Linear Algebra

Linear Algebra Notes

By

Daniel Ruiz

Last Updated: March 2020

Prelude

This set of notes covers Linear Algebra from the textbook *Linear Algebra Done Right* by *Sheldon Axler.* These notes are one of many that I have decided to make available for anyone interested. One of the benefits of this series are having a compendium of definitions, examples and personal thoughts that I can always refer back to if I need a reminder on a particular topic. In addition, I find that this medium reduces the search time for specific definitions, theorems, examples etc and thus aids in reinforcing my own knowledge when frequented.

Contents

Chapter 1

Vector Spaces

Linear Algebra is the study of linear maps on finite-dimensional vector spaces. This chapters aims to define vector spaces and discuss their elementary properties.

1.1 $\mathbb{R}^n$ and $\mathbb{C}^n$

1.1.1 Complex Numbers

Definition 1.1.1: Complex Numbers & their Structure

*A **complex number** is an ordered pair (a, b), where $a, b \in \mathbb{R}$. We will typically write (a, b) as $a + bi$, with i satisfying $i^2 = -1$ (i is typically referred to as the imaginary unit). The set of all complex numbers is denoted by $\mathbb{C}$:*

$$\mathbb{C} = \{a + bi \ : \ a, b \in \mathbb{R}\} \tag{1.1}$$

We define addition and multiplication on $\mathbb{C}$ by

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \tag{1.2}$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i \tag{1.3}$$

$\forall a, b, c, d \in \mathbb{R}$. If $a \in \mathbb{R}$, we identify $a + 0i$ with the real number a. In this sense, one can consider $\mathbb{R}$ as a subset of $\mathbb{C}$. Similarly, we will usually write $0 + bi$ as just bi and $0 + 1i$ as just i.

Proposition 1.1.1: Properties of Complex Arithmetic

1. ***Commutativity**:*

$$\alpha + \beta = \beta + \alpha \text{ and } \alpha\beta = \beta\alpha \ \forall \alpha, \beta \in \mathbb{C} \tag{1.4}$$

2. ***Associativity**:*

$$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda) \text{ and } (\alpha\beta)\lambda = \alpha(\beta\lambda) \ \forall \alpha, \beta, \lambda \in \mathbb{C} \tag{1.5}$$

3. ***Identities:***

$$\lambda + 0 = \lambda \text{ and } \lambda 1 = \lambda \ \forall \lambda \in \mathbb{C} \tag{1.6}$$

4. ***Additive Inverse:***

$$\text{For every } \alpha \in \mathbb{C}, \text{there exists a unique } \beta \in \mathbb{C} \text{ such that } \alpha + \beta = 0 \tag{1.7}$$

5. ***Multiplicative Inverse**:*

$$\text{For every } \alpha \in \mathbb{C} \text{ where } \alpha \neq 0, \text{there exists a unique } \beta \in \mathbb{C} \text{ such that } \alpha\beta = 1 \tag{1.8}$$

6. ***Distributive Property***:
$$\lambda(\alpha+\beta) = \lambda\alpha + \lambda\beta \text{ for all } \lambda, \alpha, \beta \in \mathbb{C} \tag{1.9}$$

Definition 1.1.2: Subtraction and Division on $\mathbb{C}$

Let $\alpha, \beta \in \mathbb{C}$.

- *Let $-\alpha$ denote the additive inverse of α, so that it is the unique number satisfying $\alpha + (-\alpha) = 0$. Then, we define subtraction on $\mathbb{C}$ by*
$$\beta - \alpha := \beta + (-\alpha) \tag{1.10}$$
- *For $\alpha \neq 0$, let $1/\alpha$ denote the multiplicative inverse of α. Thus, $1/\alpha$ is the unique number satisfying $\alpha(1/\alpha) = 1$. Then, we define division on $\mathbb{C}$ by*
$$\beta/\alpha := \beta(1/\alpha) \tag{1.11}$$

Notation: $\mathbb{F}$

Throughout these notes, $\mathbb{F}$ stands for either $\mathbb{R}$ or $\mathbb{C}$. The chosen notation $\mathbb{F}$ is used because both $\mathbb{R}$ and $\mathbb{C}$ are examples of fields. Hence, if we prove a theorem involving $\mathbb{F}$, it holds for both $\mathbb{R}$ and $\mathbb{C}$.

Definition 1.1.3: Scalars

Let $\mathbb{F}$ be a field. Then, we refer to elements of $\mathbb{F}$ as ***scalars****.*

Definition 1.1.4: Integral Exponentiation

Let $\alpha \in \mathbb{F}$ and m be a positive integer. We define α^m as the product of α with itself m times:
$$\alpha^m = \underbrace{\alpha \cdots \alpha}_{m \text{ times}} \tag{1.12}$$

1.1.2 Lists

Definition 1.1.5: List, Length

Let n be a nonnegative integer. A ***list*** *of* ***length*** *n is an ordered collection of n elements (which can be numbers, lists or other abstract entities) separated by commas and surrounded by parentheses. A list of length n looks like*
$$(x_1, ..., x_n). \tag{1.13}$$
*Two lists are considered equal if and only if they have the same length and the same elements in the same order. In the mathematical literature, a list of length n is also often referred to as an n-****tuple****. A list of length 0 is written as $()$.*

1.1.3 $\mathbb{F}^n$

Definition 1.1.6: $\mathbb{F}^n$ and Coordinates

We define $\mathbb{F}^n$ as the set of all lists of length n of elements of $\mathbb{F}$:

$$\mathbb{F}^n := \{(x_1, ..., x_n) : x_j \in \mathbb{F} \text{ for } j = 1, ..., n\} \tag{1.14}$$

For $(x_1, ..., x_n) \in \mathbb{F}^n$ and $j \in \{1, ..., n\}$, we say that x_j is the jth ***coordinate*** *of $(x_1, ..., x_n)$.*

Definition 1.1.7: Addition on $\mathbb{F}^n$

We define ***addition*** *in $\mathbb{F}^n$ by adding the corresponding coordinates:*

$$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n) \tag{1.15}$$

Proposition 1.1.2: Commutativity of addition on $\mathbb{F}^n$

Let $x, y \in \mathbb{F}^n$, then $x + y = y + x$.

Definition 1.1.8: Zero Vector 0

Let 0 denote the list of length n whose coordinates are all 0:

$$0 = (0, ..., 0) \tag{1.16}$$

Definition 1.1.9: Additive Inverse in $\mathbb{F}^n$

Let $x \in \mathbb{F}^n$. Then, the ***additive inverse*** *of x, denoted by $-x$, is the vector $-x \in \mathbb{F}^n$ such that*

$$x + (-x) = 0. \tag{1.17}$$

In essence, if $x = (x_1, ..., x_n)$, then $-x = (-x_1, ..., -x_n)$.

Definition 1.1.10: Scalar Multiplication in $\mathbb{F}^n$

The ***product*** *of a number λ and a vector in $\mathbb{F}^n$ is computed by multiplying each coordinate of the vector by λ:*

$$\lambda(x_1, ..., x_n) = (\lambda x_1, ..., \lambda x_n), \tag{1.18}$$

where $\lambda \in \mathbb{F}$ and $(x_1, ..., x_n) \in \mathbb{F}^n$.

1.1.4 Digression on Fields

Definition 1.1.11: Field

A ***field*** *is a set containing at least two distinct elements called 0 and 1 (representing the additive inverse and multiplicative inverses respectively) along with operations of addition and multiplication satisfying all the properties listed in Proposition 1.1.1.*

1.2 Definition of Vector Space

Definition 1.2.1: Addition Operator

*An **addition** on a set V is a function $+ : V \times V \to V$ that assigns an element $u + v \in V$ to every pair $u, v \in V$.*

Definition 1.2.2: Scalar Multiplication

*A **scalar multiplication** on a set V is a function $* : V \times \mathbb{F} \to V$ that assigns an element $\lambda v \in V$ to each $\lambda \in \mathbb{F}$ and each $v \in V$.*

Definition 1.2.3: Vector Space

*A **vector space** is a set V, along with an addition on V and a scalar multiplication on V such that the following properties hold:*

1. ***Commutativity**:*
$$u + v = v + u \ \forall u, v \in V \tag{1.19}$$
2. ***Associativity**:*
$$(u + v) + w = u + (v + w) \text{ and } (ab)v = a(bv) \ \forall u, v, w \in V \text{and } \forall a, b \in \mathbb{F} \tag{1.20}$$
3. ***Additive Identity**:*
$$\textit{There exists an element } 0 \in V \textit{such that } v + 0 = v \ \forall v \in V \tag{1.21}$$
4. ***Additive Inverse**:*
$$\textit{For every } u \in V, \textit{there exists } w \in V \textit{ such that } v + w = 0 \tag{1.22}$$
5. ***Multiplicative Identity**:*
$$1v = v \ \forall v \in V \tag{1.23}$$
6. ***Distributive Properties**:*
$$a(u + v) = au + av \text{ and } (a + b)v = av + bv \ \forall a, b \in \mathbb{F} \text{ and } \forall u, v \in V \tag{1.24}$$

*The scalar multiplication in a vector space depends on $\mathbb{F}$. Hence, to be precise, we will say that V is a **vector space over** $\mathbb{F}$.*

Definition 1.2.4: Vector / Point

*Elements of a vector space are called **vectors** or **points**.*

Definition 1.2.5: Real Vector Space, Complex Vector Space

- *A vector space over $\mathbb{R}$ is called a **real vector space**.*
- *A vector space over $\mathbb{C}$ is called a **complex vector space**.*

Definition 1.2.6: Notation: $\mathbb{F}^S$

Let S be a set. Then, $\mathbb{F}^S$ denotes the set of all functions from S to $\mathbb{F}$.

- *Let $f, g \in \mathbb{F}^S$, the **sum** $f + g \in \mathbb{F}^S$ is the function defined by*
$$(f+g)(x) = f(x) + g(x) \tag{1.25}$$
$\forall x \in S$.
- *Let $\lambda \in \mathbb{F}$ and $f \in \mathbb{F}^S$, the **product** $\lambda f \in \mathbb{F}^S$ is the function defined by*
$$(\lambda f)(x) = \lambda f(x) \tag{1.26}$$
$\forall x \in S$.

If S is a nonempty set, then $\mathbb{F}^S$, with the operations of addition and scalar multiplication defined above is a vector space over $\mathbb{F}$. For the rest of these notes, V denotes a vector space over $\mathbb{F}$.

Proposition 1.2.1: Unique Additive Identity

A vector space has a unique additive identity.

Proposition 1.2.2: Unique Additive Inverse

A vector space has a unique additive inverse.

Definition 1.2.7: Notation: $-v$, $w - v$

Let $v, w \in V$. Then

- *$-v$ denotes the additive inverse of v*
- *$w - v$ is defined to be $w + (-v)$*

Proposition 1.2.3

Let V be a vector space. Then, $0v = 0 \ \forall v \in V$.

Proposition 1.2.4

Let $\mathbb{F}$ be a field. Then, $a0 = 0 \ \forall a \in \mathbb{F}$.

Corollary 1.2.1

Let V be a vector space. Then, $(-1)v = -v \ \forall v \in V$.

