5: INNER PRODUCTS, ADJOINTS, SPECTRAL THEOREMS, SELF-ADJOINT OPERATORS

STEVEN HEILMAN

Contents

1. Review

Proposition 1.1. *Let A be an $n \times n$ matrix. Let $v_1, \ldots, v_k$ be eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_k$, respectively. If $\lambda_1, \ldots, \lambda_k$ are all distinct, then the vectors $v_1, \ldots, v_k$ are linearly independent.*

Lemma 1.2 (An Eigenvector Basis Diagonalizes T). *Let V be an n-dimensional vector space over a field $\mathbf{F}$, and let $T\colon V \to V$ be a linear transformation. Suppose V has an ordered basis $\beta := (v_1, \ldots, v_n)$. Then v_i is an eigenvector of T with eigenvalue $\lambda_i \in \mathbf{F}$, for all $i \in \{1, \ldots, n\}$, if and only if the matrix $[T]_\beta^\beta$ is diagonal with $[T]_\beta^\beta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

Lemma 1.3. *Let V be a finite-dimensional vector space over a field $\mathbf{F}$. Let β, β' be two bases for V. Let $T\colon V \to V$ be a linear transformation. Define $Q := [I_V]_\beta^{\beta'}$. Then $[T]_\beta^\beta$ and $[T]_{\beta'}^{\beta'}$ satisfy the following relation*

$$[T]_{\beta'}^{\beta'} = Q[T]_\beta^\beta Q^{-1}.$$

Theorem 1.4 (The Fundamental Theorem of Algebra). *Let $f(\lambda)$ be a real polynomial of degree n. Then there exist $\lambda_0, \lambda_1, \ldots \lambda_n \in \mathbf{C}$ such that*

$$f(\lambda) = \lambda_0 \prod_{i=1}^{n} (\lambda - \lambda_i).$$

Lemma 1.5. *Let A, B be similar matrices. Then A, B have the same characteristic polynomial.*

Date: June 1, 2015.

Theorem 1.6. *Let A be an $n \times n$ matrix. There exist scalars $a_1, \ldots, a_{n-2} \in \mathbf{F}$ such that the characteristic polynomial $f(\lambda)$ of A satisfies*

$$f(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\mathrm{Tr}(A)\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + \det(A).$$

2. Inner Product Spaces

Up until this point, we have focused on the linear properties of vector spaces. Our investigation now becomes much deeper when we consider more geometric properties of vector spaces. That is we now start to consider the size of vectors, and how close one vector can be to another. These notions are made rigorous by the introduction of norms and inner products, respectively. Also, with this more geometric information, linear algebra and analysis will start to relate much more with each other.

We first introduce the concept of a general norm, which measures the length of vectors. We then introduce the more specific concept of an inner product, which measures the angle between two vectors.

Definition 2.1 (Normed Linear Space). Let $\mathbf{F}$ denote either $\mathbf{R}$ or $\mathbf{C}$. Let V be a vector space over $\mathbf{F}$. A **normed linear space** is a vector space V equipped with a **norm**. A norm is a function $V \to \mathbf{R}$, denoted by $\|\cdot\|$, which satisfies the following properties.

(a) For all $v \in V$, for all $\alpha \in \mathbf{F}$, $\|\alpha v\| = |\alpha| \|v\|$. (Homogeneity)
(b) For all $v \in V$ with $v \neq 0$, $\|v\|$ is a positive real number; $\|v\| > 0$. And $v = 0$ if and only if $\|v\| = 0$. (Positive definiteness)
(c) For all $v, w \in V$, $\|v + w\| \leq \|v\| + \|w\|$. (Triangle Inequality)

Example 2.2. Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. Define the 2-norm on $\mathbf{R}^n$ by

$$\|x\|_2 := \sqrt{x_1^2 + \cdots + x_n^2}.$$

So, for $n = 1$, we have $\|x\|_2 = |x|$. We will see below one way to show the triangle inequality for the 2-norm.

Define the 1-norm on $\mathbf{R}^n$ by

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|.$$

Define the ∞-norm on $\mathbf{R}^n$ by

$$\|x\|_\infty := \max_{i=1,\ldots,n} |x_i|.$$

Exercise 2.3. Let $x, y \in \mathbf{R}$. Verify that $|x + y| \leq |x| + |y|$. Deduce that the triangle inequality holds for the 1-norm and the ∞-norm.

Exercise 2.4. Let V be a normed linear space. Show that, for any $v, w \in V$, $\|v - w\| \geq |\|v\| - \|w\||$.

Definition 2.5 (Complex Conjugate). Let $i := \sqrt{-1}$. Let $x, y \in \mathbf{R}$, and let $z = x + iy \in \mathbf{C}$. Define $\overline{z} := x - iy$. Define $|z| := \sqrt{x^2 + y^2}$. Note that $|z|^2 = z\overline{z}$.

Definition 2.6 (Inner Product). Let $\mathbf{F}$ denote either $\mathbf{R}$ or $\mathbf{C}$. Let V be a vector space over $\mathbf{F}$. An **inner product space** is a vector space V equipped with an **inner product**. An inner product is a function $V \times V \to \mathbf{F}$, denoted by $\langle \cdot, \cdot \rangle$, which satisfies the following properties.

(a) For all $v, v', w \in V$, $\langle v+v', w\rangle = \langle v, w\rangle + \langle v', w\rangle$. (Linearity in the first argument).
(b) For all $v, w \in V$, for all $\alpha \in \mathbf{F}$, $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$. (Homogeneity in the first argument)
(c) For all $v \in V$, if $v \neq 0$, then $\langle v, v\rangle$ is a positive real number; $\langle v, v\rangle > 0$. (Positivity)
(d) For all $v, w \in V$, $\langle v, w\rangle = \overline{\langle w, v\rangle}$. (Conjugate symmetry)

Exercise 2.7. Using the above properties, show the following things.
(e) For all $v, v', w \in V$, $\langle w, v+v'\rangle = \langle w, v\rangle + \langle w, v'\rangle$. (Linearity in the second argument)
(f) For all $v, w \in V$, for all $\alpha \in \mathbf{F}$, $\langle v, \alpha w\rangle = \overline{\alpha}\langle v, w\rangle$.
(g) For all $v \in V$, $\langle v, 0\rangle = \langle 0, v\rangle = 0$.
(h) $\langle v, v\rangle = 0$ if and only if $v = 0$.

Remark 2.8. If $\mathbf{F} = \mathbf{R}$, then property (d) says that $\langle v, w\rangle = \langle w, v\rangle$.

Example 2.9. Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, and let $y = (y_1, \ldots, y_n) \in \mathbf{R}^n$. Define the standard inner product (or dot product) on $\mathbf{R}^n$ by

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i.$$

More generally, if $\alpha_1, \ldots, \alpha_n > 0$, then the following definition also gives an inner product

$$\langle x, y\rangle_\alpha := \sum_{i=1}^{n} \alpha_i x_i y_i.$$

Example 2.10. Let $w = (w_1, \ldots, w_n) \in \mathbf{C}^n$, and let $z = (z_1, \ldots, z_n) \in \mathbf{C}^n$. Define the standard inner product (or dot product) on $\mathbf{C}^n$ by

$$\langle w, z\rangle := \sum_{i=1}^{n} w_i \overline{z_i}.$$

Example 2.11. Let $A, B \in M_{n\times n}(\mathbf{R})$. Define the standard inner product on $M_{n\times n}(\mathbf{R})$ by

$$\langle A, B\rangle := \mathrm{Tr}(B^t A).$$

Example 2.12. Let $f, g \in C([0,1], \mathbf{R})$. That is, f, g are continuous real valued functions on $[0,1]$. Define

$$\langle f, g\rangle := \int_0^1 f(t)g(t)dt.$$

Definition 2.13 (Orthogonal Vectors). Let V be an inner product space, and let $v, w \in V$. We say that v, w are **orthogonal** if $\langle v, w\rangle = 0$.

Lemma 2.14 (The Cauchy-Schwarz Inequality). *Let V be an inner product space. Then, for any $v, w \in V$,*

$$|\langle v, w\rangle| \leq \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}.$$

Proof. If $w = 0$, then both sides of our inequality are zero, so the inequality holds, and we may assume $w \neq 0$. So, $\langle w, w\rangle > 0$. Define $\alpha := \langle v, w\rangle / \langle w, w\rangle$. Note that by conjugate-symmetry of the inner product,

$$\langle v, w\rangle\langle w, v\rangle = \langle v, w\rangle\overline{\langle v, w\rangle} = |\langle v, w\rangle|^2 .$$

Also, by positivity of the inner product, $\langle v-\alpha w, v-\alpha w\rangle \geq 0$. That is,

$$\langle v,v\rangle - \alpha\langle w,v\rangle - \overline{\alpha}\langle v,w\rangle + |\alpha|^2 \langle w,w\rangle \geq 0$$

Substituting in the definition of α,

$$\langle v,v\rangle - |\langle v,w\rangle|^2/\langle w,w\rangle - |\langle v,w\rangle|^2/\langle w,w\rangle + |\langle v,w\rangle|^2/\langle w,w\rangle \geq 0.$$

Simplifying, we get $\langle v,v\rangle - |\langle v,w\rangle|^2/\langle w,w\rangle \geq 0$, as desired. □

Remark 2.15. If the choice of α looks a bit mysterious, note that if $v=\beta w$ for some $\beta\in\mathbf{R}$, then the Cauchy-Schwarz inequality becomes an equality. And if $\langle v,w\rangle = 0$, then the Cauchy-Schwarz inequality has zero on the left side and a possibly large number on the right. So, the positive number $\sqrt{\langle v,v\rangle}\sqrt{\langle w,w\rangle} - |\langle v,w\rangle|$ somehow measures how close v and w are to being parallel. Also, in the proof of the Cauchy-Schwarz inequality, αw is parallel to w and $v-\alpha w$ is orthogonal to w, so writing

$$v = (v-\alpha w)+\alpha w,$$

we see that the size of $v-\alpha w$ also measures how close v and w are to being parallel.

