any $u, v \in V$ we have

$$\langle g(v), v\rangle = \langle f^* \circ f(v) - v, v\rangle = \langle f^* \circ f(v), v\rangle - \langle v, v\rangle = \langle f(v), f(v)\rangle - \langle v, v\rangle = \|f(v)\|^2 - \|v\|^2 = 0$$

and therefore

$$0 = \langle g(u+v), u+v\rangle = \langle g(u), v\rangle + \langle g(v), u\rangle = \langle g(u), v\rangle + \langle v, g(u)\rangle = \langle g(u), v\rangle + \overline{\langle g(u), v\rangle}$$

Letting $v = g(u)$ we obtain

$$0 = \langle g(u), g(u)\rangle + \overline{\langle g(u), g(u)\rangle} = 2\langle g(u), g(u)\rangle$$

Therefore $g(u) = 0$ for all $u \in V$ and hence $g = 0$. □

Lemma 4.21

Let $f : V \to V$ be a linear transformation on a inner product space and $W \leqslant V$ a subspace. If W is f-invariant, then $W^\perp$ is f^*-invariant.

Proof. Let $u \in W$ and $v \in W^\perp$. Then $\langle u, f^*(v)\rangle = \langle f(u), v\rangle = 0$ since $f(u) \in W$ and $v \in W^\perp$. □

4.1 Exercises

Exercise 125. If A is a transition matrix between orthonormal bases, show that A is isometric (i.e., $A^*A = I$).

Exercise 126. Suppose that f is a linear transformation on an inner product space V. Prove the following.

(a) If f is self-adjoint, then all eigenvalues of f are real.

(b) If f is isometric, then all eigenvalues of f have absolute value 1.

Exercise 127. Suppose that f is a linear transformation on a finite dimensional inner product space V. Show that the range of f^* is the orthogonal complement of the kernel of f. Deduce that the rank of f is equal to the rank of f^*. Deduce that the row-rank of a square matrix is equal to its column rank.

Exercise 128. Consider the inner product space $\mathcal{P}(K)$ having inner product $\langle p(x), q(x)\rangle = \int_0^1 p(x)\overline{q(x)}\,dx$. Show that the linear transformation $\delta : \mathcal{P}(K) \to \mathcal{P}(K)$ given by differentiation has no adjoint. (Hint: Try to find what $\delta^*(1)$ should be.)

Exercise 129. Show that a triangular matrix which is self-adjoint or unitary is diagonal.

Exercise 130. Let V be a finite dimensional inner product space and f a linear transformation on V. Show that, given a vector $w \in V$, there exists a unique vector $w_1 \in V$ such that $\langle f(v), w\rangle = \langle v, w_1\rangle$ for all $v \in V$. (Hint: First show that it will be enough to consider only those v that lie in some fixed orthonormal basis of V.)

Exercise 131. Let g be a self-adjoint linear transformation on a finite dimensional inner product space V. Suppose that $\langle g(v), v\rangle = 0$ for all $v \in V$.

(a) Show that $\langle g(u), w\rangle + \langle g(w), u\rangle = 0$ for all $u, w \in V$.

(b) Deduce that g is the zero linear transformation if the space is a real space. (This is the time to use the fact that g is self-adjoint).

(c) Assume now that the space is complex; deduce that $\langle g(u), w\rangle$ is imaginary for all $u, w \in V$.

(d) Deduce that $\langle g(iu), w\rangle$ is imaginary for all $u, w \in V$ and so $\langle g(u), w\rangle = 0$ for all $u, w \in V$.

(e) Deduce that g is zero in the complex case also.

Exercise 132. Let f be a linear transformation on a finite dimensional inner product space V. Suppose that W is an f-invariant and f^*-invariant subspace of V. Show that $(f_W)^* = (f^*)_W$.

Exercise 133. Let f be an isometry on a finite dimensional inner product space V. Suppose that W is an f-invariant subspace of V. Show that f_W is also an isometry.