1.3 Subspaces

Definition 1.3.1: Subspace

A subset U of V is called a ***subspace*** *of V if U is also a vector space (using the same addition and scalar multiplication as on V). Some mathematicians use the term* ***linear subspace****, which means the same as subspace.*

Theorem 1.3.1: Conditions for a Subspace

A subset U of V is a subspace of V if and only if U satisfies the following three conditions:

1. ***Additive Identity***
$$0 \in U \tag{1.27}$$
2. ***Closed Under Addition***
$$u, w \in U \text{ implies } u + w \in U \tag{1.28}$$
3. ***Closed under Scalar Multiplication***
$$a \in \mathbb{F} \text{ and } u \in U \text{ implies } au \in U \tag{1.29}$$

1.3.1 Sums of Subspaces

Definition 1.3.2: Sum of Subsets

Suppose that $U_1, ..., U_m$ are subsets of V. The ***sum*** *of $U_1, ..., U_m$, denoted as $U_1 + ... + U_m$, is the set of all possible sums of elements of $U_1, ..., U_m$. More precisely,*

$$U_1 + ... + U_m = \{u_1 + ... + u_m : u_1 \in U_1, ..., u_m \in U_m\} \tag{1.30}$$

Proposition 1.3.1: Sum of Subspaces is the smallest containing Subspace

Suppose that $U_1, ..., U_m$ are subspaces of V. Then $U_1 + ... + U_m$ is the smallest subspace of V containing $U_1, ..., U_m$.

1.3.2 Direct Sums

Definition 1.3.3: Direct Sum

Suppose that $U_1, ..., U_m$ are subspaces of V.

- *The sum $U_1 + ... + U_m$ is called a* ***direct sum*** *if each element of $U_1 + ... + U_m$ can be written in only one way as a sum $u_1 + ... + u_m$, where each u_j is in U_j*[a][b]*.*
- *If $U_1 + ... + U_m$ is a direct sum, then $U_1 \oplus ... \oplus U_m$ denotes $U_1 + ... + U_m$, with the $\oplus$ notation serving as an indication that this is a direct sum.*

The definition of direct sum requires that every vector in the sum have a unique representation as an appropriate sum.

[a] Typically, suppose that A, B were two Abelian groups or objects with an abelian structure. Then, we can define the direct

sum as the set of ordered pairs $A \oplus B = \{(a,b) | a \in A, b \in B\}$ with the new addition structure defined coordinate-wise.

[b] The uniqueness condition in writing the summation is sufficient so that we may identify it with the usual notion of direct sum.

Proposition 1.3.2: Condition for a Direct Sum

Suppose that $U_1, ..., U_m$ are subspaces of V. Then $U_1 + ... + U_m$ is a direct sum if and only if the only way to write 0 as a sum $u_1 + ... + u_m$, where each $u_j \in U_j$, is by taking each u_j equal to 0.

Proposition 1.3.3: Direct Sum of Two Subspaces

Suppose U and W are subspaces of V. Then, $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

1.4 Exercises

Definition 1.4.1: Even Function

*A function $f : \mathbb{R} \to \mathbb{R}$ is called **even** if*

$$f(-x) = f(x) \tag{1.31}$$

for all $x \in \mathbb{R}$.

Definition 1.4.2: Odd Function

*A function $f : \mathbb{R} \to \mathbb{R}$ is called **odd** if*

$$f(-x) = -f(x) \tag{1.32}$$

for all $x \in \mathbb{R}$.

Proposition 1.4.1

Let U_e denote the set of real-valued even functions on $\mathbb{R}$ and let U_o denote the set of real-valued odd functions on $\mathbb{R}$. Then,

$$\mathbb{R}^{\mathbb{R}} = U_e \oplus U_o \tag{1.33}$$

Proof. First, we note that any function $f : \mathbb{R} \to \mathbb{R}$ can be expressed as a sum of an even and odd function. We define $f_e(x) = \frac{f(x)+f(-x)}{2}$ and $f_o(x) = \frac{f(x)-f(-x)}{2}$. Then $f = f_e + f_o$ where f_e and f_o are even and odd functions respectively. We will first observe that $\mathbb{R}^{\mathbb{R}} = U_e + U_o$. Let $f \in U_e$ and $g \in U_o$, then $f + g : \mathbb{R} \to \mathbb{R}$. Hence, we necessarily have $f + g \in \mathbb{R}^{\mathbb{R}}$. Now, let $h \in \mathbb{R}^{\mathbb{R}}$. Then, we define h_e and h_o as the even and odd components of h respectively. Since $h_e \in U_e$ and $h_o \in U_o$, then $h \in U_e + U_o$. Hence, we have established that $\mathbb{R}^{\mathbb{R}} = U_e + U_o$.

We now want to show that $\mathbb{R}^{\mathbb{R}} = U_e \oplus U_o$. We observe that U_e and U_o are subspaces of $\mathbb{R}^{\mathbb{R}}$ as they satisfy all the properties of a subspace. It's clear that aside from the zero function, a function cannot simultaneously be both even and odd[a]. Since U_e, U_o are subspaces of $\mathbb{R}^{\mathbb{R}}$ and $U_e \cap U_o = \{0\}$, by Proposition 1.3.3, $U_e + U_o$ is a direct sum, thereby establishing

$$\mathbb{R}^{\mathbb{R}} = U_e \oplus U_o$$

□

[a]Suppose that a function f was both even and odd. Then, we can conclude that $f(x) = -f(x)$ $\forall x \in \mathbb{R}$. Hence, $f(x) = 0$ $\forall x \in \mathbb{R}$, signifying that $f = 0$.

Chapter 2

Finite-Dimensional Vector Spaces

Definition 2.0.1: Notation: $\mathbb{F}$, V

- *$\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$*
- *V denotes a vector space over $\mathbb{F}$*

2.1 Span and Linear Independence

Definition 2.1.1: List of Vectors

We will usually write lists of vectors without surrounding parentheses[a]*.*

[a]For instance $(4, 1, 6), (9, 5, 7)$ is a list of length 2 of vectors in $\mathbb{R}^3$.

2.1.1 Linear Combinations and Span

Definition 2.1.2: Linear Combination

*A **linear combination** of a list $v_1, .., v_m$ of vectors in V is a vector of the form*

$$a_1v_1 + ... + a_mv_m, \tag{2.1}$$

where $a_1, .., a_m \in \mathbb{F}$.

Definition 2.1.3: Span

*The set of all linear combinations of a list of vectors $v_1, ..., v_m \in V$ is called the **span** of $v_1, ..., v_m$, denoted as span$(v_1, .., v_m)$. In essence,*

$$\text{span}(v_1, ..., v_m) = \{a_1v_1 + ... + a_mv_m : a_1, ..., a_m \in \mathbb{F}\} \tag{2.2}$$

*The span of the empty list $()$ is defined to be $\{0\}$. Some mathematicians use the term **linear span**, which means the same as span.*

Proposition 2.1.1: Span is the Smallest Containing Subspace

The span of a list of vectors in V is the smallest subspace of V containing all the vectors in the list.

Definition 2.1.4: Spans

If $\text{span}(v_1, ..., v_m)$ *equals* V*, we say that* $v_1, .., v_m$ ***spans*** V.

Definition 2.1.5: Finite-Dimensional Vector Space

*A vector space is called **finite-dimensional** if some list of vectors in it spans the space.*

Definition 2.1.6: Polynomial, $\mathcal{P}(\mathbb{F})$

- *A function* $p : \mathbb{F} \to \mathbb{F}$ *is called a **polynomial** with coefficients in* $\mathbb{F}$ *if there exist* $a_0, ..., a_m \in \mathbb{F}$ *such that*
$$p(z) = a_0 + a_1 z + a_2 z^2 + ... + a_m z^m \tag{2.3}$$
for all $z \in \mathbb{F}$.
- $\mathcal{P}(\mathbb{F})$ *is the set of all polynomials with coefficients in* $\mathbb{F}$.

With the usual operations of addition and scalar multiplication, $\mathcal{P}(\mathbb{F})$ *is a vector space over* $\mathbb{F}$.

Definition 2.1.7: Degree of a Polynomial, deg p

- *A polynomial* $p \in \mathcal{P}(\mathbb{F})$ *is said to have **degree*** m *if there exist scalars* $a_0, a_1, ..., a_m \in \mathbb{F}$ *with* $a_m \neq 0$ *such that*
$$p(z) = a_0 + a_1 z + ... + a_m z^m \tag{2.4}$$
for all $z \in \mathbb{F}$*. If* p *has degree* m*, we write* $\deg p = m$.
- *The polynomial that is identically* 0 *is said to have degree* $-\infty$.

Definition 2.1.8: $\mathcal{P}_m(\mathbb{F})$

For m *a nonnegative integer,* $\mathcal{P}_m(\mathbb{F})$ *denotes the set of all polynomials with coefficients in* $\mathbb{F}$ *and degree at most* m*. We note that* $\mathcal{P}_m(\mathbb{F})$ *is a finite-dimensional vector space for each non-negative integer* m.

Definition 2.1.9: Infinite-Dimensional Vector Space

*A vector space is called **infinite-dimensional** if it is not finite-dimensional.*

Example 2.1.1

$\mathcal{P}(\mathbb{F})$ *is infinite dimensional.*

Proof. Suppose by contradiction, that there exists a list of elements $(p_1, ..., p_n)$ that spans $\mathcal{P}(\mathbb{F})$. Then, let

$$m := \max_{j \in \mathbb{Z}_n} \deg(p_j) \tag{2.5}$$

Consider the polynomial $q(z) = z^{m+1}$. Clearly, $q \in \mathcal{P}(\mathbb{F})$ but $q \notin \text{span}(p_1, .., p_m)$. Hence, we have a contradiction. Therefore, $\mathcal{P}(\mathbb{F})$ must be infinite dimensional. □

2.1.2 Linear Independence

Definition 2.1.10: Linear Independence

- *A list $v_1, .., v_m$ of vectors in V is said to be* ***linearly independent*** *if the only choice of $a_1, ..., a_m \in \mathbb{F}$ that satisfies*

$$a_1 v_1 + ... + a_m v_m = 0 \tag{2.6}$$

  is if $a_1 = .. = a_m = 0$.

- *The empty list $()$ is defined to be linearly independent.*

This reasoning shows that $v_1, .., v_m$ is linearly independent if and only if each vector in span$(v_1, ..v_m)$ has only one representation as a linear combination of $v_1, .., v_m$.

Proposition 2.1.2

Let $(v_1, .., v_m)$ denote a linearly independent list. If some vectors are removed from this list, the remaining list is also linearly independent.

Proof. WLOG, consider the list $(v_1, .., v_j)$ where $j < m$. Then, suppose by contradiction that there exists a set of values $b_1, .., b_j \in \mathbb{F}$ that are not all zero so that

$$\sum_{i=1}^{j} b_i v_i = 0. \tag{2.7}$$

Since $(v_1, .., v_m)$ are linearly independent, we must have that the only solution to

$$\sum_{i=1}^{m} a_i v_i = 0 \tag{2.8}$$

is if $a_i = 0 \ \forall i$. However, we can clearly see that

$$\sum_{i=1}^{m} a_i v_i = \sum_{i=1}^{j} a_i v_i + \sum_{k=j+1}^{m} a_i v_m = 0 \tag{2.9}$$

is also satisfied if $a_i = b_i$ for $1 \leq i \leq j$ and $a_i = 0$ for $j + 1 \leq i \leq m$, which by construction is not all zeroes. This is a contradiction. Hence, any new list obtained by the removal of some elements from a linearly independent list must itself be linearly independent. □

Definition 2.1.11: Linearly Dependent

- *A list of vectors in V is called **linearly dependent** if it is not linearly independent.*
- *In essence, a list $v_1, ..., v_m$ of vectors in V is linearly dependent if there exist $a_1, .., a_m \in \mathbb{F}$, not all 0, such that*

$$a_1 v_1 + ... + a_m v_m = 0 \tag{2.10}$$

Lemma 2.1.1: Linear Dependence Lemma

Suppose that $v_1, .., v_m$ is a linearly dependent list in V. Then, there exists $j \in \{1, 2, ..., m\}$ such that the following hold:

(a) $v_j \in \text{span}(v_1, ..., v_{j-1})$

(b) If the j^{th} term is removed from $v_1, .., v_m$, the span of the remaining list equals $\text{span}(v_1, ..., v_m)$.