We now show that an inner product space is a normed linear space, with norm $\|v\| := \sqrt{\langle v,v\rangle}$.

Lemma 2.16. *Let $\langle,\rangle$ be an inner product on a vector space V. Then the function $\|\cdot\|\colon V\to\mathbf{R}$ defined by $\|v\| := \sqrt{\langle v,v\rangle}$ is a norm on V.*

Proof. Homogeneity and positive definiteness follow readily from the definition of the inner product, and from Exercise 2.7(h). It therefore suffices to show that the triangle inequality holds. Let $v,w\in V$. We need to show that $\|v+w\|\leq\|v\|+\|w\|$. In order to show this inequality, it suffices to show that its square holds.

$$\begin{aligned}\|v+w\|^2 &= |\langle v+w,v+w\rangle| = |\langle v,v\rangle+\langle w,w\rangle+\langle v,w\rangle+\langle w,v\rangle|\\ &\leq |\langle v,v\rangle|+|\langle w,w\rangle|+|\langle v,w\rangle|+|\langle w,v\rangle| \qquad \text{, by Exercise 2.3}\\ &\leq \|v\|^2+\|w\|^2+2\|v\|\,\|w\| \qquad \text{, by Lemma 2.14}\\ &= (\|v\|+\|w\|)^2.\end{aligned}$$

□

Consequently, the triangle inequality holds for the norm $\|\cdot\|_2$ on $\mathbf{R}^n$, for any $n\geq 1$.

3. Orthogonality

Let V be an inner product space with inner product $\langle\,,\rangle$. Recall that $v,w\in V$ are said to be orthogonal if $\langle v,w\rangle = 0$. If v,w are orthogonal, we also sometimes say that v,w are perpendicular, and we write $v\perp w$.

Lemma 3.1. *Let V be an inner product space. Suppose $v\in V$ is orthogonal to each of the vectors $v_1,\ldots,v_n\in V$. Then v is orthogonal to any linear combination of $v_1,\ldots,v_n$.*

Proof. Since $\langle v,v_i\rangle = 0$ for all $i\in\{1,\ldots,n\}$, if $\alpha_1,\ldots,\alpha_n\in\mathbf{F}$, we have

$$\langle v,\sum_{i=1}^n \alpha_i v_i\rangle = \sum_{i=1}^n \overline{\alpha_i}\langle v,v_i\rangle = 0.$$

∎

Theorem 3.2 (**Pythagorean Theorem**). *Let V be an inner product space, and let $v, w \in V$ be orthogonal. Then $\|v+w\|^2 = \|v\|^2 + \|w\|^2$.*

Proof.

$$\|v+w\|^2 = \langle v+w, v+w\rangle = \langle v,v\rangle + \langle w,w\rangle + \langle v,w\rangle + \langle w,v\rangle = \|v\|^2 + \|w\|^2 .$$

∎

Theorem 3.3 (**Generalized Pythagorean Theorem**). *Let V be an inner product space, and let $v_1, \ldots, v_n \in V$ be orthogonal to each other. That is, $\langle v_i, v_j\rangle = 0$ for all $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Then*

$$\left\|\sum_{i=1}^{n} v_i\right\|^2 = \sum_{i=1}^{n} \|v_i\|^2$$

Proof. We induct on n. The base case has been proven, so we only need to prove the inductive step. Suppose the assertion is true for a fixed n. Then, let $v_1, \ldots, v_{n+1}$ be orthogonal to each other. From Lemma 3.1, v_{n+1} is orthogonal to $\sum_{i=1}^{n} v_i$. So, applying the Pythagorean Theorem to v_{n+1} and $\sum_{i=1}^{n} v_i$, and then using the inductive hypothesis,

$$\left\|v_{n+1} + \sum_{i=1}^{n} v_i\right\|^2 = \|v_{n+1}\|^2 + \left\|\sum_{i=1}^{n} v_i\right\|^2 = \|v_{n+1}\|^2 + \sum_{i=1}^{n} \|v_i\|^2 .$$

∎

Corollary 3.4. *Let V be an inner product space, and let $v_1, \ldots, v_n \in V$ be orthogonal to each other. That is, $\langle v_i, v_j\rangle = 0$ for all $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Let $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$. Then*

$$\left\|\sum_{i=1}^{n} \alpha_i v_i\right\|^2 = \sum_{i=1}^{n} |\alpha_i|^2 \|v_i\|^2 .$$

Proof. If $\langle v_i, v_j\rangle = 0$, then $\langle \alpha_i v_i, \alpha_j v_j\rangle = 0$. So, apply Theorem 3.3 to the set of vectors $\alpha_1 v_1, \ldots, \alpha_n v_n$. ∎

Definition 3.5 (**Orthogonal Set, Orthonormal Set**). Let V be an inner product space and let $(v_1, \ldots, v_n)$ be a collection of vectors in V. The set of vectors $(v_1, \ldots, v_n)$ is said to be **orthogonal** if $\langle v_i, v_j\rangle = 0$ for all $i, j \in \{1, \ldots, n\}$ with $i \neq j$. If additionally $\langle v_i, v_i\rangle = 1$ for all $i \in \{1, \ldots, n\}$, the set of vectors $(v_1, \ldots, v_n)$ is called **orthonormal**.

Corollary 3.6. *Let V be an inner product space, and let $v_1, \ldots, v_n \in V$ be an orthonormal set of vectors. Then*

$$\left\|\sum_{i=1}^{n} \alpha_i v_i\right\|^2 = \sum_{i=1}^{n} |\alpha_i|^2 .$$

Corollary 3.7. *Any set of orthonormal vectors is linearly independent.*

3.1. **Orthonormal Bases.**

Definition 3.8 (**Orthonormal Basis**). Let V be an inner product space. An **orthonormal basis** of V is a collection $(v_1, \ldots, v_n)$ of orthonormal vectors that is also a basis for V.

Corollary 3.9. *Let V be an n-dimensional inner product space. Let $(v_1, \ldots, v_n)$ be an orthonormal set in V. Then $(v_1, \ldots, v_n)$ is an orthonormal basis of V.*

Proof. By Corollary 3.7, $(v_1, \ldots, v_n)$ is linearly independent. If we have n linearly independent vectors in an n-dimensional space, then these vectors form a basis of V. □

Theorem 3.10. *Let V be an inner product space. Let $(v_1, \ldots, v_n)$ be an orthonormal basis of V. Then, for any $v \in V$, we have*

$$v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i.$$

Proof. Let $v \in V$. Since $(v_1, \ldots, v_n)$ is a basis of V, there exist $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$ such that

$$v = \sum_{i=1}^{n} \alpha_i v_i. \qquad (*)$$

So, we need to show that $\alpha_i = \langle v, v_i \rangle$ for all $i \in \{1, \ldots, n\}$. Let $j \in \{1, \ldots, n\}$. Taking the inner product of both sides of $(*)$ with v_j, and then applying orthonormality,

$$\langle v, v_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle = \alpha_i \langle v_j, v_j \rangle = \alpha_j.$$

□

Corollary 3.11. *Let V be an inner product space. Let $\beta = (v_1, \ldots, v_n)$ be an orthonormal basis of V. Then, the coordinate vector $[v]^\beta$ has the form*

$$[v]^\beta = \begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}.$$

Remark 3.12. Let V, W be finite-dimensional inner product spaces. Let $\beta = (v_1, \ldots, v_n)$ be an orthonormal basis of V and let γ be an orthonormal basis of W. Let $T \colon V \to W$ be a linear transformation. Then we can compute $[T]_\beta^\gamma$ using inner products, since its columns are $[T(v_1)]^\gamma, \ldots, [T(v_n)]^\gamma$.

Example 3.13. Consider $C([0,1], \mathbf{C})$. As usual, let $i := \sqrt{-1}$. Let $f, g \in C([0,1], \mathbf{C})$, and consider the standard inner product

$$\langle f, g \rangle := \int_0^1 f(t) \overline{g(t)} dt.$$

Let $k \in \mathbf{Z}$, $t \in [0,1]$. Define $v_k(t) := e^{2\pi i k t}$. We claim that the set $\{v_k\}_{k \in \mathbf{Z}}$ is an orthonormal set. (In a suitable sense, it is also an orthonormal basis, but we cannot cover this topic here; for more, look into Fourier analysis.) Let $j, k \in \mathbf{Z}$ and observe

$$\langle v_j, v_k \rangle = \int_0^1 v_j(t) \overline{v_k(t)} dt = \int_0^1 e^{2\pi i j t} e^{-2\pi i k t} dt = \int_0^1 e^{2\pi i (j-k) t} dt$$

So, if $j = k$, we get $\langle v_j, v_k \rangle = \int_0^1 dt = 1$. And if $j \neq k$, we have $j - k \in \mathbf{Z}$, so

$$\langle v_j, v_k \rangle = \frac{1}{2\pi i(j-k)}(e^{2\pi i(j-k)} - 1) = \frac{1}{2\pi i(j-k)}(1-1) = 0.$$

Let T_n denote the set of trigonometric polynomials of the form $a_0 + a_1 e^{2\pi i t} + \cdots + a_n e^{2\pi i n t}$, $a_1, \ldots, a_n \in \mathbf{C}$. Then T_n is a subspace of $C([0,1], \mathbf{C})$, so T_n is also an inner product space. So from Theorem 3.10, if $f \in T_n$, we have

$$f(t) = \sum_{j=0}^{n} (\int_0^1 f(s) e^{-2\pi i j s} ds) e^{2\pi i j t},$$

and from Corollary 3.6, we have **Plancherel's formula**

$$\int_0^1 |f(t)|^2 \, dt = \sum_{j=0}^{n} \left| \int_0^1 f(s) e^{-2\pi i j s} ds \right|^2.$$

The scalars $\int_0^1 f(s) e^{-2\pi i j s} ds$, $j \in \{0, \ldots, n\}$ are called the **Fourier coefficients** of f.

