Rice University
Department of Mathematics

An Introduction to Linear Algebra

Based on MATH 354 at Rice University

Author
Gabriel Gress

May 30, 2021

Contents

Chapter 0

Introduction

These notes cover the core introductory topics in linear algebra. In particular, it defines and states the major theorems for finite-dimensional vector spaces, as well as defining important terms for all mathematics such as eigenvalues, inner products, operator theory, and more.

The lecture notes are based off two main sources. The overall outline and the major statements of theorems and definitions are based off lecture notes from Dr. Várilly-Alvarado during the Fall 2019 teaching of Rice's course MATH 354 – *Honors Linear Algebra*. These notes are supplemented by notes and exercises from Axler's *Linear Algebra Done Right*.

Chapter 1

Fields and Other Prerequisites

1.1 Fields

Definition 1.1.1: Fields

Let F be a field. Then we define $F^n = \{(x_1, \ldots, x_n) \mid x_i \in F, i = 1, \ldots, n\}$ and the following properties are true:

- elements of F^n can be added: if $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$
- Has a neutral element: $(0, \ldots, 0) =: \overline{0}$
- Has an additive inverses: $-x = (-x_1, \ldots, -x_n)$ where $x + (-x) = \overline{0}$.
- Elements of F^n can be "scaled" by elements of F: Let $\lambda \in F, x = (x_1, \ldots, x_n) \in F^n$. Then we define $\lambda \cdot x$ as $(\lambda \cdot x_1, \ldots, \lambda \cdot x_n)$.

Warning: F^n is NOT a field, unless $n = 1$ (because we cannot well-define a multiplying property).

1.2 Proofs by Induction

Theorem 1.2.1: Principle of Mathematical Induction II

Let $P(n)$ be a statement indexed by $n \in N$. Suppose that

1. $P(1)$ is true. (The *base case*)
2. If $P(k)$ is true for some $k \in \mathbb{N}$, then $P(k+1)$ is true. (The *inductive hypothesis*)

Then $P(n)$ is true for all $n \in N$.

Theorem 1.2.2: Principle of Mathematical Induction III

Let $Q(n)$ be a statement indexed by $n \in N$. Suppose that

1. $Q(k_0)$ is true for some $k_0 \in \mathbb{N}$. (The *base case*)
2. If $Q(k)$ is true for some $k \in \mathbb{N}$, then $Q(k+1)$ is true. (The *inductive hypothesis*)

Then $Q(n)$ is true for all $n \geq k_0$.

Chapter 2

Vector Spaces

Definition 2.0.1: Vector Space

A **vector space** has four key elements:

- Our space of vectors V
- A field F
- An addition operation $+ : V \times V \to V$
- A scalar multiplication operation $\cdot : F \times V \to V$

where operations (3) and (4) must satisfy:

- Commutativity: $u + v = v + u \ \forall u, v \in V$
- Associativity: $(u + v) + w = u + (v + w)$, $\forall u, v, w \in V$ AND $(\lambda_1 \cdot \lambda_2) \cdot v = \lambda_1 \cdot (\lambda_2 \cdot v)$
- Additive identity: $\exists \overline{0} \in V$ such that $v + \overline{0} = v$, $\forall v \in V$.
- Scalar multiplicative identity: $1_F \cdot v = v$, $\forall v \in V$
- Distributivity:

$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda * v \quad \lambda \in F,\ u, v \in V$$
$$(\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v \quad \lambda_1, \lambda_2 \in F,\ v \in V$$

We call elements $v \in V$ **vectors**.

Example 2.0.1: Vector Spaces

Some pedagogical examples:

- $(\mathbb{R}^2, \mathbb{R}, +, \cdot)$
- $(\mathbb{C}^2, \mathbb{C}, +, \cdot)$ (via complex scalar multiplication)
- $(\mathbb{C}^2, \mathbb{R}, +, \cdot)$ (via real scalar multiplication)
- $(F^n, F, +, \cdot)$
- $(F^\infty, F, +, \cdot)$ where $F^\infty = \{(x_1, x_2, \ldots) \mid x_i \in F, i = 1, 2, \ldots\}$ (note that this space is infinite dimensional)
- Let F be a field, and S a set. Let $V = F^S :=$ { all functions $f : S \to F$}. Addition is defined in V as follows: Let $f, g : S \to F$. Then $\forall s \in S$

  $$(f + g) : S \to F \text{ and } (f + g)(s) := f(s) + g(s)$$

  Scalar multiplication is defined as follows: Let $\lambda \in F; f \in V; f : S \to F$, so $(\lambda \cdot f) : S \to F$. Then $\forall s \in S$

  $$(\lambda \cdot f)(s) := \lambda \cdot f(s) \forall s \in S$$

- Let $F = \mathbb{R}$ and $V =$ {polynomials of degree ≤ 19 with coefficients in $\mathbb{R}$}. Then V is a vector space.