Theorem 2.1.1: Length of Linearly Independent List $\leq$ Length of Spanning List

In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proposition 2.1.3: Finite-Dimensional Subspaces

Every subspace of a finite-dimensional vector space is finite-dimensional.

2.2 Bases

Definition 2.2.1: Basis

*A **basis** of V is a list of vectors in V that is linearly independent and spans V.*

Corollary 2.2.1: Basis Span

Let V be a finite-dimensional vector space. Suppose that $v_1, .., v_n$ forms a basis for V, then $V = \text{span}(v_1, ..., v_n)$.

Definition 2.2.2: Standard Basis

*The list $(1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1)$ is a basis of $\mathbb{F}^n$, called the **standard basis** of $\mathbb{F}^n$.*

Proposition 2.2.1: Criterion for Basis

A list $v_1, ..., v_n$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form

$$v = a_1 v_1 + ... + a_n v_n \tag{2.11}$$

where $a_1, ..., a_n \in \mathbb{F}$.

Proposition 2.2.2: Spanning List Contains a Basis

Every spanning list in a vector space can be reduced to a basis of the vector space.

Proposition 2.2.3: Basis of Finite-Dimensional Vector Space

Every finite-dimensional vector space has a basis.

Proposition 2.2.4: Linearly Independent List Extends to a Basis

Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Theorem 2.2.1: Every Subspace of V is part of a Direct Sum equal to V

Suppose that V is finite-dimensional and U is a subspace of V. Then there is a subspace W of V such that $V = U \oplus W$.

2.3 Dimension

Proposition 2.3.1: Basis Length does not Depend on Basis

Any two bases of a finite-dimensional vector space have the same length.

Definition 2.3.1: Dimension, dim V

- *The **dimension** of a finite-dimensional vector space is the length of any basis of the vector space.*
- *The dimension of V (if V is finite-dimensional) is denoted by $\dim V$.*

Proposition 2.3.2: Dimension of a Subspace

If V is finite-dimensional and U is a subspace of V, then $\dim U \leq \dim V$.

Proposition 2.3.3: Linearly Independent List of the Right Length is a Basis

Suppose that V is finite-dimensional. Then, every linearly independent list of vectors in V with length $\dim V$ is a basis of V.

Proposition 2.3.4: Spanning List of the Right Length is a Basis

Suppose that V is finite-dimensional. Then, every spanning list of vectors in V with length $\dim V$ is a basis of V.

Theorem 2.3.1: Dimension of a Sum

If U_1 and U_2 are subspaces of a finite-dimensional vector space, then

$$dim(U_1 + U_2) = dim(U_1) + dim(U_2) - dim(U_1 \cap U_2) \tag{2.12}$$

2.4 Exercises

Proposition 2.4.1

Suppose V is finite-dimensional, with dim $V = n \geq 1$. Then, there exists 1-dimensional subspaces $U_1, .., U_n$ of V such that

$$V = U_1 \oplus ... \oplus U_n \tag{2.13}$$

Proof. Let $v_1, ..., v_n$ denote a basis for V. Then, we define $U_i = \text{span}(v_i)$. It's clear that U_i is a subset of V, $0 \in U_i$ and all U_i's are closed under addition and scalar multiplication. Hence, we observe that U_i are subspaces of V.

The length of the list of vectors spanning U_i's is 1, hence dim $U_i = 1$ $\forall i$. We want to first show that $V = U_1 + ... + U_n$. We note that

$$\sum_{i=1}^{n} \text{span}(v_i) = \text{span}(v_1, v_2, ..., v_n), \tag{2.14}$$

Since $v_1, .., v_n$ forms a basis for V, then $V = span(v_1, ..., v_n)$. Hence, we have that $U_1 + ... + U_n = V$.

Lastly, consider an element $u_1 + ... + u_n \in U_1 + ... + U_n$, where $u_j \in U_j$. Then, this can take the form $u_1 + ... + u_n = \sum_{i=1}^{n} a_i v_i$ for some $a_i \in \mathbb{F}$. Since $v_1, .., v_n$ form a basis for $U_1 + ... + U_n$, then the only way to satisfy $\sum_{i=1}^{n} a_i v_i = 0$ is if $a_i = 0$ $\forall i$. In other words, $\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} a_i v_i = 0$ is only satisfied if $u_i = 0$ $\forall i$. Then, by Proposition 1.3.2, $U_1 + ... + U_n$ is a direct sum. Consequently,

$$V = U_1 \oplus ... \oplus U_n \tag{2.15}$$

□

Proposition 2.4.2

Suppose $U_1, .., U_m$ are finite-dimensional subspaces of V such that $U_1 + .. + U_m$ is a direct sum. Then, $U_1 \oplus ... \oplus U_m$ is finite-dimensional and

$$dim(U_1 \oplus ... \oplus U_m) = dim(U_1) + ... + dim(U_m) \tag{2.16}$$

Chapter 3

Linear Maps

Definition 3.0.1: Notation: $\mathbb{F}, V, W$

- *$\mathbb{F}$ denotes $\mathbb{R}$ or $\mathbb{C}$.*
- *V and W denote vector spaces over $\mathbb{F}$.*

3.1 The Vector Space of Linear Maps

Definition 3.1.1: Linear Map

*A **linear map** from V to W is a function $T : V \to W$ with the following properties:*

1. ***Additivity***
$$T(u + v) = Tu + Tv \text{ for all } u, v \in V \tag{3.1}$$
2. ***Homogeneity***
$$T(\lambda v) = \lambda(Tv) \text{ for all } \lambda \in \mathbb{F} \text{ and all } v \in V. \tag{3.2}$$

*Some mathematicians use the term **linear transformation** which means the same as linear map. For linear maps, we will often use the notation Tv as well as the more standard functional notation $T(v)$.*

Definition 3.1.2: Notation: $\mathcal{L}(V, W)$ or $\mathrm{Hom}_{\mathbb{F}}(V, W)$

Let V, W be vector spaces over $\mathbb{F}$. The set of all linear maps from V to W will be denoted by $\mathcal{L}(V, W)$. We note that a lot of literature also denotes this by $\mathrm{Hom}_{\mathbb{F}}(V, W)$.

Common Linear Maps

- ***Zero***
We'll let 0 denote the function that takes each element of some vector space to the additive identity of another vector space. Specifically, $0 \in \mathcal{L}(V, W)$ is defined by
$$0v = 0 \ \forall v \in V \tag{3.3}$$
The zero on the left is a function from V to W, whereas the 0 on the right side is the additive identity in W.

- ***Identity***
  *The **identity map**, denoted I, is the function on some vector space that takes each element to itself. Specifically, $I \in \mathcal{L}(V,V)$ is defined by*
  $$Iv = v \ \forall v \in V \tag{3.4}$$

- ***Differentiation***
  Define $D \in \mathcal{L}\big(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R})\big)$ by
  $$Dp = p' \tag{3.5}$$

Proposition 3.1.1: Linear Maps and Basis of Domain

Suppose that $v_1, .., v_n$ is a basis of V and $w_1, ..., w_n \in W$. Then, there exists a unique linear map $T : V \to W$ such that
$$Tv_j = w_j \tag{3.6}$$
for each $j = 1, ..., n$.

3.1.1 Algebraic Operations on $\mathcal{L}(V,W)$

Definition 3.1.3: Addition on Scalar Multiplication on $\mathcal{L}(V,W)$

*Suppose that $S, T \in \mathcal{L}(V,W)$ and $\lambda \in \mathbb{F}$. The **sum** $S+T$ and the **product** λT are the linear maps from V to W defined by*
$$(S+T)(v) = Sv + Tv \text{ and } (\lambda T)(v) = \lambda(Tv) \tag{3.7}$$
for all $v \in V$.

Theorem 3.1.1: $\mathcal{L}(V,W)$ is a Vector Space

With the operations of addition and scalar multiplication as defined above, $\mathcal{L}(V,W)$ is a vector space.

Definition 3.1.4: Product of Linear Maps

*If $T \in \mathcal{L}(U,V)$ and $S \in \mathcal{L}(V,W)$, then the **product** $ST \in \mathcal{L}(U,W)$ is defined by*
$$(ST)(u) = S(Tu) \tag{3.8}$$
for $u \in U$. Hence, ST is the usual composition $S \circ T$ of two functions, but when both functions are linear, most mathematicians write ST instead of $S \circ T$.

Proposition 3.1.2: Algebraic Properties of Products of Linear Maps

- ***Associativity***
  $$(T_1T_2)T_3 = T_1(T_2T_3) \tag{3.9}$$
  whenever T_1, T_2, and T_3 are linear maps such that the products make sense (meaning that T_3 maps into domain of T_2 and T_2 maps into the domain of T_1).

- ***Identity***
  $$TI_V = I_WT = T \tag{3.10}$$

whenever $T \in \mathcal{L}(V, W)$. We have denoted I_V as the identity map on V and I_W as the identity map on W.

- ***Distributive Properties***

$$(S_1 + S_2)T = S_1T + S_2T \text{ and } S(T_1 + T_2) = ST_1 + ST_2 \tag{3.11}$$

whenever $T, T_1, T_2 \in \mathcal{L}(V, W)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$.

Proposition 3.1.3: Linear Maps take 0 to 0

Let T be a linear map from V to W. Then $T(0) = 0$.

3.2 Null Spaces and Ranges

3.2.1 Null Space and Injectivity

Definition 3.2.1: Null Space, null T / Kernel

*Let $T \in \mathcal{L}(V, W)$. The **null space** of T, denoted null T, is the subset of V consisting of those vectors that T maps to 0:*

$$\text{null } T = \{v \in V : Tv = 0\} \tag{3.12}$$

*Some mathematicians use the term **kernel** instead of null space.*

Proposition 3.2.1: The Null Space is a Subspace

Let $T \in \mathcal{L}(V, W)$. Then null T is a subspace of V.

Proposition 3.2.2: Injectivity is Equivalent to Null Space equals $\{0\}$

Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if null $T = \{0\}$.

Definition 3.2.2: Injective

*A function $T : V \to W$ is called **injective** if $Tu = Tv$ implies that $u = v$. Many mathematicians also use the term* one-to-one *to mean injectivity.*

3.2.2 Range and Surjectivity

Definition 3.2.3: Range

*Let $T \in \mathcal{L}(V, W)$. The **range** of T is the subset of W consisting of those vectors that are of the form Tv for some $v \in V$:*

$$\text{range } T = \{Tv : v \in V\} \tag{3.13}$$

Some mathematicians also use the term ***image*** *to denote the range*[a]*, though this is used in the wider context for any map* $f : A \to B$*, where* A, B *are sets and* f *is an arbitrary function.*

[a]In the broader sense, range is equivalent to image but our consideration is restricted to linear maps between vector spaces.