Exercise 134. Let V be a two dimensional real inner product space and let f be an isometry of V. Show that f can be represented by a matrix of the form $\begin{bmatrix} \cos\theta & -\sin\theta \\ \epsilon\sin\theta & \epsilon\cos\theta \end{bmatrix}$ where $\epsilon = \pm 1$.

5 Spectral theorem

We now come to the question of when a linear transformation can be diagonalised. We have seen necessary and sufficient conditions in terms of the minimal polynomial of the transformation. The spectral theorem gives a sufficient condition for diagonalisability (without reference to the minimal polynomial).

Theorem 4.22: Spectral theorem for normal linear transformations

Let V be a finite dimensional, complex inner product space vector space and let $f : V \to V$ be a linear transformation. If f is normal, then there exists an orthonormal basis $\mathcal{B}$ for V such that $[f]_{\mathcal{B}}$ is diagonal.

Proof. We use (strong) induction on $n = \dim(V)$. If $n = 1$, then the statement is trivially true. Assume now that $n > 1$ and that the statement holds for all cases in which the dimension is less than n. Let $\lambda \in \mathbb{C}$ be an eigenvalue of f and V_λ the corresponding eigenspace. By Proposition 4.12 we have $V = V_\lambda \oplus V_\lambda^\perp$. Note that $\dim(V_\lambda) < \dim(V)$ and $\dim(V_\lambda^\perp) < \dim(V)$. We will show that both V_λ and $V_\lambda^\perp$ are f-invariant, and then apply Lemma 2.16. That V_λ is f-invariant is clear (see Exercise 42). To show that $V_\lambda^\perp$ is f-invariant, we note first that V_λ is f^*-invariant since (using that f is normal):

$$u \in V_\lambda \implies f(f^*(u)) = f^*(f(u)) = f^*(\lambda u) = \lambda f^*(u) \implies f^*(u) \in V_\lambda$$

That $V_\lambda^\perp$ is f-invariant then follows from Lemma 4.21 since $(f^*)^* = f$. Let $f_1 : V_\lambda \to V_\lambda$ and $f_2 : V_\lambda^\perp \to V_\lambda^\perp$ be the restrictions of f to V_λ and $V_\lambda^\perp$ respectively. By the induction hypothesis, there exist orthonormal bases $\mathcal{B}_1$ and $\mathcal{B}_2$ for V_λ and $V_\lambda^\perp$ respectively, such that $[f_i]_{\mathcal{B}_i}$ is diagonal. By Lemma 2.16 $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for V and $[f]_{\mathcal{B}} = [f_1]_{\mathcal{B}_1} \oplus [f_2]_{\mathcal{B}_2}$. In particular, $[f]_{\mathcal{B}}$ is diagonal. $\square$

Theorem 4.23: Spectral theorem for normal matrices

Let $A \in M_n(\mathbb{C})$ be such that $AA^* = A^*A$. There exists a matrix $U \in M_n(\mathbb{C})$ such that $U^*U = I$ (i.e., U is unitary) and U^*AU is diagonal.

Proof. Define $f : M_{n\times 1} \to M_{n\times 1}$ by $f(X) = AX$ and apply Theorem 4.22. Letting U be the matrix whose columns are the elements of $\mathcal{B}$, we have $[f]_{\mathcal{B}} = U^{-1}AU$. That $U^{-1} = U^*$ follows from the fact that $\mathcal{B}$ is an orthonormal basis. □

Remark. The columns of U form an orthonormal basis and the diagonal entries or U^*AU are exactly the eigenvalues of A.

Example 4.24. Let $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Then $A^*A = AA^*$. By the spectral theorem there is unitary matrix $U \in M_2(\mathbb{C})$ such that U^*AU is diagonal. To find such a U we calculate an orthonormal basis of eigenvectors. The eigenvalues of the matrix A are $1-i, 1+i$. An orthonormal basis for the $(1-i)$-eigenspace is $\{\left[\begin{smallmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{smallmatrix}\right]\}$. An orthonormal basis for the $(1+i)$-eigenspace is $\{\left[\begin{smallmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{smallmatrix}\right]\}$. So we can take

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}$$

Note however that the matrix A is *not* diagonalisable over $\mathbb{R}$. That is, there does not exist an invertible matrix $P \in M_2(\mathbb{R})$ such that $P^{-1}AP$ is diagonal.