4. Gram-Schmidt Orthogonalization

Definition 4.1 (Unit Vector). Let V be a normed linear space, and let $v \in V$. If $\|v\| = 1$, we say that v is a **unit vector**.

Remark 4.2. Let $v \neq 0$. Then $v / \|v\|$ is a unit vector.

Definition 4.3 (Projection onto a vector). Let v, w be vectors in an inner product space, with $w \neq 0$. Define the **orthogonal projection** of v onto w by

$$\mathrm{P}_w(v) := \frac{\langle v, w \rangle}{\langle w, w \rangle} w = \left\langle v, \frac{w}{\|w\|} \right\rangle \frac{w}{\|w\|}.$$

Note that P_w is a linear transformation.

As we saw in the proof of the Cauchy-Schwarz inequality, if $v, w \in V$ and $w \neq 0$, we can write

$$v = (v - P_w(v)) + P_w(v).$$

And $v - P_w(v)$ is orthogonal to w, while $P_w(v)$ is parallel to w. Therefore, $v - P_w(v)$ is orthogonal to $P_w(v)$.

Definition 4.4 (Projection onto a subspace). Let V be an inner product space. Let $W \subseteq V$ be an n-dimensional subspace of V. Let $w_1, \ldots, w_n$ be an orthogonal set of nonzero vectors in W. Let $v \in V$. Define the **orthogonal projection** of v onto W by

$$P_W(v) := \sum_{i=1}^{n} \left\langle v, \frac{w_i}{\|w_i\|} \right\rangle \frac{w_i}{\|w_i\|}.$$

Note that $P_W \colon V \to V$ is a linear transformation, and $R(P_W) \subseteq W$.

Remark 4.5. $P_W(v) = v$ if and only if $v \in W$ by Theorem 3.10. Also, the definition of $P_W(v)$ does not depend on the orthogonal set of nonzero vectors $w_1, \ldots, w_n$. This follows by applying Theorem 3.10 to the orthonormal set $(w_1 / \|w_1\|, \ldots, w_n / \|w_n\|)$.

Remark 4.6. Let $w_1, \ldots, w_n$ be an orthogonal set of nonzero vectors in W. As before, given $v \in V$ and W an n-dimensional subspace of V, we can write

$$v = (v - P_W(v)) + P_W(v).$$

Note that $P_W(v) \in W$, and $(v - P_W(v))$ is orthogonal to w_i for each $i \in \{1, \ldots, n\}$. So, by Lemma 3.1, $(v - P_W(v))$ is orthogonal to any vector in W.

Given a set of linearly independent vectors, we can create an orthonormal set of vectors from the linearly independent set by using projections and Remark 4.6. The procedure for creating these orthonormal sets is known as Gram-Schmidt orthogonalization.

Theorem 4.7 (Gram-Schmidt Orthogonalization). *Let $v_1, \ldots, v_n$ be a linearly independent set of vectors in an inner product space V. Then we can create an orthogonal set of vectors in V as follows. Define*

$$w_1 := v_1.$$

$$w_2 := v_2 - P_{w_1}(v_2).$$

$$w_3 := v_3 - P_{\mathrm{span}(w_1, w_2)}(v_3).$$

And so on. In general, for $k \in \{2, \ldots, n\}$, define

$$w_k := v_k - P_{\mathrm{span}(w_1, \ldots, w_{k-1})}(v_k).$$

Then for each $k \in \{1, \ldots, n\}$, $(w_1, \ldots, w_k)$ is an orthogonal set of nonzero vectors in V. Also, $\mathrm{span}(w_1, \ldots, w_k) = \mathrm{span}(v_1, \ldots, v_k)$ for each $k \in \{1, \ldots, n\}$. Finally, note that the set $(w_1/\|w_1\|, \ldots, w_n/\|w_n\|)$ is an orthonormal set of vectors in V with the same span as $v_1, \ldots, v_n$.

Proof. Note that $w_2 \perp w_1$ from Remark 4.6. We will show that $\{w_1, \ldots, w_k\}$ is an orthogonal set of nonzero vectors, and $\mathrm{span}(w_1, \ldots, w_k) = \mathrm{span}(v_1, \ldots, v_k)$ by induction on k. (The base case $k = 1$ holds since $v_1 \neq 0$.) Assume $\{w_1, \ldots, w_k\}$ is an orthogonal set of nonzero vectors, and $\mathrm{span}(w_1, \ldots, w_k) = \mathrm{span}(v_1, \ldots, v_k)$ for some k. Consider w_{k+1}. Using the definition of w_{k+1}, the inductive hypothesis, and Remark 4.5,

$$w_{k+1} = v_{k+1} - P_{\mathrm{span}(w_1, \ldots, w_k)}(v_{k+1}) = v_{k+1} - P_{\mathrm{span}(v_1, \ldots, v_k)}(v_{k+1}). \qquad (*)$$

From Remark 4.6, w_{k+1} is orthogonal to any vector in $\mathrm{span}(v_1, \ldots, v_k) = \mathrm{span}(w_1, \ldots, w_k)$. Also, $w_{k+1} \neq 0$, since $v_{k+1} \notin \mathrm{span}(v_1, \ldots, v_k)$, by linear independence. That is, $v_{k+1} \neq P_{\mathrm{span}(v_1, \ldots, v_k)}(v_{k+1})$ by Remark 4.5. Therefore, $\{w_1, \ldots, w_{k+1}\}$ is an orthogonal set of nonzero vectors. We now show the spanning property. From $(*)$ and the definition of the projection, $w_{k+1} \in \mathrm{span}(v_1, \ldots, v_{k+1})$. So,

$$\mathrm{span}(w_1, \ldots, w_{k+1}) \subseteq \mathrm{span}(v_1, \ldots, v_{k+1}).$$

Now, note that the span on the right is $(k+1)$-dimensional by Corollary 3.7, as is the span on the left. So we must have equality. The induction step is complete, and we are done. □

Example 4.8. Consider $P_2([-1, 1])$, the set of real polynomials of degree at most 2 on the interval $[-1, 1]$. Let $f, g \in P_2([-1, 1])$. We use the inner product

$$\langle f, g \rangle := \int_{-1}^{1} f(t)g(t)dt.$$

Let's start with the standard basis $(1, t, t^2)$ where t is a real variable, and let's create an orthonormal basis from the standard one. Define $v_1 := 1$, $v_2 := t$, and $v_3 := t^2$. Define

$$w_1 := v_1 = 1.$$

Note that $\langle w_1, w_1 \rangle = 2$, so w_1 has norm $\sqrt{2}$. Then, define

$$w_2 := v_2 - \langle v_2, w_1 \rangle w_1/2 = t.$$

We can verify that $\langle t, 1 \rangle = 0$. Note also that $\langle t, t \rangle = \int_{-1}^{1} t^2 dt = 2/3$, so $(\sqrt{3/2})w_2$ has norm 1. So, define

$$\begin{aligned} w_3 &:= v_3 - \langle v_3, w_2 \rangle w_2 / \|w_2\|^2 - \langle v_3, w_1 \rangle w_1 / \|w_1\|^2 \\ &= t^2 - (\int_{-1}^{1} x^3 dx) t (3/2) - (1/2)(\int_{-1}^{1} x^2 dx) \\ &= t^2 - 1/3. \end{aligned}$$

We can verify that $\langle t^2 - 1/3, 1 \rangle = 0$ and $\langle t^2 - 1/3, t \rangle = 0$. Also, $\int_{-1}^{1} (t^2 - 1/3)^2 dt = 8/45$, so $t^2 - 1/3$ has norm $\sqrt{8/45}$.

In conclusion, an orthonormal basis for $P_2([-1, 1])$, is

$$\left\{ \sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}} t, \sqrt{\frac{45}{8}} \left(t^2 - \frac{1}{3} \right) \right\}.$$

Corollary 4.9. *Every finite dimensional inner product space has an orthonormal basis.*

Proof. Recall that every finite-dimensional vector space has a basis, by the definition of finite-dimensionality. Given this basis $v_1, \ldots, v_n$, apply the Gram-Schmidt Orthogonalization (Theorem 4.7) to get an orthonormal set $w_1 / \|w_1\|, \ldots, w_n / \|w_n\|$. By Corollary 3.9, the vectors produced from the Gram-Schmidt process are an orthonormal basis. □

Corollary 4.10. *Let V be an inner product space, and let $W \subseteq V$ be a finite-dimensional subspace. Then there exists a linear transformation $P \colon V \to V$ such that $P^2 = P$, $R(P) \subseteq W$, and $P(w) = w$ for any $w \in W$. That is, P is a projection onto W.*

Proof. From Corollary 4.9, let $w_1, \ldots, w_n$ be an orthonormal basis for W. As in Definition 4.4, define

$$P(v) = P_W(v) := \sum_{i=1}^{n} \langle v, w_i \rangle w_i.$$

□

4.1. **Orthogonal Complements.**

Definition 4.11 (Orthogonal Subspaces). Let V_1, V_2 be two subspaces of an inner product space V. If $v_1 \perp v_2$ for all $v_1 \in V_1, v_2 \in V_2$, we say that V_1 is **orthogonal** to V_2, and we write $V_1 \perp V_2$.