Proposition 3.2.3: The Range is a Subspace

If $T \in \mathcal{L}(V, W)$*, then range* T *is a subspace of* W*.*

Definition 3.2.4: Surjective

A function $T : V \to W$ *is called* ***surjective*** *if its range equals* W*. Many mathematicians also use the term onto to mean surjectivity.*

3.2.3 Fundamental Theorem of Linear Maps

Theorem 3.2.1: Fundamental Theorem of Linear Maps

Suppose that V *is finite-dimensional and* $T \in \mathcal{L}(V, W)$*. Then range* T *is finite-dimensional and*

$$\text{dim } V = \text{dim null } T + \text{dim range } T. \tag{3.14}$$

Proposition 3.2.4: A map to a smaller dimensional space is not injective

Suppose that V *and* W *are finite-dimensional vector spaces such that dim* $V >$ *dim* V*. Then no linear map from* V *to* W *is injective.*

Proposition 3.2.5: A map to a larger dimensional space is not surjective

Suppose V *and* W *are finite-dimensional vector spaces such that dim* $V <$ *dim* W*. Then no linear map from* V *to* W *is surjective.*

Definition 3.2.5: Homogeneous / Inhomogeneous System of Linear Equations

Fix positive integers m *and* n*. Let* $A_{j,k} \in \mathbb{F}$ *for* $j = 1, ..., m$ *and* $k = 1, ..., n$*. A system of linear equations is defined as the collection of equalities:*

$$\sum_{k=1}^{n} A_{1,k} x_k = c_1 \tag{3.15}$$

$$\vdots \tag{3.16}$$

$$\vdots \tag{3.17}$$

$$\sum_{k=1}^{n} A_{m,k} x_k = c_m \tag{3.18}$$

where $x_i, c_i \in \mathbb{F}$*. We say that this system of linear equations is* ***homogeneous*** *iff* $c_1 = ... = c_m = 0 \ \forall i$ *and* ***inhomogeneous*** *otherwise.*

Proposition 3.2.6: Homogeneous System of Linear Equations

A homogeneous system of linear equations with more variables than equations has nonzero solutions.

Proposition 3.2.7: Inhomogeneous System of Linear Equations

An inhomogeneous system of linear equations with more equations than variables has no solution for some choice of the constant terms.

3.3 Matrices

3.3.1 Representing a Linear Map by a Matrix

Definition 3.3.1: Matrix, $A_{j,k}$

*Let m and n denote positive integers. An m-by-n **matrix** A is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns:*

$$A = \begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & \vdots & \vdots \\ A_{m,1} & \dots & A_{m,n} \end{bmatrix} \tag{3.19}$$

The notation $A_{j,k}$ denotes the entry in row j, column k of A.

Definition 3.3.2: Matrix of a Linear Map, $\mathcal{M}(T)$

*Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, ..., v_n$ is a basis of V and $w_1, ..., w_m$ is a basis of W. The **matrix of** T with respect to these bases is the m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by*

$$Tv_k = A_{1,k}w_1 + ... + A_{m,k}w_m. \tag{3.20}$$

If the bases are not clear from the context, then the notation $\mathcal{M}\big(T, (v_1, .., v_n), (w_1, ..., w_m)\big)$ is used. We note that if T maps from an n-dimensional vector space to an m-dimensional vector space, then $\mathcal{M}(T)$ is an m-by-n matrix. If T is a linear map from $\mathbb{F}^n$ to $\mathbb{F}^m$, then unless stated otherwise, assume the bases in question are the standard ones.

3.3.2 Addition and Scalar Multiplication of Matrices

In this section, we assume that V and W are finite-dimensional and that a basis has been chosen for each of these vector spaces. Hence, for each linear map from V to W, we can talk about its matrix (w.r.t chosen bases).

Definition 3.3.3: Matrix Addition

*The **sum of two matrices of the same size** is the matrix obtained by adding the corresponding entries in the matrices:*

$$\begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & \vdots & \vdots \\ A_{m,1} & \dots & A_{m,n} \end{bmatrix} + \begin{bmatrix} C_{1,1} & \dots & C_{1,n} \\ \vdots & \vdots & \vdots \\ C_{m,1} & \dots & C_{m,n} \end{bmatrix} = \begin{bmatrix} A_{1,1} + C_{1,1} & \dots & A_{1,n} + C_{1,n} \\ \vdots & \vdots & \vdots \\ A_{m,1} + C_{m,1} & \dots & A_{m,n} + C_{m,n} \end{bmatrix} \tag{3.21}$$

Proposition 3.3.1: The matrix of the sum of Linear Maps

Suppose that we fix the same basis for all three linear maps $S, T, S+T \in \mathcal{L}(V, W)$. Then, $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

Definition 3.3.4: Scalar Multiplication of a Matrix

The product of a scalar and a matrix is the matrix obtained by multiplying each entry in the matrix by the scalar:

$$\lambda \begin{bmatrix} A_{1,1} & \dots & A_{1,n} \\ . & . & . \\ A_{m,1} & \dots & A_{m,n} \end{bmatrix} = \begin{bmatrix} \lambda A_{1,1} & \dots & \lambda A_{1,n} \\ . & . & . \\ \lambda A_{m,1} & \dots & \lambda A_{m,n} \end{bmatrix}. \tag{3.22}$$

In essence, $(\lambda A)_{j,k} = \lambda A_{j,k}$ for any $\lambda \in \mathbb{F}$.

Proposition 3.3.2: The Matrix of a Scalar times a Linear Map

Suppose that $\lambda \in \mathbb{F}$ and $T \in \mathcal{L}(V, W)$. Then, $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Definition 3.3.5: Notation: $\mathbb{F}^{m,n}$

For m and n positive integers, the set of all m-by-n matrices with entries in $\mathbb{F}$ is denoted by $\mathbb{F}^{m,n}$.

Proposition 3.3.3: dim $\mathbb{F}^{m,n} = mn$

Suppose that m and n are positive integers. With addition and scalar multiplication defiend as above, $\mathbb{F}^{m,n}$ is a vector space with dimension mn.

3.3.3 Matrix Multiplication

Definition 3.3.6: Matrix Multiplication

Suppose that $A \in \mathbb{F}^{m,n}$ and $C \in \mathbb{F}^{n,p}$. Then AC is defined to be the m-by-p matrix whose entry in row j, column k, is given by the following equation

$$(AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k} \tag{3.23}$$

In other words, the entry in row j, column k of AC is computed by taking row j of A and column k of C, multiplying together the corresponding entries, and then summing.

Proposition 3.3.4: The Matrix of the Product of Linear Maps

If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$, then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

Definition 3.3.7: Notation $A_{j,\cdot}, A_{\cdot,k}$

Suppose A is an m-by-n matrix.

- *If $1 \leq j \leq m$, then $A_{j,\cdot}$ denotes the 1-by-n matrix consisting of row j of A.*
- *If $1 \leq k \leq n$, then $A_{\cdot,k}$ denotes the m-by-1 matrix consisting of column k of A.*

Proposition 3.3.5: Entry of Matrix Product Equals Row times Column

Suppose that $A \in \mathbb{F}^{m,n}$ and $C \in \mathbb{F}^{n,p}$. Then

$$(AC)_{j,k} = A_{j,\cdot}C_{\cdot,k} \tag{3.24}$$

for $1 \leq j \leq m$ and $1 \leq k \leq p$.

Proposition 3.3.6: Column of Matrix Product equals Matrix times Column

Suppose that $A \in \mathbb{F}^{m,n}$ and $C \in \mathbb{F}^{n,p}$. Then

$$(AC)_{\cdot,k} = AC_{\cdot,k} \tag{3.25}$$

for $1 \leq k \leq p$.

Proposition 3.3.7: Linear Combination of Columns

Suppose that $A \in \mathbb{F}^{m,n}$ and $c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ is an n-by-1 matrix. Then,

$$Ac = c_1 A_{\cdot,1} + \ldots + c_n A_{\cdot,n} \tag{3.26}$$

In essence, Ac is a linear combination of the columns of A, with the scalars that multiply the columns coming from c.

3.4 Invertibility and Isomorphic Vector Spaces

3.4.1 Invertible Linear Maps

Definition 3.4.1: Invertible Linear Map / Inverse

- *A linear map $T \in \mathcal{L}(V, W)$ is called* ***invertible*** *if there exists a linear map $S \in \mathcal{L}(W, V)$ such that ST equals the identity map on V and TS equals the identity map on W.*
- *A linear map $S \in \mathcal{L}(W, V)$ satisfying $ST = I_V$ and $TS = I_W$ is called an* ***inverse*** *of T. Here, I_V denotes the identity map on V and I_W denotes the identity map on W.*

Proposition 3.4.1: Inverse is Unique

An invertible linear map has a unique inverse.

Definition 3.4.2: Notation: T^{-1}

If T is invertible, then its inverse is denoted by T^{-1}. In other words, if $T \in \mathcal{L}(V, W)$ is invertible, then T^{-1} is the unique element of $\mathcal{L}(V, W)$ such that $T^{-1}T = I$ and $TT^{-1} = I$.

Proposition 3.4.2: Invertibility is Equivalent to Injectivity and Surjectivity

A linear map is invertible if and only if it is injective and surjective.

3.4.2 Isomorphic Vector Spaces

Definition 3.4.3: Isomorphism, Isomorphic

- *An **isomorphism** is an invertible linear map.*
- *Two vector spaces are called **isomorphic** if there is an isomorphism from one vector space onto the other one.*

Proposition 3.4.3: Dimension shows whether vector spaces are isomorphic

Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.

Proposition 3.4.4: $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$ are isomorphic

Suppose that $v_1, ..., v_n$ is a basis of V and $w_1, ..., w_m$ is a basis of W. Then $\mathcal{M}$ is an isomorphism between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$.

Proposition 3.4.5: dim $\mathcal{L}(V, W) = (\mathbf{dim} V)(\mathbf{dim} W)$

Suppose that V and W are finite-dimensional. Then $\mathcal{L}(V, W)$ is finite-dimensional and

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W) \tag{3.27}$$

3.4.3 Linear Maps Thought of as Matrix Multiplication

Definition 3.4.4: Matrix of a Vector, $\mathcal{M}(v)$

*Suppose that $v \in V$ and $v_1, ..., v_n$ is a basis of V. The **matrix of** v with respect to this basis is the n-by-1 matrix*

$$\mathcal{M}(v) = \begin{bmatrix} c_1 \\ : \\ c_n \end{bmatrix} \tag{3.28}$$

where $c_1, ..., c_n$ are the scalars such that

$$v = c_1 v_1 + ... + c_n v_n \tag{3.29}$$

Hence, the matrix $\mathcal{M}(v)$ of a vector $v \in V$ depends on the basis $\{v_1, ..., v_n\}$ of V as well as on v.