Remark 2.0.1

If $S = \{1, \ldots, n\}$ then $V = F^S$ which is equivalent to functions $f : \{1, \ldots, n\} \to F$, which is equivalent to F^n. Thus they are isomorphic.

2.1 First Consequences of Axioms of a Vector Space

Let $(V, F, +, \cdot)$ be a vector space.

Proposition 2.1.1

The additive identity of V is unique

Proof. Suppose V has two additive identities $\overline{0}, \overline{0'}$. Then $\overline{0} = \overline{0} + \overline{0'} = \overline{0'}$. ∎

Proposition 2.1.2

Additive inverses are unique

Proof. Let $v \in V$ be any vector. Suppose that w, w' are additive inverses of v. Then

$$\begin{aligned} w &= w + \overline{0} \\ &= w + (v + w') \\ &= (w + v) + w' \\ &= \overline{0} + w' \\ &= w' \end{aligned}$$

∎

Proposition 2.1.3

For all $v \in V, 0 \cdot v = \overline{0}$

Proof.

$$\begin{aligned} 0 \cdot v &= (0+0) \cdot v \\ &= 0 \cdot v + 0 \cdot v \end{aligned}$$

Add $-0 \cdot v$ to both sides:

$$\begin{aligned} -(0 \cdot v) + (0 \cdot v) &= -(0 \cdot v) + (0 \cdot v) + (0 \cdot v) \\ &\implies \overline{0} = \overline{0} + (0 \cdot v) \\ &\implies \overline{0} = 0 \cdot v \end{aligned}$$

■

Exercise 2.1.1

Prove that $-v = (-1) * v$.

2.2 Subspaces

Definition 2.2.1: Subspace

Let $(V, F, +, \cdot)$ be a vector space. A subset $U \subseteq V$ is a subspace if $(U, F, +, \cdot)$ is itself a vector space under the same operations as V.

Lemma 2.2.1: Conditions for subspaces

$U \subseteq V$ is a subspace if and only if

- $\overline{0}$ is still in U
- U is closed under addition: if $u, v \in U$ then $u + v \in U$
- U is closed under scalar multiplication; if $v \in U$ and $\lambda \in F$ then $\lambda \cdot v \in U$

The three conditions ensure that the additive identity of V is in U, and that both addition and scalar multiplication make sense in U.

Example 2.2.1: Example of subspace

Let $F = \mathbb{R}$ and $V = \{f \mid f : (0, 3) \to \mathbb{R}\}$. Let

$$U = \{f \in V \mid f \text{ differentiable, } f'(2) = 0\} \subseteq V.$$

Then $U \subset V$ is a subspace.

Proof of Example. The zero vector of V is $\overline{0} : (0, 3) \to \mathbb{R}$ defined by $\overline{0} : x \mapsto 0$. The zero vector is differentiable and zero at $x = 2$, so the zero vector is in our set U.

Now we want to show that U is closed under addition. Let $f, g \in U$. By the linearity of differentiation, $(f + g)$ is differentiable, and so satisfies the first property of U. If $f'(2) = 0$ and $g'(2) = 0$, then $\frac{d}{dx}(f(x) + g(x)) \mid_{x=2} = 0 + 0 = 0$

Finally, we want to show that U is closed under scalar multiplication. Let $f \in U$ and $\lambda \in \mathbb{R}$. Consider $\lambda \cdot f(x)$. Because λ is a scalar, we know that $\lambda \cdot f(x)$ is still differentiable. Moreover, the derivative of the new function is simply $\lambda \cdot f'(x)$, and at $x = 2$, $\lambda \cdot f'(2) = \lambda \cdot 0 = 0$ and so U is still closed under scalar multiplication. ■

2.3 Sums of Subspaces

Let $(V, F, +, \cdot)$ be a vector space, and let $U_1, \ldots, U_m$ be subspaces. Define

$$U_1 + \ldots + U_m := \{u_1 + \ldots + u_m \mid u_i \in U_i \text{ for } i = 1, \ldots, m\}$$

to be the sum of two subspaces.

Example 2.3.1

Let $V = \mathbb{R}^3$, $F = \mathbb{R}$ and consider

$$U_1 = \{(x, 0, 0) \mid x \in \mathbb{R}\} \in \mathbb{R}^3$$
$$U_2 = \{(0, y, 0) \mid y \in \mathbb{R}\} \in \mathbb{R}^3$$

. Then

$$U_1 + U_2 = \{(x, y, 0) \mid x, y \in \mathbb{R}\}.$$

Theorem 2.3.1: Sum of subspaces

If $U_1, \ldots, U_m \subseteq V$ are subspaces of $(V, F, +, \cdot)$, then $U_1 + \ldots + U_m$ is the "smallest" subspace of V that contains $U_1, \ldots, U_m$. That is, if there exists $U' \subset V$ such that $U_1, \ldots, U_m \subset U'$, then $U_1 + \ldots + U_m \subset U'$.