Lemma 4.12. *Let V_1, V_2 be two subspaces of an inner product space V. If $V_1 \perp V_2$, then $V_1 \cap V_2 = \{0\}$.*

Proof. Since V_1, V_2 are subspaces, $0 \in V_1$ and $0 \in V_2$, so $0 \in V_1 \cap V_2$. Now, let $v \in V_1 \cap V_2$. We will show that $v = 0$. Then $v \in V_2$. But since $v \in V_2$ and $V_2 \perp V_1$, we have $\langle v, v_1 \rangle = 0$ for all $v_1 \in V_1$. In particular, since $v \in V_1$, we have $\langle v, v \rangle = 0$. By the positive definiteness property of the inner product, we conclude that $v = 0$. That is, $V_1 \cap V_2 = \{0\}$. □

Definition 4.13 (**Orthogonal Complement**)**.** Let V_1 be a subspace of an inner product space V. Define the **orthogonal complement** of V_1 in V by

$$V_1^{\perp} := \{v \in V : \langle v, v_1 \rangle = 0, \forall\, v_1 \in V_1\}.$$

Exercise 4.14. Show that $\{0\}^{\perp} = V$ and $V^{\perp} = \{0\}$.

Exercise 4.15. Let V_1 be a subspace of an inner product space V. Show that $V_1^{\perp}$ is a subspace of V.

The following Theorem gives an algorithm for computing orthogonal complements.

Theorem 4.16. *Let V be an n-dimensional inner product space, and let $W \subseteq V$ be a k-dimensional subspace. Let $v_1, \ldots, v_k$ be a basis of W, and let $v_1, \ldots, v_n$ be an extension of that basis to V. (We proved that this extension exists in Chapter 1). Let $w_1, \ldots, w_n$ be the orthonormal vectors produced by Gram-Schmidt orthogonalization. Then $w_1, \ldots, w_k$ is an orthonormal basis of W, and $w_{k+1}, \ldots, w_n$ is an orthonormal basis of $W^{\perp}$.*

Proof. From Theorem 4.7, $\mathrm{span}(w_1, \ldots, w_k) = \mathrm{span}(v_1, \ldots, v_k)$. Since W is k-dimensional, we conclude that $w_1, \ldots, w_k$ is a basis for W. Since $w_1, \ldots, w_k$ is also orthonormal, it is therefore an orthonormal basis of W.

Also, the vectors $w_{k+1}, \ldots, w_n$ are orthonormal, and therefore they are linearly independent by Corollary 3.7. So, it remains to show that $w_{k+1}, \ldots, w_n$ spans $W^{\perp}$. Let $j \in \{k+1, \ldots, n\}$. By the Gram-Schmidt process, w_j is orthogonal to each of the vectors $w_1, \ldots, w_k$. By Lemma 3.1, w_j is then orthogonal to all of W. So, $w_j \in W^{\perp}$. So, $\mathrm{span}(w_{k+1}, \ldots, w_n) \subseteq W^{\perp}$. It remains to show that every $w \in W^{\perp}$ is in $\mathrm{span}(w_{k+1}, \ldots, w_n)$.

Let $w \in W^{\perp}$. Since $w \in V$ and $(w_1, \ldots, w_n)$ is an orthonormal basis of V, we have by Theorem 3.10,

$$w = \sum_{i=1}^{n} \langle w, w_i \rangle w_i.$$

Since $w \in W^{\perp}$, $\langle w, w_i \rangle = 0$ for each $i \in \{1, \ldots, k\}$. That is,

$$w = \sum_{i=k+1}^{n} \langle w, w_i \rangle w_i.$$

So, $w \in \mathrm{span}(w_{k+1}, \ldots, w_n)$, as desired. □

Example 4.17. We continue the definitions and notation from Example 4.8. Consider $W \subseteq P_2([-1, 1])$, where W is the span of 1 and t. Let's compute $W^{\perp}$. To do this, we complete the set $(1, t)$ to a basis $(1, t, t^2)$. From Example 4.8, we then used Gram-Schmidt orthogonalization using $v_1 = 1$, $v_2 = t$ and $v_3 = t^2$. We found that the resulting orthonormal basis for $P_2([-1, 1])$ is

$$\left\{ \sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}t, \sqrt{\frac{45}{8}}\left(t^2 - \frac{1}{3}\right) \right\}.$$

So, $\sqrt{1/2}$ and $\sqrt{(3/2)}t$ are an orthonormal basis of W. And therefore, $W^{\perp}$ is a one-dimensional space described as

$$W^{\perp} = \{\alpha(t^2 - 1/3) \colon \alpha \in \mathbf{R}\}.$$

Corollary 4.18 (Dimension Theorem for orthogonal complements). *Let W be a subspace of a finite-dimensional inner product space V. Then*

$$\dim(W) + \dim(W^{\perp}) = \dim(V).$$

Corollary 4.19. *Let W be a subspace of a finite-dimensional inner product space V. Then every $v \in V$ can be written uniquely as $v = w + n$ where $w \in W$ and $n \in W^{\perp}$.*

Proof. By Theorem 4.16, $\exists$ an orthonormal basis $w_1, \ldots, w_m$ of V such that $w_1, \ldots, w_k$ is an orthonormal basis of W, and $w_{k+1}, \ldots, w_m$ is an orthonormal basis of $W^{\perp}$. Let $v \in V$. by Theorem 3.10,

$$v = \sum_{i=1}^{m} \langle v, w_i \rangle w_i = \sum_{i=1}^{k} \langle v, w_i \rangle w_i + \sum_{i=k+1}^{m} \langle v, w_i \rangle w_i.$$

Define

$$w := \sum_{i=1}^{k} \langle v, w_i \rangle w_i, \qquad n := \sum_{i=k+1}^{m} \langle v, w_i \rangle w_i.$$

Then $v = w + n$, $w \in W$, $n \in W^{\perp}$. (As an aside, note that $w = P_W(v)$, and $n = v - P_W(v)$.) We now show the desired uniqueness statement. Suppose $v = w' + n'$ with $w' \in W$ and $n' \in W^{\perp}$. We will be done once we show that $w = w'$ and $n = n'$. Since $w + n = w' + n'$, we have

$$w - w' = n - n'. \qquad (*)$$

The vector on the left of $(*)$ is in W, and the vector on the right of $(*)$ is in $W^{\perp}$. By Lemma 4.12, $W \cap W^{\perp} = \{0\}$. So, both sides of $(*)$ must be zero. That is, $w = w'$ and $n = n'$, as desired. $\square$

Theorem 4.20 (Orthogonal projections minimize length). *Let W be a subspace of a finite-dimensional inner product space V. Let $v \in V$, and let $w = P_W(v)$ be the orthogonal projection of v onto W. Then, for any $w' \in W$ with $w' \neq w$, we have $\|v - w\| < \|v - w'\|$.*

Proof. From Corollary 4.19, write $v = w + n$, where $w := P_W(v) \in W$, and $n := v - w \in W^{\perp}$. Then $\|v - w\| = \|n\|$. Now, write

$$v - w' = (v - w) + (w - w').$$

Since $w, w' \in W$, $w - w' \in W$. Since $v - w = n \in W^{\perp}$, $\langle v - w, w - w' \rangle = 0$. So, by the Pythagorean Theorem (Theorem 3.2),

$$\|v - w'\|^2 = \|v - w\|^2 + \|w - w'\|^2.$$

So, $\|v - w'\|^2 > \|v - w\|^2$ since $w \neq w'$, as desired. $\square$

5. Adjoints

5.1. Linear Functionals.

Definition 5.1 (Linear Functional). Let V be a vector space over a field $\mathbf{F}$. A **linear functional** is a linear transformation $T\colon V \to \mathbf{F}$.

Linear functionals are also known as dual vectors, covectors, or 1-forms. In order to understand some vector space V, it is often of interest to understand the set of all linear functionals on V. Such a classification becomes quite subtle especially for infinite dimensional spaces. However, as we show below, the case of finite-dimensional inner product spaces is fairly tame.

Example 5.2. Define $T\colon \mathbf{R}^3 \to \mathbf{R}$ by $T(a,b,c) := a+2b+3c$. Then T is a linear functional.

Example 5.3. Define $T\colon C([0,1],\mathbf{R}) \to \mathbf{R}$ by $T(f) := \int_0^1 f(t)dt$. Then T is a linear functional.

Example 5.4. Define $T\colon C([0,1],\mathbf{R}) \to \mathbf{R}$ by $T(f) := f(1/3)$. Then T is a linear functional.

Example 5.5. Let V be a finite-dimensional inner product space over $\mathbf{R}$. Let $w \in V$, and define $T\colon V \to \mathbf{R}$ by $T(v) := \langle v, w\rangle$. Then T is a linear functional.

As we now show, the previous example essentially classifies all linear functionals on a finite-dimensional inner product space.