Proposition 3.4.6: $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(v_k)$

Suppose that $T \in \mathcal{L}(V,W)$ *and* $\{v_1, ..., v_n\}$ *is a basis of* V *and* $w_1, ..., w_m$ *is a basis of* W*. Let* $1 \leq k \leq n$*. Then, the* k^{th} *column of* $\mathcal{M}(T)$*, which is denoted by* $\mathcal{M}(T)_{\cdot,k}$ *equals* $\mathcal{M}(v_k)$*.*

Proposition 3.4.7: Linear Maps act Like Matrix Multiplication

Suppose that $T \in \mathcal{L}(V,W)$ *and* $v \in V$*. Suppose* $v_1, ..., v_n$ *is a basis of* V *and* $w_1, ..., w_m$ *is a basis of* W*. Then*

$$\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v) \tag{3.30}$$

3.4.4 Operators

Definition 3.4.5: Operator, $\mathcal{L}(V)$

- *A linear map from a vector space to itself is called an* ***operator****.*
- *The notation* $\mathcal{L}(V)$ *denotes the set of all operators on* V*. In other words,* $\mathcal{L}(V) = \mathcal{L}(V,V)$*.*

Definition 3.4.6: Injectivity is Equivalent to Surjectivity in Finite Dimensions

Suppose that V *is finite-dimensional and* $T \in \mathcal{L}(V)$*, Then, the following are equivalent:*

(a) T *is invertible;*

(b) T *is injective;*

(c) T *is surjective.*

3.5 Products and Quotients of Vector Spaces

3.5.1 Products of Vector Spaces

Definition 3.5.1: Product of Vector Spaces

Suppose that $V_1, ..., V_m$ *are vector spaces over* $\mathbb{F}$*.*

- *The* ***product*** $V_1 \times ... \times V_m$ *is defined by*

$$V_1 \times ... \times V_m = \{(v_1, ..., v_m) : v_1 \in V, ..., v_m \in V_m\} \tag{3.31}$$

- *Addition on* $V_1 \times ... \times V_m$ *is defined by*

$$(u_1, ..., u_m) + (v_1, ..., v_m) = (u_1 + v_1, ..., u_m + v_m) \tag{3.32}$$

- *Scalar multiplication on* $V_1 \times ... \times V_m$ *is defined by*

$$\lambda(v_1, ..., v_m) = (\lambda v_1, ..., \lambda v_m) \tag{3.33}$$

Proposition 3.5.1: Product of Vector Spaces is a Vector Space

Suppose that $V_1, ..., V_m$ are vector spaces over $\mathbb{F}$. Then $V_1 \times ... \times V_m$ is a vector space over $\mathbb{F}$.

Proposition 3.5.2: Dimension of a Product is the Sum of Dimensions

Suppose that $V_1, ..., V_m$ are finite-dimensional vector spaces. Then, $V_1 \times ... \times V_m$ is finite-dimensional and

$$\dim(V_1 \times ... \times V_m) = \dim V_1 + ... + \dim V_m \tag{3.34}$$

3.5.2 Products and Direct Sums

Proposition 3.5.3: Products and Direct Sums

Suppose that $U_1, ..., U_m$ are subspaces of V. Define a linear map $\Gamma : U_1 \times ... \times U_m \to U_1 + ... + U_m$ by

$$\Gamma(u_1, ..., u_m) = u_1 + ... + u_m. \tag{3.35}$$

Then $U_1 + ... + U_m$ is a direct sum if and only if Γ is injective.

Proposition 3.5.4: A Sum is a Direct Sum if and only if Dimensions add up

Suppose V is finite-dimensional and $U_1, ..., U_m$ are subspaces of V. Then $U_1 + ... + U_m$ is a direct sum if and only if

$$\dim(U_1 + ... + U_m) = \dim U_1 + ... + \dim U_m \tag{3.36}$$

3.5.3 Quotients of Vector Spaces

Definition 3.5.2: Notation: $v + U$

Suppose that $v \in V$ and U is a subspace of V. Then, $v + U$ is the subset of V defined by

$$v + U = \{v + u : u \in U\} \tag{3.37}$$

In the mathematical literature, this object is given the name of a left coset. Since, addition is abelian for the cases of U that we are considering, left or right coset need not be specified.

Example 3.5.1: Parallel Lines with Cosets

Suppose that

$$U = \{(x, 2x) \in \mathbb{R}^2 : x \in \mathbb{R}\} \tag{3.38}$$

Then U is the line in $\mathbb{R}^2$ through the origin with slope 2. Thus,

$$(17, 20) + U \tag{3.39}$$

is the line in $\mathbb{R}^2$ that contains the point $(17, 20)$ and has slope 2.

Definition 3.5.3: Affine Subset, Parallel

- *An* ***affine subset*** *of V is a subset of V of the form $v + U$ for some $v \in V$ and some subspace U of V.*
- *For $v \in V$ and U a subspace of V, the affine subset $v + U$ is said to be* ***parallel*** *to U.*

Definition 3.5.4: Quotient Space, V/U

Suppose U is a subspace of V. Then the ***quotient space*** *V/U is the set of all affine subsets of V parallel to U. In other words,*

$$V/U = \{v + U : v \in V\} \tag{3.40}$$

Example 3.5.2: Quotient Space: Set of all Lines

If $U = \{(x, 2x) \in \mathbb{R}^2 : x \in \mathbb{R}\}$, then $\mathbb{R}^2/U$ is the set of all lines in $\mathbb{R}^2$ that have slope 2.

Proposition 3.5.5: Two Affine Subsets Parallel to U are Equal or Disjoint

Suppose that U is a subspace of V and $v, w \in V$. Then, the following are equivalent:

(a) $v - w \in U$;

(b) $v + U = w + U$;

(c) $(v + U) \cap (w + U) \neq \emptyset$

Definition 3.5.5: Addition and Scalar Multiplication on V/U

Suppose that U is a subspace of V. Then, addition and scalar multiplication are defined on V/U by

$$(v + U) + (w + U) = (v + w) + U \tag{3.41}$$

$$\lambda(v + U) = (\lambda v) + U \tag{3.42}$$

for $v, w \in V$ and $\lambda \in \mathbb{F}$.

Theorem 3.5.1: Quotient Space is a Vector Space

Suppose that U is a subspace of V. Then, V/U, with the operations of addition and scalar multiplication as defined above, is a vector space.

Definition 3.5.6: Quotient map, π

Suppose U is a subspace of V. The ***quotient map*** *π is the linear map $\pi : V \to V/U$ defined by*

$$\pi(v) = v + U \tag{3.43}$$

for $v \in V$.

Proposition 3.5.6: Dimension of a Quotient Space

Suppose that V is finite-dimensional and U is a subspace of V. Then

$$\dim V/U = \dim V - \dim U \tag{3.44}$$

Definition 3.5.7: $\tilde{T}$

Suppose that $T \in \mathcal{L}(V, W)$. Define $\tilde{T} : V/(\text{null } T) \to W$ by

$$\tilde{T}(v + \text{null } T) = Tv \tag{3.45}$$

Since coset representations are not unique, we need to ensure that this map is well defined. To that end, suppose that $v, w \in V$ such that

$$v + \text{null } T = w + \text{null } T. \tag{3.46}$$

Then, we want $Tv = Tw$ or this map doesn't make sense. Since $v + \text{null } T = w + \text{null } T$, then by Proposition 3.5.5 we have $v - w \in \text{null } T$. Hence, we have $T(v - w) = 0$, from which we conclude that $T(v) = T(w)$. Thus, the definition of $\tilde{T}$ makes sense.

Proposition 3.5.7: Null Space and Range of $\tilde{T}$

Suppose that $T \in \mathcal{L}(V, W)$. Then

(a) *$\tilde{T}$ is a linear map from $V/(\text{null } T)$ to W;*

(b) *$\tilde{T}$ is injective;*

(c) *range $\tilde{T}$ = range T;*

(d) *$V/(\text{null } T)$ is isomorphic to range T.*

3.6 Duality

3.6.1 The Dual Space and the Dual Map

Definition 3.6.1: Linear Functional

*A **linear functional** on V is a linear map from V to $\mathbb{F}$. In other words, a linear functional is an element of $\mathcal{L}(V, \mathbb{F})$.*

Definition 3.6.2: Dual Space, V'

*The **dual space** of V, denoted V', is the vector space of all linear functionals on V. In other words, $V' = \mathcal{L}(V, \mathbb{F})$.*

Proposition 3.6.1: dim V' = dim V

Suppose that V is finite-dimensional. Then V' is also finite-dimensional and $\dim V' = \dim V$.

Definition 3.6.3: Dual Basis

*If $v_1, ..., v_n$ is a basis for V, then the **dual basis** of $v_1, .., v_n$ is the list $\psi_1, ..., \psi_n$ of elements of V', where each ψ_j is the linear functional on V such that*

$$\psi_j(v_k) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases} \tag{3.47}$$

Proposition 3.6.2: Dual Basis is a Basis of the Dual Space

Suppose V is finite-dimensional. Then the dual basis of a basis of V is a basis of V'.

Definition 3.6.4: Dual Map, T'

*If $T \in \mathcal{L}(V, W)$, then the **dual map** of T is the linear map $T' \in \mathcal{L}(W', V')$ defined by $T'(\psi) = \psi \circ T$ for $\psi \in W'$.*

Proposition 3.6.3: Algebraic Properties of Dual Maps

Let U, V, W be vector spaces.

- $(S + T)' = S' + T'$ *for all* $S, T \in \mathcal{L}(V, W)$.
- $(\lambda T)' = \lambda T'$ *for all* $\lambda \in \mathbb{F}$ *and all* $T \in \mathcal{L}(V, W)$.
- $(ST)' = T'S'$ *for all* $T \in \mathcal{L}(U, V)$ *and all* $S \in \mathcal{L}(V, W)$

Some books use the notation V^* and T^* for duality instead of V' and T'. However, this notation will be reserved for the adjoint.

3.6.2 The Null Space and Range of the Dual of a Linear Map

Definition 3.6.5: Annihilator, U^0

*For $U \subset V$, the **annihilator** of U, denoted U^0, is defined by*

$$U^0 = \{\psi \in V' : \psi(u) = 0 \text{ for all } u \in U\} \tag{3.48}$$

Proposition 3.6.4: The Annihilator is a Subspace

Suppose that $U \subset V$. Then U^0 is a subspace of V'.

Proposition 3.6.5: Dimension of the Annihilator

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U + \dim U^0 = \dim V \tag{3.49}$$

Proposition 3.6.6: The Null Space of T'

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) *null $T' = (\text{range } T)^0$;*

(b) *dim null $T' =$ dim null T + dim W - dim V.*

Proposition 3.6.7: T surjective is equivalent to T' injective

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is surjective if and only if T' is injective.

Proposition 3.6.8: The Range of T'

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then

(a) *dim range $T' =$ dim range T;*

(b) *range $T' = (\text{null} T)^0$.*

Proposition 3.6.9: T injective is equivalent to T' surjective

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then T is injective if and only if T' is surjective.

3.6.3 The Matrix of the Dual of a Linear Map

Definition 3.6.6: Transpose, A^t

*The **transpose** of a matrix A, denoted A^t, is the matrix obtained from A by interchanging the rows and columns. More specifically, if A is an m-by-n matrix, then A^t is the n-by-m matrix whose entries are given by the equation*

$$(A^t)_{k,j} = A_{j,k} \tag{3.50}$$

Proposition 3.6.10: The Transpose of the Product of Matrices

If A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^t = C^t A^t \tag{3.51}$$

Theorem 3.6.1: The Matrix of T' is the Transpose of the Matrix of T

Suppose that $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(T') = (\mathcal{M}(T))^t$.

3.6.4 The Rank of a Matrix

Definition 3.6.7: Row Rank, Column Rank

Suppose that A is an m-by-n matrix with entries in $\mathbb{F}$.

- *The **row rank** of A is the dimension of the span of the rows of A in $\mathbb{F}^{1,n}$.*
- *The **column rank** of A is the dimension of the span of the columns of A in $\mathbb{F}^{m,1}$.*

Proposition 3.6.11: Dimension of range T equals column rank of $\mathcal{M}(T)$

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Then dim range T equals the column rank of $\mathcal{M}(T)$.

Proposition 3.6.12: Row Rank equals Column Rank

Suppose $A \in \mathbb{F}^{m,n}$. Then the row rank of A equals the column rank of A.

Definition 3.6.8: Rank

*The **rank** of a matrix $A \in \mathbb{F}^{m,n}$ is the column rank of A.*

Chapter 4

Polynomials

4.1 Polynomials

4.1.1 Complex Conjugate and Absolute Value

Definition 4.1.1: Re z, Im z

Suppose that $z = a + bi$, where $a, b \in \mathbb{R}$.