For real inner product spaces we have the following.

Theorem 4.25: Spectral theorem for symmetric linear transformations

Let V be a finite dimensional real inner product space and let $f : V \to V$ be a self-adjoint linear transformation. Then there exists an orthonormal basis $\mathcal{B}$ of V such that $[f]_{\mathcal{B}}$ is diagonal.

Outline of proof. We use induction on $n = \dim(V)$. If $n = 1$, the result holds trivially.
Since f is self-adjoint, all eigenvalues are real (Exercise 126). Let $\lambda \in \mathbb{R}$ be an eigenvalue of f and let $u \in V$ be such that $f(u) = \lambda u$. Let $W = \operatorname{span}(u)$. Then $V = W \oplus W^\perp$ (Proposition 4.12) and W and $W^\perp$ are both f-invariant (Lemma 4.21). By the induction hypothesis, there exists an orthonormal basis $\mathcal{C} = \{c_1, \ldots, c_{n-1}\}$ for $W^\perp$ such that $D = [f|_{W^\perp}]_{\mathcal{C}}$ is diagonal. Letting $\mathcal{B} = \{c_1, \ldots, c_{n-1}, u/\|u\|\}$ we have that $[f]_{\mathcal{B}} = D \oplus [\lambda]$ (Lemma 2.16). It remains to show that $\mathcal{B}$ is orthonormal. This follows from the fact that both $\mathcal{C}$ and $\{u/\|u\|\}$ are orthonormal and that $\langle c_i, u\rangle = 0$ for all i. □

5.1 Exercises

Exercise 135. Show that if $A = UDU^*$ where D is a diagonal matrix and U is unitary, then A is a normal matrix. (The spectral theorem implies that the converse is true).

Exercise 136. Show that a linear transformation $f : V \to V$ on a complex inner product space V is normal if and only if $\langle f(u), f(v)\rangle = \langle f^*(u), f^*(v)\rangle$ for all $u, v \in V$.

Exercise 137. (a) Show that every normal matrix A has a square root; that is, a matrix B so that $B^2 = A$.

(b) Must every complex square matrix have a square root?

Exercise 138. Two linear transformations f and g on a finite dimensional complex inner product space are **unitarily equivalent** if there is a unitary linear transformation u such that $g = u^{-1}fu$. Two matrices are **unitarily equivalent** if their linear transformations, with respect to some fixed orthonormal basis, are **unitarily equivalent**.
Decide whether or not the following matrices are unitarily equivalent.

(a) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$

Exercise 139. Are f and f^* always unitarily equivalent?

Exercise 140. If f is a normal linear transformation on a finite dimensional complex inner product space, and if $f^2 = f^3$, show that $f = f^2$. Show also that f is self-adjoint.

Exercise 141. If f is a normal linear transformation on a finite dimensional complex inner product space show that $f^* = p(f)$ for some polynomial p.

Exercise 142. If f and g are normal linear transformations on a finite dimensional complex inner product space and $fg = gf$, show that $f^*g = gf^*$. (Harder) Prove that the same result holds assuming only that f is normal.

Exercise 143. Let f be a linear transformation on a finite dimensional complex inner product space. Suppose that f commutes with f^*f; that is, that $f(f^*f) = (f^*f)f$. We aim to show that f is normal.

(a) Show that f^*f is normal.

(b) Choose an orthonormal basis so that the matrix of f^*f takes the block diagonal form $\operatorname{diag}(A_1, \dots, A_m)$ where $A_i = \lambda_i I_{m_i}$ and $\lambda_i = \lambda_j$ only if $i = j$.

(c) Show that f has matrix, with respect to this basis, of the block diagonal form $\operatorname{diag}(B_1, \dots, B_m)$ for some $m_i \times m_i$ matrices B_i.

(d) Deduce that $B_i^* B_i = A_i$ and so that $B_i^* B_i = B_i B_i^*$.

(e) Deduce that f is normal.