Theorem 5.6 (Riesz Representation Theorem). *Let V be a finite-dimensional inner product space over a field $\mathbf{F}$. Let $T\colon V \to \mathbf{F}$ be a linear functional. Then there exists a unique vector $w \in V$ such that, for all $v \in V$, $T(v) = \langle v, w\rangle$.*

Proof. From Corollary 4.9, V has an orthonormal basis $v_1, \ldots, v_n$. Define

$$w := \sum_{j=1}^{n} v_j \overline{T(v_j)}.$$

Let $v \in V$. From Theorem 3.10,

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i.$$

Since T is linear, we get

$$T(v) = \sum_{i=1}^{n} \langle v, v_i\rangle T(v_i)$$

Since $v_1, \ldots, v_n$ is an orthonormal basis,

$$\langle v, w\rangle = \left\langle \sum_{i=1}^{n} \langle v, v_i\rangle v_i, \sum_{j=1}^{n} v_j \overline{T(v_j)} \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle v, v_i\rangle T(v_j)\langle v_i, v_j\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle T(v_i) = T(v).$$

This completes the existence part of the proof. We now prove uniqueness.

Suppose there exists $w' \in V$ such that $T(v) = \langle v, w'\rangle$. We will show that $w = w'$. For all $v \in V$, $T(v) = \langle v, w\rangle = \langle v, w'\rangle$. That is,

$$\forall\, v \in V \;\; \langle v, w - w'\rangle = 0. \qquad (*)$$

Choosing $v = w - w'$ shows that $\langle w - w', w - w'\rangle = 0 = \|w - w'\|^2$. Therefore, $w - w' = 0$, as desired. $\square$

5.2. **Adjoints.** Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let V be a finite-dimensional inner product space over $\mathbf{F}$ with inner product $\langle\,,\rangle_V$, and let W be a finite-dimensional inner product space over $\mathbf{F}$ with inner product $\langle\,,\rangle_W$. Let $T\colon V\to W$ be a linear transformation. Given any $w\in W$, define a linear functional $T_w\colon V\to\mathbf{F}$ by

$$T_w(v) := \langle T(v), w\rangle_W.$$

Note that T_w is actually a linear function. To see this, let $v, v'\in V$ and let $\alpha\in\mathbf{F}$. Then

$$T_w(v+v') = \langle T(v+v'), w\rangle_W = \langle T(v), w\rangle_W + \langle T(v'), w\rangle_W = T_w(v) + T_w(v').$$

$$T_w(\alpha v) = \langle T(\alpha v), w\rangle_W = \alpha\langle T(v), w\rangle_W = \alpha T_w(v).$$

Definition 5.7 (**Adjoint**)**.** Since $T_w\colon V\to\mathbf{F}$ is a linear functional, we can apply Theorem 5.6 to get a unique vector in V, which we denote by $T^*(w)$, such that for all $v\in V$,

$$T_w(v) = \langle v, T^*(w)\rangle_V.$$

As we will see shortly, T^* is a linear transformation from W to V, which we call the **adjoint** of T. Also, recalling the definition of T_w, we have

$$\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V.$$

Remark 5.8. Note that $T\colon V\to W$, whereas $T^*\colon W\to V$.

Remark 5.9. We have added subscripts to the above inner products to emphasize that the inner product in W could be different from the inner product in V. From now on, we will drop these subscripts.

Lemma 5.10. *Let $T\colon V\to W$ be a linear transformation between finite-dimensional inner product spaces. Then $T^*\colon W\to V$ is a linear transformation*

Proof. Let $w, w'\in W$ and let $\alpha\in\mathbf{F}$. We will first show that $T^*(w+w') = T^*(w)+T^*(w')$. By the definition of T^*, for all $v\in V$,

$$\langle T(v), w+w'\rangle = \langle v, T^*(w+w')\rangle.$$

So, rearranging things and applying the definition of T^* again,

$$\langle v, T^*(w+w')\rangle = \langle T(v), w\rangle + \langle T(v), w'\rangle = \langle v, T^*(w)\rangle + \langle v, T^*(w')\rangle = \langle v, T^*(w)+T^*(w')\rangle.$$

By the uniqueness part of the Riesz Representation Theorem (Theorem 5.6), we therefore have $T^*(w+w') = T^*(w)+T^*(w')$.

We now show that $T^*(\alpha w) = \alpha T^*(w)$.

$$\langle v, T^*(\alpha w)\rangle = \langle T(v), \alpha w\rangle = \overline{\alpha}\langle T(v), w\rangle = \overline{\alpha}\langle v, T^*(w)\rangle = \langle v, \alpha T^*(w)\rangle$$

By the uniqueness part of the Riesz Representation Theorem (Theorem 5.6), we therefore have $T^*(\alpha w) = \alpha T^*(w)$. □

Definition 5.11 (**Adjoint of a Matrix**)**.** Let A be an $m\times n$ matrix with $A_{jk}\in\mathbf{C}$, $1\leq j\leq m$, $1\leq k\leq n$. The **adjoint** of A, denoted by $A^\dagger$, is an $n\times m$ matrix with entries $(A^\dagger)_{jk} := \overline{A_{kj}}$, $1\leq j\leq n$, $1\leq k\leq m$.

Theorem 5.12. *Let $T\colon V\to W$ be a linear transformation between inner product spaces V, W. Let $\beta = (v_1,\ldots,v_n)$ be an orthonormal basis of V and let $\gamma = (w_1,\ldots,w_m)$ be an orthonormal basis of W. Then*

$$[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^\dagger.$$

Proof. Let $w \in W$. From Corollary 3.11, recall that

$$[w]^{\gamma} = \begin{pmatrix} \langle w, w_1 \rangle \\ \vdots \\ \langle w, w_m \rangle \end{pmatrix}.$$

Also, from Remark 3.12, $[T]_{\beta}^{\gamma}$ has columns $[T(v_1)]^{\gamma}, \ldots, [T(v_n)]^{\gamma}$. That is,

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} \langle T(v_1), w_1 \rangle & \langle T(v_2), w_1 \rangle & \cdots & \langle T(v_n), w_1 \rangle \\ \langle T(v_1), w_2 \rangle & \langle T(v_2), w_2 \rangle & \cdots & \langle T(v_n), w_2 \rangle \\ \vdots & \vdots & \cdots & \vdots \\ \langle T(v_1), w_m \rangle & \langle T(v_2), w_m \rangle & \cdots & \langle T(v_n), w_m \rangle \end{pmatrix}.$$

Similarly, $[T^*]_{\gamma}^{\beta}$ is an $n \times m$ matrix with (i, j) entry $\langle T^*(w_j), v_i \rangle$. And

$$\langle T^*(w_j), v_i \rangle = \overline{\langle v_i, T^*(w_j) \rangle} = \overline{\langle T(v_i), w_j \rangle}.$$

That is,

$$[T^*]_{\gamma}^{\beta} = \begin{pmatrix} \overline{\langle T(v_1), w_1 \rangle} & \overline{\langle T(v_1), w_2 \rangle} & \cdots & \overline{\langle T(v_1), w_m \rangle} \\ \overline{\langle T(v_2), w_1 \rangle} & \overline{\langle T(v_2), w_2 \rangle} & \cdots & \overline{\langle T(v_2), w_m \rangle} \\ \vdots & \vdots & \cdots & \vdots \\ \overline{\langle T(v_n), w_1 \rangle} & \overline{\langle T(v_n), w_2 \rangle} & \cdots & \overline{\langle T(v_n), w_m \rangle} \end{pmatrix}.$$

In conclusion $[T^*]_{\gamma}^{\beta} = ([T]_{\beta}^{\gamma})^{\dagger}$. □

Corollary 5.13. *Let* $\mathbf{F}$ *denote* $\mathbf{R}$ *or* $\mathbf{C}$*. Let* A *be an* $m \times n$ *matrix with elements in* $\mathbf{F}$*. Let* $\mathbf{F}^n$ *and* $\mathbf{F}^m$ *respectively denote the usual vector spaces* $\mathbf{F}^n$ *and* $\mathbf{F}^m$ *with their standard inner products. Then the adjoint of* $L_A \colon \mathbf{F}^n \to \mathbf{F}^m$ *is* $L_{A^{\dagger}}$*.*

Proof. Note that $L_A \colon \mathbf{F}^m \to \mathbf{F}^n$ and $L_{A^{\dagger}} \colon \mathbf{F}^n \to \mathbf{F}^m$. Let β be the standard basis of $\mathbf{F}^m$ and let γ be the standard basis of $\mathbf{F}^n$. From Theorem 5.12,

$$[L_A^*]_{\gamma}^{\beta} = ([L_A]_{\beta}^{\gamma})^{\dagger} = A^{\dagger} = [L_{A^{\dagger}}]_{\gamma}^{\beta}.$$

So, $L_A^* = L_{A^{\dagger}}$, as desired. □

Remark 5.14. Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let A be an $m \times n$ matrix with elements in $\mathbf{F}$. Then for any $v \in \mathbf{F}^n$ and for any $w \in \mathbf{F}^m$,

$$\langle Av, w \rangle = \langle v, A^{\dagger} w \rangle.$$

In this equality, the inner product on the left is the standard inner product on $\mathbf{F}^m$, and the inner product on the right is the standard inner product on $\mathbf{F}^n$. Note that if we change the inner product, then the adjoint could possibly change as well. For example, suppose $n = 2$ and we use the inner product

$$\langle (v_1, v_2), (w_1, w_2) \rangle' := (v_1, v_2) \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} (w_1, w_2)^t, \qquad (v_1, v_2), (w_1, w_2) \in \mathbf{R}^2.$$

Then it is not true that $\langle Av, w \rangle' = \langle v, A^{\dagger} w \rangle'$. For example, choose $v = (1, 0)$, $w = (1, 0)$, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $\langle Av, w \rangle' = \langle (1, 0), (1, 0) \rangle' = (1, 0)(1, 1/2)^t = 1$, while $\langle v, A^{\dagger} w \rangle' = \langle (1, 0), (1, 1) \rangle' = (1, 0)(3/2, 3/2)^t = 3/2$.