- *The **real part** of z, denoted Re z, is defined by Re $z = a$.*
- *The **imaginary part** of z, denoted Im z, is defined by Im $z = b$.*

Hence, for every complex number z, we have

$$z = Re\ z + (Im\ z)i \tag{4.1}$$

Definition 4.1.2: Complex Conjugate, $\bar{z}$

*Let $z \in \mathbb{C}$. The **complex conjugate** of $z \in \mathbb{C}$, denoted $\bar{z}$, is defined by*

$$\bar{z} = Re\ z - (Im\ z)i \tag{4.2}$$

Definition 4.1.3: Absolute Value, $|z|$

*Let $z \in \mathbb{C}$. The **absolute value** of $z \in \mathbb{C}$, denoted $|z|$, is defined by*

$$|z| = \sqrt{(Re\ z)^2 + (Im\ z)^2} \tag{4.3}$$

Proposition 4.1.1: Properties of Complex Numbers

Suppose that $w, z \in \mathbb{C}$. Then

1. ***Sum of z and $\bar{z}$***
$$z + \bar{z} = 2Re\ z; \tag{4.4}$$
2. ***Difference of z and $\bar{z}$***
$$z - \bar{z} = 2(Im\ z)i; \tag{4.5}$$

3. ***Product of z and $\bar{z}$***

$$z\bar{z} = |z|^2 \tag{4.6}$$

4. ***Additivity and Multiplicativity of Complex Conjugate***

$$\overline{w+z} = \bar{w} + \bar{z} \text{ and } \overline{wz} = \bar{w}\bar{z}; \tag{4.7}$$

5. ***Conjugate of Conjugate***

$$\bar{\bar{z}} = z; \tag{4.8}$$

6. ***Real and Imaginary Parts are Bounded by $|z|$***

$$|Re\ z| \leq |z| \text{ and } |Im\ z| \leq |z| \tag{4.9}$$

7. ***Absolute Value of the Complex Conjugate***

$$|\bar{z}| = |z| \tag{4.10}$$

8. ***Multiplicativity of Absolute Value***

$$|wz| = |w||z|; \tag{4.11}$$

9. ***Triangle Inequality***

$$|w+z| \leq |w| + |z| \tag{4.12}$$

4.1.2 Uniqueness of Coefficients for Polynomials

Proposition 4.1.2: If a Polynomial is the Zero Function, then all Coefficients are 0

Suppose $a_0, ..., a_m \in \mathbb{F}$. If

$$a_0 + a_1 z + ... + a_m z^m = 0 \tag{4.13}$$

for every $z \in \mathbb{F}$, then $a_0 = ... = a_m = 0$.

4.1.3 The Division Algorithm for Polynomials

Proposition 4.1.3: Division Algorithm for Polynomials

Suppose that $p, s \in \mathcal{P}(\mathbb{F})$, with $s \neq 0$. Then there exist unique polynomials $q, r \in \mathcal{P}(\mathbb{F})$ such that

$$p = sq + r \tag{4.14}$$

and deg $r <$ deg s.

4.1.4 Zeros of Polynomials

Definition 4.1.4: Zero of a Polynomial

*A number $\lambda \in \mathbb{F}$ is called a **zero** (or **root**) of a polynomial $p \in \mathcal{P}(\mathbb{F})$ if*

$$p(\lambda) = 0 \tag{4.15}$$

Definition 4.1.5: Factor

*A polynomial $s \in \mathcal{P}(\mathbb{F})$ is called a **factor** of $p \in \mathcal{P}(\mathbb{F})$ if there exists a polynomial $q \in \mathcal{P}(\mathbb{F})$ such that $p = sq$.*

Proposition 4.1.4: Each Zero of a Polynomial Corresponds to a Degree-1 Factor

Suppose $p \in \mathcal{P}(\mathbb{F})$ and $\lambda \in \mathbb{F}$. Then $p(\lambda) = 0$ if and only if there is a polynomial $q \in \mathcal{P}(\mathbb{F})$ such that

$$p(z) = (z - \lambda)q(z) \tag{4.16}$$

for every $z \in \mathbb{F}$.

Proposition 4.1.5: A Polynomial has At Most as Many Zeros as its Degree

Suppose $p \in \mathcal{P}(\mathbb{F})$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct zero in $\mathbb{F}$.

4.1.5 Factorization of Polynomials over $\mathbb{C}$

Theorem 4.1.1: Fundamental Theorem of Algebra

Every nonconstant polynomial with complex coefficients has a zero.

Proposition 4.1.6: Factorization of a Polynomial over $\mathbb{C}$

If $p \in \mathcal{P}(\mathbb{C})$ is nonconstant polynomial, then p has a unique factorization (except for the order of the factors) of the form

$$p(z) = c(z - \lambda_1)...(z - \lambda_m) \tag{4.17}$$

where $c, \lambda_1, ..., \lambda_m \in \mathbb{C}$.

4.1.6 Factorization of Polynomials over $\mathbb{R}$

Proposition 4.1.7: Polynomials with Real Coefficients have Zeros in Pairs

Suppose $p \in \mathcal{P}(\mathbb{C})$ is a polynomial with real coefficients. If $\lambda \in \mathbb{C}$ is a zero of p, then so is $\bar{\lambda}$.

Proposition 4.1.8: Factorization of a Quadratic Polynomial

Suppose $b, c \in \mathbb{R}$. *Then there is a polynomial factorization of the form*

$$x^2 + bx + c = (x - \lambda_1)(x - \lambda_2) \tag{4.18}$$

with $\lambda_1, \lambda_2 \in \mathbb{R}$ *if and only if* $b^2 \geq 4c$.

Proposition 4.1.9: Factorization of Polynomial over $\mathbb{R}$

Suppose $p \in \mathcal{P}(\mathbb{R})$ *is a nonconstant polynomial. Then* p *has a unique factorization (except for the order of the factors) of the form*

$$p(x) = c(x - \lambda_1)...(x - \lambda_m)(x^2 + b_1 x + c_1)...(x^2 + b_M x + c_M) \tag{4.19}$$

where $c, \lambda_1, ..., \lambda_m, b_1, ..., b_M, c_1, ..., c_M \in \mathbb{R}$, *with* $b_j^2 < 4c_j$ *for each* j.

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

5.1 Invariant Subspaces

Definition 5.1.1: Invariant Subspace

Suppose that $T \in \mathcal{L}(V)$. A subspace U of V is called ***invariant*** *under T if $u \in U$ implies $Tu \in U$.*

In essence, U is invariant under T if $T|_U$ is an operator on U.

5.1.1 Eigenvalues and Eigenvectors

Definition 5.1.2: Eigenvalue

Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbb{F}$ is called an ***eigenvalue*** *of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.*

Proposition 5.1.1: Equivalent Conditions to be an Eigenvalue

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Then the following are equivalent:

(a) *λ is an eigenvalue of T;*

(b) *$T - \lambda I$ is not injective;*

(c) *$T - \lambda I$ is not surjective;*

(d) *$T - \lambda I$ is not invertible.*

Definition 5.1.3: Eigenvector

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of T. A vector $v \in V$ is called an ***eigenvector*** *of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.*

Corollary 5.1.1

Let $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ be an eigenvalue of T. A vector $v \in V$ such that $v \neq 0$ is an eigenvector of T corresponding to λ if and only if $v \in null(T - \lambda I)$.

Proposition 5.1.2: Linearly Independent Eigenvectors

Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, ..., \lambda_m$ are distinct eigenvalues of T and $v_1, ..., v_m$ are corresponding eigenvectors. Then $v_1, ..., v_m$ is linearly independent.

Proposition 5.1.3: Number of Eigenvalues

Suppose V is finite-dimensional. Then each operator on V has at most dim V distinct eigenvalues.

5.1.2 Restriction and Quotient Operators

Definition 5.1.4: $T|_U$ and T/U

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T.

- *The **restriction operator** $T|_U \in \mathcal{L}(U)$ is defined by*

$$T|_U(u) = Tu \tag{5.1}$$

for $u \in U$.

- *The **quotient operator** $T/U \in \mathcal{L}(V/U)$ is defined by*

$$(T/U)(v + U) = Tv + U \tag{5.2}$$

for $v \in V$.

5.2 Eigenvectors and Upper-Triangular Matrices

5.2.1 Polynomials Applied to Operators

Definition 5.2.1: T^m

Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

- *T^m is defined by*

$$T^m = \underbrace{T \cdots T}_{m \text{ times}}. \tag{5.3}$$

- *T^0 is defined to be the identity operator I on V.*
- *If T is invertible with inverse T^{-1}, then T^{-m} is defined by*

$$T^{-m} = (T^{-1})^m \tag{5.4}$$

Definition 5.2.2: $p(T)$

Suppose $T \in \mathcal{L}(V)$ *and* $p \in \mathcal{P}(\mathbb{F})$ *is a polynomial given by*

$$p(z) = a_0 + a_1 z + ... + a_m z^m \tag{5.5}$$

for $z \in \mathbb{F}$*. Then* $p(T)$ *is the operator defined by*

$$p(T) = a_0 I + a_1 T + a_2 T^2 + ... + a_m T^m \tag{5.6}$$

Corollary 5.2.1

Let $T \in \mathcal{L}(V)$*, then the function* $P_T : \mathcal{P}(\mathbb{F}) \to \mathcal{L}(V)$ *given by*

$$P_T(p) = p(T) \tag{5.7}$$

is linear.

Definition 5.2.3: Product of Polynomials

If $p, q \in \mathcal{P}(\mathbb{F})$*, then* $pq \in \mathcal{P}(\mathbb{F})$ *is the polynomial defined by*

$$(pq)(z) = p(z)q(z) \tag{5.8}$$

for $z \in \mathbb{F}$*.*

Proposition 5.2.1: Multiplicative Properties

Suppose $p, q \in \mathcal{P}(\mathbb{F})$ *and* $T \in \mathcal{L}(V)$*. Then*

(a) $(pq)(T) = p(T)q(T)$;

(b) $p(T)q(T) = q(T)p(T)$.

5.2.2 Existence of Eigenvalues

Proposition 5.2.2: Operators on Complex Vector Spaces have an Eigenvalue

Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

5.2.3 Upper-Triangular Matrices

Definition 5.2.4: Matrix of an Operator, $\mathcal{M}(T)$

Suppose $T \in \mathcal{L}(V)$ *and* $v_1, ..., v_n$ *is a basis for* V*. The* ***matrix of*** T *with respect to this basis is the* n*-by-*n *matrix*

$$\mathcal{M}(T) = \begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{bmatrix} \tag{5.9}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + ... + A_{n,k}v_n \tag{5.10}$$

If the basis is not clear from the context, then the notation $\mathcal{M}\big(T,(v_1,...,v_n)\big)$.

Definition 5.2.5: Diagonal of a Matrix

*The **diagonal** of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner.*

Definition 5.2.6: Upper-Triangular Matrix

*A matrix is called **upper-triangular** if all the entries below the diagonal equal 0.*

Proposition 5.2.3: Conditions for Upper-Triangular Matrix

Suppose $T \in \mathcal{L}(V)$ and $v_1,...,v_n$ is a basis of V. Then the following are equivalent:

(a) *the matrix of T with respect to $v_1,...,v_n$ is upper triangular;*

(b) *$Tv_j \in \text{span}(v_1,...,v_j)$ for each $j = 1,...,n$.*

(c) *$\text{span}(v_1,...,v_j)$ is invariant under T for each $j = 1,...,n$.*

Proposition 5.2.4: Over $\mathbb{C}$, every operator has an upper-triangular matrix

Suppose V is finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

Proposition 5.2.5: Determination of Invertibility from Upper-Triangular Matrix

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.

Proposition 5.2.6: Determination of Eigenvalues from Upper-Triangular Matrix

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

5.3 Eigenspaces and Diagonal Matrices

Definition 5.3.1: Diagonal Matrix

*A **diagonal matrix** is a square matrix that is 0 everywhere except possibly along the diagonal.*

Definition 5.3.2: Eigenspace, $E(\lambda, T)$

Suppose that $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. The ***eigenspace*** *of T corresponding to λ, denoted $E(\lambda, T)$, is defined by*

$$E(\lambda, T) = null(T - \lambda I). \tag{5.11}$$

In other words, $E(\lambda, T)$ is the set of all eigenvectors of T corresponding to λ, along with the 0 vector.

Proposition 5.3.1: Sum of Eigenspaces is a Direct Sum

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose also that $\lambda_1, ..., \lambda_m$ are distinct eigenvalues of T. Then

$$E(\lambda_1, T) + ... + E(\lambda_m, T) \tag{5.12}$$

is a direct sum. Furthermore,

$$\dim E(\lambda_1, T) + ... + \dim E(\lambda_m, T) \leq \dim V. \tag{5.13}$$

Definition 5.3.3: Diagonalizable

An operator $T \in \mathcal{L}(V)$ is called ***diagonalizable*** *if the operator has a diagonal matrix with respect to some basis of V.*

Theorem 5.3.1: Conditions Equivalent to Diagonalizability

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:

(a) *T is diagonalizable;*

(b) *V has a basis consisting of eigenvectors of T;*

(c) *There exists 1-dimensional subspaces $U_1, ..., U_n$ of V, each invariant under T, such that*

$$V = U_1 \oplus \cdots \oplus U_n \tag{5.14}$$

(d) *$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$;*

(e) *$\dim V = \dim E(\lambda_1, T) + \cdots \dim E(\lambda_m, T)$.*

Proposition 5.3.2: Enough Eigenvalues Implies Diagonalizability

If $T \in \mathcal{L}(V)$ has $\dim V$ distinct eigenvalues, then T is diagonalizable.

Chapter 6

Inner Product Spaces

6.1 Inner Products and Norms

6.1.1 Inner Products

Definition 6.1.1: Dot Product

Let $x, y \in \mathbb{R}^n$. The ***dot product*** *of x and y, denoted $x \cdot y$, is defined by*

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n \tag{6.1}$$

where $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$.

Definition 6.1.2: Inner Product

An ***inner product*** *on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v \rangle \in \mathbb{F}$ and has the following properties:*

- ***Positivity***[a]
$$\langle v, v \rangle \geq 0 \text{ for all } v \in V \tag{6.2}$$
- ***Definiteness***
$$\langle v, v \rangle = 0 \text{ if and only if } v = 0 \tag{6.3}$$
- ***Additivity in First Slot***
$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \text{ for all } u, v, w \in V \tag{6.4}$$
- ***Homogeneity in First Slot***
$$\langle \lambda u, v \rangle = \lambda \langle u, v \rangle \text{ for all } \lambda \in \mathbb{F} \text{ and all } u, v \in V \tag{6.5}$$
- ***Conjugate Symmetry***
$$\langle u, v \rangle = \overline{\langle v, u \rangle} \text{ for all } u, v \in V \tag{6.6}$$

It should be noted that although most mathematicians define an inner product as above, many physicists use a definition requiring homogeneity in the second slot instead of the first slot.

[a] If $\lambda \in \mathbb{C}$, then the notation $\lambda \geq 0$ means that λ is real and nonnegative.

Example 6.1.1: Euclidean Inner Product

*The **Euclidean Inner product** on $\mathbb{F}^n$ is defined by*

$$\big\langle (w_1, ..., w_n), (z_1, ..., z_n) \big\rangle = w_1 \overline{z_1} + \cdots w_n \overline{z_n} \tag{6.7}$$

Definition 6.1.3: Inner Product Space

*An **inner product space** is a vector space V along with an inner product on V.*

For the rest of the chapter, V denotes an inner product space over $\mathbb{F}$.

Proposition 6.1.1: Basic Properties of an Inner Product

(a) *For each fixed $u \in V$, the function that takes v to $\big\langle v, u \big\rangle$ is a linear map from V to $\mathbb{F}$.*

(b) $\big\langle 0, u \big\rangle = 0$ *for every* $u \in V$.

(c) $\big\langle u, 0 \big\rangle = 0$ *for every* $u \in V$.

(d) $\big\langle u, v + w \big\rangle = \big\langle u, v \big\rangle + \big\langle u, w \big\rangle$ *for all* $u, v, w \in V$.

(e) $\big\langle u, \lambda v \big\rangle = \overline{\lambda} \big\langle u, v \big\rangle$ *for all* $\lambda \in \mathbb{F}$. *and* $u, v \in V$.

6.1.2 Norms

Definition 6.1.4: Norm, $||v||$

*For $v \in V$, the **norm** of v, denoted $||v||$, is defined by*

$$||v|| = \sqrt{\big\langle v, v \big\rangle} \tag{6.8}$$

Proposition 6.1.2: Basic Properties of the Norm

Suppose that $v \in V$.

(a) $||v|| = 0$ *if and only if* $v = 0$.

(b) $||\lambda v|| = |\lambda| \; ||v||$ *for all* $\lambda \in \mathbb{F}$.[a]

[a] Since we typically consider $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, $|\lambda|$ denotes the magnitude of the complex number.

Definition 6.1.5: Orthogonal

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

Proposition 6.1.3: Orthogonality and 0

(a) 0 *is orthogonal to every vector in* V.

(b) 0 *is the only vector in* V *that is orthogonal to itself.*

Proposition 6.1.4: Pythagorean Theorem

Suppose that u, v *are orthogonal vectors in* V. *Then*

$$||u+v||^2 = ||u||^2 + ||v||^2 \tag{6.9}$$

Proposition 6.1.5: Orthogonal Decomposition

Suppose $u, v \in V$, *with* $v \neq 0$. *Set* $c = \frac{<u,v>}{||v||^2}$ *and* $w = u - \frac{<u,v>}{||v||^2}v$. *Then*

$$\langle w, v\rangle = 0 \text{ and } u = cv + w \tag{6.10}$$

Theorem 6.1.1: Cauchy-Schwartz Inequality

Suppose $u, v \in V$. *Then*

$$|\langle u, v\rangle| \leq ||u||\ ||v|| \tag{6.11}$$

This inequality is an equality if and only if one of u, v *is a scalar multiple of the other.*

Proof. Let $u, v \in V$ such that $u, v \neq 0$. By the definition of an inner product, we have that

$$\left\langle u - \frac{\langle u, v\rangle}{||v||^2}v, u - \frac{\langle u, v\rangle}{||v||^2}v \right\rangle \geq 0. \tag{6.12}$$

It follows from the additive properties of inner products that

$$0 \leq \langle u, u\rangle - \frac{\langle u, v\rangle\langle v, u\rangle}{||v||^2}. \tag{6.13}$$

Hence,

$$\langle u, v\rangle\langle v, u\rangle \leq ||u||^2||v||^2, \tag{6.14}$$

from which one finally takes a square root to achieve the desired result. □

Proposition 6.1.6: Triangle Inequality

Suppose $u, v \in V$. *Then*

$$||u+v|| \leq ||u|| + ||v||. \tag{6.15}$$

This inequality is an equality if and only if one of u, v *is a nonnegative multiple of the other.*

Proof. Let $u, v \in V$. Then we observe that

$$\begin{aligned} ||u+v||^2 &= \langle u+v, u+v \rangle \\ &= \langle u,u \rangle + \langle u,v \rangle + \langle v,u \rangle + \langle v,v \rangle \\ &= ||u||^2 + ||v||^2 + 2\mathrm{Re}[\langle u,v \rangle] \\ &\leq ||u||^2 + ||v||^2 + 2|\langle u,v \rangle| \\ &\leq ||u||^2 + ||v||^2 + 2||u||\ ||v|| \\ &= (||u|| + ||v||)^2 \end{aligned} \tag{6.16}$$

Hence, it follows that $||u+v|| \leq ||u|| + ||v||$. In the first inequality, we used the fact that $\mathrm{Re}(z) \leq |z|$ for all $z \in \mathbb{C}$. The second inequality follows from the Cauchy-Schwartz Inequality. ☐

Proposition 6.1.7: Parallelogram Equality

Suppose $u, v \in V$. Then

$$||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2) \tag{6.17}$$

6.2 Orthonormal Bases

Definition 6.2.1: Orthonormal

- *A list of vectors is called* ***orthonormal*** *if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.*
- *In essence, a list $e_1, ..., e_m$ of vectors in V is orthonormal if*

$$\langle e_i, e_j \rangle = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \tag{6.18}$$

Proposition 6.2.1: The Norm of an Orthonormal Linear Combination

If $e_1, ..., e_m$ is an orthonormal list of vectors in V, then

$$||a_1e_1 + \cdots + a_me_m||^2 = |a_1|^2 + \cdots + |a_m|^2 \tag{6.19}$$

for all $a_1, ..., a_m \in \mathbb{F}$.

Corollary 6.2.1: An Orthonormal List is Linearly Independent

Every orthonormal list of vectors is linearly independent.

Definition 6.2.2: Orthonormal Basis

An ***orthonormal basis*** *of V is an orthonormal list of vectors in V that is also a basis of V.*

Proposition 6.2.2: An Orthonormal List of the Right Length is an Orthonormal Basis

Every orthonormal list of vectors in V with length $\dim V$ is an orthonormal basis of V.

Proposition 6.2.3: Writing a Vector as Linear Combination of Orthonormal Basis

Suppose $e_1, ..., e_n$ is an orthonormal basis of V and $v \in V$. Then

$$v = \langle v, e_1 \rangle e_1 + \cdots \langle v, e_n \rangle e_n \tag{6.20}$$

and

$$||v||^2 + |\langle v, e_1 \rangle|^2 + \cdots |\langle v, e_n \rangle|^2 \tag{6.21}$$

Proposition 6.2.4: Gram-Schmidt Procedure

Suppose that $v_1, ..., v_m$ is a linearly independent list of vectors in V. Let $e_1 = v_1/||v_1||$. For $j = 2, ..., m$ define e_j inductively by

$$e_j = \frac{v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}}{||v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}||}. \tag{6.22}$$

Then $e_1, ..., e_m$ is an orthonormal list of vectors in V such that

$$span(v_1, ..., v_j) = span(e_1, ..., e_j) \tag{6.23}$$

for $j = 1, ..., m$.

Proposition 6.2.5: Existence of Orthonormal Basis

Every finite-dimensional inner product space has an orthonormal basis.

Proposition 6.2.6: Orthonormal List Extends to Orthonormal Basis

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proposition 6.2.7: Upper-Triangular Matrix with respect to Orthonormal Basis

Suppose $T \in \mathcal{L}(V)$. If T has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Theorem 6.2.1: Schur's Theorem

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

6.2.1 Linear Functionals on Inner Product Spaces

Theorem 6.2.2: Riesz Representation Theorem

Suppose V is finite-dimensional and ψ is a linear functional on V. Then there is a unique vector $u \in V$ such that

$$\psi(v) = \langle v, u \rangle \tag{6.24}$$

for every $v \in V$.

6.3 Orthogonal Complements and Minimization Problems

6.3.1 Orthogonal Complements

Definition 6.3.1: Orthogonal Complement, $U^\perp$

*If U is a subset of V, then the **orthogonal complement**of U, denoted $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:*

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 \text{ for every } u \in U\} \tag{6.25}$$

Proposition 6.3.1: Basic Properties of Orthogonal Complement

Consider the inner product space V.