Exercise 5.15. Find the adjoint of $L_A\colon \mathbf{R}^2 \to \mathbf{R}^2$ in the above example, where $\mathbf{R}^2$ is equipped with the inner product $\langle\, ,\rangle'$.

Exercise 5.16. Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let $T\colon V \to W$, $S\colon V \to W$ and let $R\colon U \to V$ be linear transformations between inner product spaces U, V, W over $\mathbf{F}$. Verify the following facts

(a) $(T + S)^* = T^* + S^*$.
(b) For all $\alpha \in \mathbf{F}$, $(\alpha T)^* = \overline{\alpha} T^*$.
(c) $(T^*)^* = T$.
(d) $(TR)^* = R^* T^*$.
(e) If T is invertible, then $(T^{-1})^* = (T^*)^{-1}$.

Exercise 5.17. Let A be an $m \times n$ matrix. Show that $\mathrm{rank}(A) = \mathrm{rank}(A^\dagger)$.

Exercise 5.18. Let A be an $n \times n$ matrix with elements in $\mathbf{C}$. Then $\det(A^\dagger) = \overline{\det(A)}$.

6. Normal Operators

One of the ultimate goals of this course is to take an arbitrary linear transformation and either diagonalize it, or show that it cannot be diagonalized. Such a result provides the starting point for many further investigations. We cannot fully realize this goal in this course. However, we will identity a large class of linear transformations (i.e. operators) that can be diagonalized, and that appear often in practice. Two such classes of operators are normal and self-adjoint operators. This course will conclude by showing that these two classes of operators can be diagonalized. Such a diagonalization result is referred to as a **spectral theorem**. The **spectrum** of a linear operator is its set of eigenvalues. This terminology may seem a bit strange, since we usually refer to the spectrum of an electromagnetic wave. However, this conflation of terminology is no coincidence. For example, the spectral theorem for infinite-dimensional vector spaces (which is outside the scope of this course) demonstrates mathematically the discreteness of the energy emissions of the hydrogen atom. In particular, there is a self-adjoint operator whose set of eigenvalues is the energy emission spectrum of the hydrogen atom. And the eigenvectors give the (infinite set of) atomic orbitals that you learned in chemistry class, the first of which you called s,p,d and f orbitals. (Beware: in infinite-dimensional spaces, self-adjointness becomes more complicated than in the finite-dimensional case.)

Definition 6.1 (**Normal Operator**)**.** Let V be a finite-dimensional inner product space. Let $T\colon V \to V$ be a linear transformation. Recall that $T^*\colon V \to V$ is also a linear transformation. We say that T is a **normal operator** if $TT^* = T^*T$.

Example 6.2. Define $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ by $T(x, y) = (y, -x)$ for all $x, y \in \mathbf{R}$. Then $T^*(x, y) = (-y, x)$, by the definition of T^*. Observe,

$$TT^*(x, y) = T(-y, x) = (x, y).$$

$$T^*T(x, y) = T^*(y, -x) = (x, y).$$

So, $T^*T = TT^*$, so T is normal.

Definition 6.3 (**Normal Matrix**)**.** Let A be an $n \times n$ matrix. We say that A is **normal** if $AA^\dagger = A^\dagger A$.

Example 6.4. Every diagonal matrix is normal.

Proposition 6.5. *Let $T\colon V \to V$ be a linear transformation on a finite-dimensional inner product space V. Let β be an orthonormal basis of V. Then $T\colon V \to V$ is normal if and only if $[T]_\beta^\beta$ is normal.*

Proof. Suppose T is normal. Then $TT^* = T^*T$. Taking the matrix representation of this identity,

$$[T]_\beta^\beta [T^*]_\beta^\beta = [TT^*]_\beta^\beta = [T^*T]_\beta^\beta = [T^*]_\beta^\beta [T]_\beta^\beta.$$

From Theorem 5.12, $[T^*]_\beta^\beta$ is the adjoint of $[T]_\beta^\beta$. So, $[T]_\beta^\beta$ is normal.

So, we proved the forward implication. To prove the reverse implication, note that the above steps can be reversed. □

Lemma 6.6. *Let* **F** *denote* **R** *or* **C**. *Let V be a finite-dimensional inner product space over* **F**. *Let $T\colon V \to V$ be a normal operator. Assume that there exists $v \in V$ and $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$. Then $T^* v = \overline{\lambda} v$.*

Proof. It suffices to show that $\|T^*(v) - \overline{\lambda} v\|^2 = 0$. That is, we verify that

$$\langle T^*(v) - \overline{\lambda} v, T^*(v) - \overline{\lambda} v\rangle = 0.$$

Expanding both sides, we equivalently want

$$\langle T^*(v), T^*(v)\rangle - \overline{\lambda}\langle v, T^*(v)\rangle - \lambda\langle T^*(v), v\rangle + |\lambda|^2 \langle v, v\rangle = 0.$$

Using the definition of adjoint and that T is normal, the left-most term is $\langle TT^* v, v\rangle = \langle T^*Tv, v\rangle$. Using the adjoint definition some more, we equivalently want

$$\langle T^*T(v), v\rangle - \overline{\lambda}\langle T(v), v\rangle - \lambda\langle v, T(v)\rangle + |\lambda|^2 \langle v, v\rangle = 0.$$

Using $T(v) = \lambda v$, we equivalently want

$$\lambda\langle T^*(v), v\rangle - |\lambda|^2 \langle v, v\rangle - |\lambda|^2 \langle v, v\rangle + |\lambda|^2 \langle v, v\rangle = 0.$$

Applying the adjoint definition again and simplifying, we want

$$\lambda\langle v, T(v)\rangle - |\lambda|^2 \langle v, v\rangle = 0.$$

Finally, using $T(v) = \lambda v$, we have $\lambda\langle v, T(v)\rangle = |\lambda|^2 \langle v, v\rangle$, completing the proof. □

The following Lemma shows that Proposition 1.1 becomes strengthened when T is normal.

Lemma 6.7. *Let $T\colon V \to V$ be a normal operator on a finite-dimensional inner product space V. Let v_1, v_2 be two eigenvectors of T with distinct eigenvalues λ_1, λ_2, respectively. Then v_1 is orthogonal to v_2.*

Proof. Since $T(v_1) = \lambda_1 v_1$ and $T(v_2) = \lambda_2 v_2$, we have $T^*(v_1) = \overline{\lambda_1} v_1$ and $T^*(v_2) = \overline{\lambda_2} v_2$ by Lemma 6.6. So,

$$\lambda_1 \langle v_1, v_2\rangle = \langle T(v_1), v_2\rangle = \langle v_1, T^*(v_2)\rangle = \lambda_2 \langle v_1, v_2\rangle.$$

Since $\lambda_1 \neq \lambda_2$, we must have $\langle v_1, v_2\rangle = 0$, as desired. □

Remark 6.8. Most linear transformations will not be normal, since they will typically have non-orthogonal eigenvectors.

We now prove a converse to Lemma 6.7, which is also a variation on Lemma 1.2 for normal T.

Lemma 6.9. *Let $T: V \to V$ be a linear transformation on a finite-dimensional inner product space V. Suppose β is an orthonormal basis of V consisting of eigenvectors of T. Then T is normal.*

Proof. From Lemma 1.2, $[T]_\beta^\beta$ is diagonal. In particular, $[T]_\beta^\beta$ is normal. So, from Proposition 6.5, T is normal. □

Theorem 6.10 (**The Spectral Theorem for Normal Operators**). *Let $T: V \to V$ be a normal operator on a finite-dimensional inner product space V over $\mathbf{C}$. Then there exists an orthonormal basis β of V consisting of eigenvectors of T. In particular, T is diagonalizable.*

Proof. We will prove the theorem by induction on the dimension n of V. Consider first the case $n = 1$. Let β consist of exactly one nonzero unit vector $v \in V$. Since V is one dimensional, for any $w \in V$, there exists $\alpha \in \mathbf{C}$ such that $w = \alpha v$. So, if $T(v) = w$ for some $w \in V$, we have $T(v) = \alpha v$, so that v is an eigenvector of T. In conclusion, the theorem holds for $n = 1$.

Now, suppose the theorem holds for a fixed $n \geq 1$, and consider the case $\dim(V) = n + 1$. Let $f(\lambda)$ be the characteristic polynomial of some matrix representation of T. (Recall that any two matrix representations of T are similar by Lemma 1.3, and two similar matrices have the same characteristic polynomial by Lemma 1.5. So, the matrix representation that we use for T does not affect f.) From the Fundamental Theorem of Algebra (Theorem 1.4), f has $n + 1$ zeros. In particular, f has one zero. So, T has at least one eigenvalue $\lambda_1 \in \mathbf{C}$, and at least one eigenvector $v_1 \in V$, $v_1 \neq 0$ with $T(v_1) = \lambda_1 v_1$. Replacing v_1 with $v_1 / \|v_1\|$ if necessary, we may assume that $\|v_1\| = 1$.