(a) *If U is a subset of V, then $U^\perp$ is a subspace of V.*

(b) $\{0\}^\perp = V$.

(c) $V^\perp = \{0\}$.

(d) *If U is a subset of V, then $U \cap U^\perp \subset \{0\}$.*

(e) *If U and W are subsets of V and $U \subset W$, then $W^\perp \subset U^\perp$.*

Proposition 6.3.2: Direct Sum of a Subspace and its Orthogonal Complement

Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^\perp \tag{6.26}$$

Proposition 6.3.3: The Orthogonal Complement of the Orthogonal Complement

Suppose U is a finite-dimensional subspace of V. Then

$$U = (U^\perp)^\perp \tag{6.27}$$

Definition 6.3.2: Orthogonal Projection, P_U

Suppose U is a finite-dimensional subspace of V. The ***orthogonal projection*** *of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: For every $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then $P_U v = u$.*

Proposition 6.3.4: Properties of the Orthogonal Projection P_U

Suppose U is a finite-dimensional subspace of V and $v \in V$. Then

(a) $P_U \in \mathcal{L}(V)$;

(b) $P_U u = u$ *for every* $u \in U$;

(c) $P_U w = 0$ *for every* $w \in U^\perp$;

(d) *range* $P_U = U$;

(e) *null* $P_U = U^\perp$;

(f) $v - P_U v \in U^\perp$;

(g) $(P_U)^2 = P_U$;

(h) $||P_U v|| \leq ||v||$;

(i) *For every orthonormal basis* $e_1, ..., e_m$ *of* U,

$$P_U v = \langle v, e_1 \rangle e_1 + \cdots \langle v, e_m \rangle e_m. \tag{6.28}$$

6.3.2 Minimization Problems

Proposition 6.3.5: Minimizing the Distance to a Subspace

Suppose that U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$||v - P_U v|| \leq ||v - u||. \tag{6.29}$$

Furthermore, the inequality above is an equality of and only if $u = P_U v$.

Chapter 7

Operators on Inner Product Spaces

7.1 Self-Adjoint and Normal Operators

7.1.1 Adjoints

Definition 7.1.1: Adjoint, T^*

*Suppose $T \in \mathcal{L}(V, W)$. The **adjoint** of T is the function $T^* : W \to V$ such that*

$$\langle Tv, w \rangle = \langle v, T^* w \rangle \tag{7.1}$$

for every $v \in V$ and every $w \in W$.

Lemma 7.1.1: Inner Product Property

Let V be an inner product space and let $x, y \in V$. If $\langle v, x \rangle = \langle v, y \rangle$ for all $v \in V$, then $x = y$.

> *Proof.* If $\langle v, x \rangle = \langle v, y \rangle$ for all $v \in V$, then
>
> $$\langle v, x - y \rangle = 0 \ \forall v \in V \tag{7.2}$$
>
> We note that if $\langle v, z \rangle = 0$ for all $v \in V$, then $z = 0$. One can see this from the fact that $V^\perp = \{0\}$. Hence, from our relation, we have that $x - y = 0$, concluding that $x = y$. □

Proposition 7.1.1: The Adjoint is a Linear Map

If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.

Proposition 7.1.2: Properties of The Adjoint

Let U, V, W be inner product spaces over $\mathbb{F}$.

(a) $(S + T)^* = S^* + T^*$ *for all* $S, T \in \mathcal{L}(V, W)$;

(b) $(\lambda T)^* = \overline{\lambda} T^*$ *for all* $\lambda \in \mathbb{F}$ *and* $T \in \mathcal{L}(V, W)$

(c) $(T^*)^* = T$ *for all* $T \in \mathcal{L}(V, W)$;

(d) $I^* = I$, *where* I *is the identity operator on* V;

(e) $(ST)^* = T^* S^*$ *for all* $T \in \mathcal{L}(V, W)$ *and* $S \in \mathcal{L}(W, U)$

Proposition 7.1.3: Null Space and Range of T^*

Suppose that $T \in \mathcal{L}(V, W)$. *Then*

(a) *null* $T^* = (\text{range } T)^\perp$;

(b) *range* $T^* = (\text{null } T)^\perp$;

(c) *null* $T = (\text{range } T^*)^\perp$;

(d) *range* $T = (\text{null } T^*)^\perp$.

Definition 7.1.2: Conjugate Transpose

The ***conjugate transpose*** *of an* m*-by-*n *matrix is the* n*-by-*m *matrix obtained by interchanging the rows and columns and then taking the complex conjugate of each entry.*

Proposition 7.1.4: The Matrix of T^*

Let $T \in \mathcal{L}(V, W)$. *Suppose* $e_1, ..., e_n$ *is an orthonormal basis of* V *and* $f_1, ..., f_m$ *is an orthonormal basis of* W. *Then*

$$\mathcal{M}(T^*, (f_1, ..., f_m), (e_1, ..., e_n)) \tag{7.3}$$

is the conjugate transpose of

$$\mathcal{M}(T, (e_1, ..., e_n), (f_1, ..., f_m)) \tag{7.4}$$

In essence, once a basis is fixed, then $T_{ij} = \overline{(T^*)_{ji}}$ *for all* i, j.

7.1.2 Self-Adjoint Operators

Definition 7.1.3: Self-Adjoint

An operator $T \in \mathcal{L}(V)$ *is called* ***self-adjoint*** *if* $T = T^*$. *In other words,* $T \in \mathcal{L}(V)$ *is self-adjoint if and only if*

$$\langle Tv, w \rangle = \langle v, Tw \rangle \tag{7.5}$$

for all $v, w \in V$. *Some mathematicians use the term* ***Hermitian*** *instead of self-adjoint.*

Proposition 7.1.5: Eigenvalues of Self-Adjoint Operators are Real

Every eigenvalue of a self-adjoint operator is real.

Proposition 7.1.6: Over $\mathbb{C}$, Tv is orthogonal to v for all v only for the 0 operator

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Suppose

$$\langle Tv, v \rangle = 0 \tag{7.6}$$

for all $v \in V$. Then $T = 0$.

Proposition 7.1.7: Over $\mathbb{C}, \langle Tv, v \rangle$ is real for all v only for self-adjoint operators

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then T is self-adjoint if and only if

$$\langle Tv, v \rangle \in \mathbb{R} \tag{7.7}$$

for every $v \in V$.

Proposition 7.1.8: If $T = T^*$ and $\langle Tv, v \rangle = 0$ for all v, then $T = 0$

Suppose T is a self-adjoint operator on V such that

$$\langle Tv, v \rangle = 0 \tag{7.8}$$

for all $v \in V$. Then $T = 0$.

7.1.3 Normal Operators

Definition 7.1.4: Normal Operator

- *An operator on an inner product space is called **normal** if it commutes with its adjoint.*
- *In essence, $T \in \mathcal{L}(V)$ is normal if*
$$TT^* = T^*T \tag{7.9}$$

Proposition 7.1.9: T is normal if and only if $||Tv|| = ||T^*v||$ for all v

An operator $T \in \mathcal{L}(V)$ is normal if and only if

$$||Tv|| = ||T^*v|| \tag{7.10}$$

for all $v \in V$.

Proposition 7.1.10: For T normal, T and T^* have the same eigenvectors

Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. Then v is also an eigenvector of T^* with eigenvalue $\overline{\lambda}$.

Proposition 7.1.11: Orthogonal Eigenvectors for Normal Operators

Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

7.2 The Spectral Theorem

7.2.1 The Complex Spectral Theorem

Theorem 7.2.1: Complex Spectral Theorem

Suppose $\mathbb{F} = \mathbb{C}$ *and* $T \in \mathcal{L}(V)$*. Then the following are equivalent:*

(a) T *is normal.*

(b) V *has an orthonormal basis consisting of eigenvectors of* T*.*

(c) T *has a diagonal matrix with respect to some orthonormal basis of* V*.*

7.2.2 The Real Spectral Theorem

Proposition 7.2.1: Invertible Quadratic Expressions

Suppose $T \in \mathcal{L}(V)$ *is self-adjoint and* $b, c \in \mathbb{R}$ *are such that* $b^2 < 4c$*. Then*

$$T^2 + bT + cI \tag{7.11}$$

is invertible.

Proposition 7.2.2: Self-Adjoint Operators have Eigenvalues

Suppose $V \neq \{0\}$ *and* $T \in \mathcal{L}(V)$ *is a self-adjoint operator. Then* T *has an eigenvalue.*

Proposition 7.2.3: Self-Adjoint Operators and Invariant Subspaces

Suppose $T \in \mathcal{L}(V)$ *is self-adjoint and* U *is a subspace of* V *that is invariant under* T*. Then*

(a) $U^\perp$ *is invariant under* T*.*

(b) $T|_U \in \mathcal{L}(U)$ *is self-adjoint.*

(c) $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ *is self-adjoint.*

Theorem 7.2.2: Real Spectral Theorem

Suppose $\mathbb{F} = \mathbb{R}$ *and* $T \in \mathcal{L}(V)$*. Then the following are equivalent:*

(a) T *is self-adjoint.*

(b) V *has an orthonormal basis consisting of eigenvectors of* T*.*

(c) T *has a diagonal matrix with respect to some orthonormal basis of* V*.*

7.3 Positive Operators and Isometries

7.3.1 Positive Operators

Definition 7.3.1: Positive Operator

*An operator $T \in \mathcal{L}(V)$ is called **positive** if T is sef-adjoint and*

$$\langle Tv, v \rangle \geq 0 \tag{7.12}$$

for all $v \in V$. Some mathematicians also use the term positive semidefinite operator.

Definition 7.3.2: Square Root

*An operator R is called a **square root** of an operator T if $R^2 = T$.*

Proposition 7.3.1: Characterization of Positive Operators

Let $T \in \mathcal{L}(V)$. Then the following are equivalent:

(a) T is positive;

(b) T is self-adjoint and all the eigenvalues of T are nonnegative;

(c) T has a positive square root;

(d) T has a self-adjoint square root;

(e) There exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$.

Proposition 7.3.2: Each Positive Operator has Only One Positive Square Root

Every positive operator on V has a unique positive square root.

7.3.2 Isometries

Definition 7.3.3: Isometry

- *An operator $S \in \mathcal{L}(V)$ is called an **isometry** if*

$$||Sv|| = ||v|| \tag{7.13}$$

for all $v \in V$.

- *In other words, an operator is an isometry if it preserves norms.*

Theorem 7.3.1: Characterization of Isometries

Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is an isometry;

(b) $\langle Su, Sv \rangle = \langle u, v \rangle$ *for all* $u, v \in V$;

(c) $Se_1, ..., Se_n$ *is orthonormal for every orthonormal list of vectors* $e_1, ..., e_n$ *in* V;

(d) *There exists an orthonormal basis* $e_1, ..., e_n$ *of* V *such that* $Se_1, ..., Se_n$ *is orthonormal;*

(e) $S^*S = I$;

(f) $SS^* = I$;

(g) S^* *is an isometry;*

(h) S *is invertible and* $S^{-1} = S^*$

Proposition 7.3.3: Description of Isometries when $\mathbb{F} = \mathbb{C}$

Suppose V *is a complex inner product space and* $S \in \mathcal{L}(V)$. *Then the following are equivalent:*

(a) S *is an isometry.*

(b) *There is an orthonormal basis of* V *consisting of eigenvectors of* S *whose corresponding eigenvalues all have absolute value* 1.