Since $T(v_1) = \lambda_1 v_1$, Lemma 6.6 shows that $T^*(v_1) = \overline{\lambda_1} v_1$. Let $W := \{a v_1 : a \in \mathbf{C}\}$ denote the span of v_1. Observe that $W \subseteq V$ is a one-dimensional subspace. Let $W^\perp := \{v \in V : \langle v, v_1 \rangle = 0\}$ denote the orthogonal complement of W. Recall that $W^\perp$ is a subspace of V by Exercise 4.15, and $\dim(W^\perp) = n + 1 - 1 = n$ by Corollary 4.18.

We would like to apply the inductive hypothesis to T, where we restrict the domain of T to the subspace $W^\perp$. In order for the inductive hypothesis to apply, we need to show that the restriction of T to $W^\perp$ satisfies the hypotheses of the theorem. That is, we need to show:

(a) $T(W^\perp) \subseteq W^\perp$, i.e. that $W^\perp$ is invariant under T.
(b) $T^*(W^\perp) \subseteq W^\perp$.
(c) T and T^* are adjoints of each other when we consider them as operators on $W^\perp$.

Proof of (a). Let $w \in W^\perp$, so that $\langle w, v_1 \rangle = 0$. Then

$$0 = \lambda_1 \langle w, v_1 \rangle = \langle w, T^*(v_1) \rangle = \langle T(w), v_1 \rangle.$$

So, $T(w) \in W^\perp$, as desired.

Proof of (b). Let $w \in W^\perp$, so that $\langle w, v_1 \rangle = 0$. Then

$$0 = \overline{\lambda_1} \langle w, v_1 \rangle = \langle w, T(v_1) \rangle = \langle T^*(w), v_1 \rangle.$$

So, $T^*(w) \in W^\perp$, as desired.

Proof of (c). Let $v, w \in W^\perp$. We need to show that there exists $x \in W^\perp$ such that

$$\langle T(v), w \rangle = \langle v, x \rangle. \qquad (*)$$

Since T^* is the adjoint of T, we know that $x := T^*(w)$ is the unique vector in V such that $(*)$ holds, by the Riesz Representation Theorem (Theorem 5.6). So, we need to show that $T^*(w) \in W^\perp$. But this follows from part (b).

Having proven parts (a),(b) and (c), we can finally apply the inductive hypothesis to T, where we restrict the domain of T to $W^\perp$. That is, there exists an orthonormal basis $(v_2, \ldots, v_{n+1})$ of $W^\perp$ consisting of eigenvectors of T. Since $v_1 \in W$, v_1 is orthogonal to the vectors $v_2, \ldots, v_{n+1}$. So, the set of vectors $v_1, \ldots, v_{n+1}$ is an orthonormal set (recalling $\|v_1\| = 1$). Since $\dim(V) = n+1$, Corollary 3.9 says $v_1, \ldots, v_{n+1}$ is a basis of V, as desired. □

7. Self-Adjoint Operators

Definition 7.1 (**Self-Adjoint Operator**)**.** Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let V be a finite-dimensional inner product space over $\mathbf{F}$. Let $T: V \to V$ be a linear transformation. Then T is called a **self-adjoint operator** if $T^* = T$. A square matrix A is said to be **self-adjoint** if $A = A^\dagger$.

Remark 7.2. Let $T: V \to V$ be a linear transformation on a finite-dimensional inner product space V. If T is self-adjoint, then T is normal. But if T is normal, then T is not necessarily self-adjoint.

Example 7.3. The linear transformation $T: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(x, y) = (y, -x)$ is normal but not self-adjoint, since it has adjoint $T^*(x, y) = (-y, x)$. However, the linear transformation $T: \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(x, y) = (y, x)$ is self-adjoint.

Remark 7.4. Let $T: V \to V$ be a linear transformation on a finite-dimensional inner product space V over $\mathbf{F}$. If T is self-adjoint and if $\mathbf{F} = \mathbf{C}$, then T is sometimes called **Hermitian**. If T is self-adjoint and if $\mathbf{F} = \mathbf{R}$, then T is **symmetric**. A square complex matrix A with $A = A^\dagger$ is also called **Hermitian**. And a square real matrix with $A = A^\dagger$ is called **symmetric**, since $A = A^\dagger$ becomes $A = A^t$.

Theorem 7.5. *Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let V be a finite-dimensional inner product space over $\mathbf{F}$. Let $T: V \to V$ be a self-adjoint linear transformation. Then all eigenvalues of T are real.*

Proof. Let $\lambda \in \mathbf{C}$ be any eigenvalue of T. So, there exists $v \in V$ with $v \neq 0$ such that $T(v) = \lambda v$. Lemma 6.6 shows that $T^*(v) = \overline{\lambda} v$. Since $T = T^*$, we conclude that $\lambda = \overline{\lambda}$, so that $\lambda \in \mathbf{R}$, as desired. □

Remark 7.6. Similarly, all eigenvalues of a Hermitian matrix are real.

Proposition 7.7. *Let $T: V \to V$ be a linear transformation on a finite-dimensional inner product space V. Let β be an orthonormal basis of V. Then $T: V \to V$ is self-adjoint if and only if $[T]_\beta^\beta$ is self-adjoint.*

Proof. Suppose T is self-adjoint. Then $T = T^*$. From Theorem 5.12, $[T^*]_\beta^\beta$ is the adjoint of $[T]_\beta^\beta$. So, $[T]_\beta^\beta$ is self-adjoint. We proved the forward implication. To prove the reverse implication, note that the above steps can be reversed. □

Corollary 7.8. *Let A be an $n \times n$ complex Hermitian matrix, and let $f(\lambda) := \det(A - \lambda I)$ be the characteristic polynomial of A. Then there exist $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$ such that $f(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda)$.*

Proof. From the Fundamental Theorem of Algebra (Theorem 1.4), there exist $\lambda_0, \ldots, \lambda_n \in \mathbf{C}$ such that $f(\lambda) = \lambda_0 \prod_{i=1}^n (\lambda_i - \lambda)$. Recall that coefficient of the degree n term of $f(\lambda)$ is $(-1)^n$ by Theorem 1.6. So, $\lambda_0 = 1$. From Remark 7.6, $\lambda_i \in \mathbf{R}$ for all $i \in \{1, \ldots, n\}$. □

Lemma 7.9. *Let* $\mathbf{F}$ *denote either* $\mathbf{R}$ *or* $\mathbf{C}$*. Let* $T\colon V \to V$ *be a linear transformation on a finite-dimensional inner product space* V *over* $\mathbf{F}$*. Suppose there exists an orthonormal basis* β *of* V *consisting of eigenvectors of* T *with real eigenvalues. Then* T *is self-adjoint.*

Proof. From Lemma 1.2, $[T]_\beta^\beta$ is diagonal with real entries. In particular, $[T]_\beta^\beta$ is self-adjoint So, from Proposition 7.7, T is self-adjoint. □

Theorem 7.10 (The Spectral Theorem for Self-Adjoint Operators). *Let* $\mathbf{F}$ *denote either* $\mathbf{R}$ *or* $\mathbf{C}$*. Let* $T\colon V \to V$ *be a self-adjoint operator on a finite-dimensional inner product space* V *over* $\mathbf{F}$*. Then there exists an orthonormal basis* β *of* V *consisting of eigenvectors of* T*. In particular,* T *is diagonalizable. Moreover, all eigenvalues of* T *are real.*

Proof. Since T is self-adjoint, T is normal. So, if $\mathbf{F} = \mathbf{C}$, the result follows directly from the Spectral Theorem for normal operators (Theorem 6.10). Then we apply Theorem 7.5 to finish. If $\mathbf{F} = \mathbf{R}$, we repeat the proof of Theorem 6.10, replacing $\mathbf{C}$ everywhere by $\mathbf{R}$. The crucial new ingredient in the proof is that, in the inductive step and in the base case of the induction, T has some real eigenvalue $\lambda_1 \in \mathbf{R}$ by Theorem 7.5. □

Remark 7.11. Note that Theorem 6.10 requires V to be a vector space over $\mathbf{C}$. But Theorem 7.10 requires V to be a vector space over $\mathbf{R}$ or $\mathbf{C}$. So, every symmetric operator on a real inner product space is diagonalizable.

Remark 7.12. In conclusion, self-adjoint operators are really nice, since they have an orthonormal basis of eigenvectors (so they can be diagonalized), and all of their eigenvalues are real.

8. Orthogonal and Unitary Operators (Bonus Section)

Definition 8.1 (Unitary Operators). Let $\mathbf{F}$ denote $\mathbf{R}$ or $\mathbf{C}$. Let V be a finite-dimensional inner product space over $\mathbf{F}$. Let $T\colon V \to V$ be a linear transformation. Then T is called a **unitary operator** if $TT^* = T^*T = I_V$. A square matrix A is called **unitary** if $AA^\dagger = A^\dagger A = I$.

Remark 8.2. Let $T\colon V \to V$ be a linear transformation on a finite-dimensional inner product space V. If T is unitary, then T is normal.

Remark 8.3. Let $T\colon V \to V$ be a linear transformation on a finite-dimensional inner product space V over $\mathbf{F}$. If T is unitary and $\mathbf{F} = \mathbf{R}$, then T is called **orthogonal**. A square real matrix A with $AA^\dagger = A^\dagger A = I$ is also called **orthogonal**, since we then have $AA^t = A^t A = I$.

Theorem 8.4. *Let* $\mathbf{F}$ *denote* $\mathbf{R}$ *or* $\mathbf{C}$*. Let* V *be a finite-dimensional inner product space over* $\mathbf{F}$*. Let* $T\colon V \to V$ *be a unitary operator. Then all eigenvalues of* T *have absolute value 1.*

Proof. Let $\lambda \in \mathbf{C}$ be any eigenvalue of T. So, there exists $v \in V$ with $v \neq 0$ such that $T(v) = \lambda v$. Lemma 6.6 shows that $T^*(v) = \overline{\lambda} v$. So, using $T^*T = I_V$,

$$|\lambda|^2 \langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle T(v), T(v)\rangle = \langle T^*T(v), v\rangle = \langle v, v\rangle.$$

Since $v \neq 0$, we conclude that $|\lambda|^2 = 1$, as desired. □

Remark 8.5. Similarly, all eigenvalues of a unitary matrix have absolute value 1.

Proposition 8.6. *Let $T\colon V \to V$ be a linear transformation on a finite-dimensional inner product space V. Let β be an orthonormal basis of V. Then $T\colon V \to V$ is unitary if and only if $[T]_\beta^\beta$ is unitary.*

Proof. Suppose T is unitary. Then $TT^* = T^*T = I_V$. Taking the matrix representation,

$$[T]_\beta^\beta[T^*]_\beta^\beta = [TT^*]_\beta^\beta = [T^*T]_\beta^\beta = [T^*]_\beta^\beta[T]_\beta^\beta = [I_V]_\beta^\beta = I.$$

From Theorem 5.12, $[T^*]_\beta^\beta$ is the adjoint of $[T]_\beta^\beta$. So, $[T]_\beta^\beta$ is unitary. We proved the forward implication. To prove the reverse implication, note that the above steps can be reversed. □

Corollary 8.7. *Let A be an $n \times n$ unitary matrix, and let $f(\lambda) := \det(A - \lambda I)$ be the characteristic polynomial of A. Then there exist $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ with $|\lambda_i| = 1$ for all $i \in \{1, \ldots, n\}$ such that $f(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda)$.*

Proof. From the Fundamental Theorem of Algebra (Theorem 1.4), there exist $\lambda_0, \ldots, \lambda_n \in \mathbf{C}$ such that $f(\lambda) = \lambda_0 \prod_{i=1}^n (\lambda_i - \lambda)$. Recall that coefficient of the degree n term of $f(\lambda)$ is $(-1)^n$ by Theorem 1.6. So, $\lambda_0 = 1$. From Remark 8.5, $|\lambda_i| = 1$ for all $i \in \{1, \ldots, n\}$. □

Lemma 8.8. *Let $\mathbf{F}$ denote either $\mathbf{R}$ or $\mathbf{C}$. Let $T\colon V \to V$ be a linear transformation on a finite-dimensional inner product space V over $\mathbf{F}$. Suppose there exists an orthonormal basis β of V consisting of eigenvectors of T with eigenvalues of absolute value 1. Then T is unitary.*

Proof. From Lemma 1.2, $[T]_\beta^\beta$ is diagonal with entries of absolute value 1. In particular, $[T]_\beta^\beta$ is unitary. So, from Proposition 8.6, T is unitary. □

Theorem 8.9 (The Spectral Theorem for Unitary Operators). *Let $T\colon V \to V$ be a unitary operator on a finite-dimensional inner product space V over $\mathbf{C}$. Then there exists an orthonormal basis β of V consisting of eigenvectors of T. In particular, T is diagonalizable. Moreover, all eigenvalues of T have absolute value 1.*

Proof. Since T is unitary, T is normal. So, the result follows directly from the Spectral Theorem for normal operators (Theorem 6.10). Then we apply Theorem 8.4 to finish. □

Remark 8.10. Note that Theorem 8.9 requires V to be a vector space over $\mathbf{C}$. In the case that V is a vector space over $\mathbf{R}$, the corresponding spectral theorem becomes very restricted, since the only real numbers with absolute value one are 1 and -1. So, if we want to diagonalize an orthogonal operator over $\mathbf{R}$, T must have all eigenvalues 1 or -1. Even though we can diagonalize an orthogonal operator over $\mathbf{C}$ by Theorem 8.9, we can essentially never diagonalize an orthogonal operator over $\mathbf{R}$. Nevertheless, let's present the result for diagonalization over $\mathbf{R}$.

If $\mathbf{F} = \mathbf{R}$, and if T is both self-adjoint and unitary, then the Spectral Theorem for self-adjoint operators (Theorem 7.10) together with Theorem 8.4 show: there exists an orthonormal basis β of V consisting of eigenvectors of T. So, T is diagonalizable, and all eigenvalues of T are 1 or -1. Conversely, suppose T is a linear operator on an inner product space V over $\mathbf{R}$, and suppose there exists a basis β of V consisting of eigenvectors of T with eigenvalues 1 or -1. Then $[T]_\beta^\beta$ is orthogonal. So T is orthogonal by Proposition 8.6. By Lemma 7.9, T is also self-adjoint.

9. Appendix: Notation

Let A, B be sets in a space X. Let m, n be a nonnegative integers. Let $\mathbf{F}$ be a field.

$$
\begin{aligned}
\mathbf{Z} &:= \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}, \text{ the integers} \\
\mathbf{N} &:= \{0, 1, 2, 3, 4, 5, \ldots\}, \text{ the natural numbers} \\
\mathbf{Q} &:= \{m/n \colon m, n \in \mathbf{Z}, n \neq 0\}, \text{ the rationals} \\
\mathbf{R} &\text{ denotes the set of real numbers} \\
\mathbf{C} &:= \{x + y\sqrt{-1} : x, y \in \mathbf{R}\}, \text{ the complex numbers} \\
\overline{x + y\sqrt{-1}} &:= x - y\sqrt{-1}, \quad x, y \in \mathbf{R}\text{ , the complex conjugate} \\
\emptyset &\text{ denotes the empty set, the set consisting of zero elements} \\
\in &\text{ means ``is an element of.'' For example, } 2 \in \mathbf{Z} \text{ is read as ``2 is an element of } \mathbf{Z}\text{.''} \\
\forall &\text{ means ``for all''} \\
\exists &\text{ means ``there exists''} \\
\mathbf{F}^n &:= \{(x_1, \ldots, x_n) \colon x_i \in \mathbf{F}, \forall\, i \in \{1, \ldots, n\}\} \\
A \subseteq B &\text{ means } \forall\, a \in A\text{, we have } a \in B\text{, so } A \text{ is contained in } B \\
A \setminus B &:= \{x \in A \colon x \notin B\} \\
A^c &:= X \setminus A, \text{ the complement of } A \\
A \cap B &\text{ denotes the intersection of } A \text{ and } B \\
A \cup B &\text{ denotes the union of } A \text{ and } B \\
C(\mathbf{R}) &\text{ denotes the set of all continuous functions from } \mathbf{R} \text{ to } \mathbf{R} \\
P_n(\mathbf{R}) &\text{ denotes the set of all real polynomials in one real variable of degree at most } n \\
P(\mathbf{R}) &\text{ denotes the set of all real polynomials in one real variable} \\
M_{m\times n}(\mathbf{F}) &\text{ denotes the vector space of } m \times n \text{ matrices over the field } \mathbf{F} \\
I_n &\text{ denotes the } n \times n \text{ identity matrix} \\
\det &\text{ denotes the determinant function} \\
S_n &\text{ denotes the set of permutations on } \{1, \ldots, n\} \\
\mathrm{sign}(\sigma) &:= (-1)^N \text{ where } \sigma \in S_n \text{ can be written as the composition of } N \text{ transpositions} \\
\mathrm{Tr} &\text{ denotes the trace function}
\end{aligned}
$$

9.1. Set Theory.

Set Theory. Let V, W be sets, and let $f \colon V \to W$ be a function. Let $X \subseteq V$, $Y \subseteq W$.

$$f(X) := \{f(v) \colon v \in V\}.$$

$$f^{-1}(Y) := \{v \in V \colon f(v) \in Y\}.$$

The function $f \colon V \to W$ is said to be **injective** (or **one-to-one**) if: for every $v, v' \in V$, if $f(v) = f(v')$, then $v = v'$.

The function $f \colon V \to W$ is said to be **surjective** (or **onto**) if: for every $w \in W$, there exists $v \in V$ such that $f(v) = w$.

The function $f\colon V \to W$ is said to be **bijective** (or a **one-to-one correspondence**) if: for every $w \in W$, there exists exactly one $v \in V$ such that $f(v) = w$. A function $f\colon V \to W$ is bijective if and only if it is both injective and surjective.

Two sets X, Y are said to have the same **cardinality** if there exists a bijection from V onto W.

The **identity map** $I\colon X \to X$ is defined by $I(x) = x$ for all $x \in X$. To emphasize that the domain and range are both X, we sometimes write I_X for the identity map on X. Let $f\colon X \to X$. We write f^2 to denote f composed with itself: $f \circ f$. More generally, for any $n \in \mathbf{N}$, we write f^n to denote f composed with itself n times: $f \circ f \circ \cdots \circ f$.

Let V, W be vector spaces over a field $\mathbf{F}$. Then $\mathcal{L}(V, W)$ denotes the set of linear transformations from V to W, and $\mathcal{L}(V)$ denotes the set of linear transformations from V to V. Let $T\colon V \to W$ be a linear transformation between inner product spaces. Then $T^*\colon W \to V$ denotes the adjoint of T.

UCLA Department of Mathematics, Los Angeles, CA 90095-1555
E-mail address: heilman@math.ucla.edu