Proof. First we show that $U_1 + \ldots + U_m$ is a subspace of V. It is easy to see that $\overline{0} = \overline{0} + \ldots + \overline{0}$ and so the zero vector is in our space. We also know that

$$(u_1 + \ldots + u_m) + (v_1 + \ldots + v_m) = (u_1 + v_1) + (u_2 + v_2) + \ldots + (u_m + v_m) \in U_1 + \ldots + U_m$$

Finally,

$$\lambda \cdot (u_1 + \ldots + u_m) = \lambda \cdot u_1 + \ldots + \lambda \cdot u_m \in U_1 + \ldots + U_m.$$

which gives us that $U_1 + \ldots + U_m$ is a subspace of V as desired.

Now we show that $U_i \subseteq U_1 + \ldots + U_m$. Let $u_i \in U_i$. Then

$$u_2 = \overline{0} + \ldots + \overline{0} + u_i + \overline{0} + \ldots + \overline{0} \implies u_i \in U_1 + \ldots + U_m$$

as desired.

Finally, we want to show the statement. Let $U' \subseteq V$ be a subspace with $U_1, \ldots, U_m \subseteq U'$, and let $u_1 + \ldots + u_m \in U_1 + \ldots + U_m$. Our work above gives us that $u_i \in U_i$. Because we have $U_i \subset U'$, that implies that each $u_i \in U'$, and so

$$u_1 + \ldots + u_m \in U'$$

as desired. ■

2.4 Direct sums

Definition 2.4.1: Direct sum

Let $U_1, \ldots, U_m$ be subspace of $(V, F, +, \cdot)$. If each $v \in U_1 + \ldots + U_m$ can be written in *exactly* one way as

$$v = u_1 + \ldots + u_m \quad u_i \in U_i$$

then we say $U_1 + \ldots + U_m$ is a **direct sum**, and denote it by

$$U_1 \oplus U_2 \oplus \ldots \oplus U_m.$$

Example 2.4.1

Let $V = \mathbb{R}^3$, $F = \mathbb{R}$, and consider $U_1 = \{(x, y, 0) \mid x, y \in R\}$ and $U_2 = \{(0, 0, z) | z \in \mathbb{R}\}$. Then $U_1 \bigoplus U_2$ is direct.

But consider $U_3 = \{(0, y, y) \mid y \in \mathbb{R}\}$. Then $U_1 + U_2 + U_3$ is NOT direct.

Lemma 2.4.1: Criterion for direct sums

$U_1 \oplus \ldots \oplus U_m \iff [u_1 + \ldots + u_m = 0 \implies u_1, \ldots, u_m = 0]$ where $u_i \in U_i$.

Proof. First we will prove the forward direction. Observe that $\bar{0} = \bar{0} + \ldots + \bar{0}$. Because it is a direct sum, it is the only way to write the zero vector, and so each $u_i \in U_i$ must be $\bar{0}$.

For the reverse direction, suppose that the only way to write $\bar{0} = u_1 + \ldots + u_m$ is to take $u_i = \bar{0}$. Let $v \in U_1 + \ldots + U_m$. Suppose that

$$v = u_1 + \ldots + u_m \quad u_i \in U_i$$
$$v = u_1' + \ldots + u_m' \quad u_i' \in U_i$$

Subtraction yields $\bar{0} = (u_1 - u_1') + \ldots + (u_m - u_m')$. By hypothesis, each parentheses must be zero, and so $u_i = u_i'$ and hence the representation of v is unique. ■

Lemma 2.4.2: Direct sum of two subspaces

Let $U, W \subseteq V$ be two subspaces. Then

$$U \oplus W \iff U \cap W = \{\vec{0}\}.$$

2.5 Linear Independence and Span

Linear Combination

Definition 2.5.1: Linear Combination

Let $(V, F, +, \cdot)$ be a vector space. A **linear combination of** $v_1, \ldots, v_m \in V$ is a vector of the form

$$a_1 v_1 + a_2 v_2 + \ldots a_m v_m \quad a_i \in F$$

Definition 2.5.2: Span

The **span of** $v_1, \ldots, v_m$ is the set of ALL possible combination of the vectors:

$$\mathrm{Span}(v_1, \ldots, v_m) := \{a_1 v_1 + \ldots + a_m v_m \mid a_i \in F\}.$$

Proposition 2.5.1

$\text{Span}(v_1, \ldots, v_m)$ is the smallest subspace of V that contains $v_1, \ldots, v_m$.

If $\text{Span}(v_1, \ldots, v_m) = V$, we say that $v_1, \ldots, v_m$ "span" the space V.

We call the set of unit vectors of a space the **standard basis**.

Definition 2.5.3

A vector space V is **finite-dimensional** if $V = span(v_1, \ldots, v_m)$.

If a vector space is not finite-dimensional, then we say it is **infinite-dimensional**. To maintain convention, if $V = F[x]$, then $\deg(0) = -\infty$.

Theorem 2.5.1

$F[x]$ is infinite-dimensional.

Linear Independence

Definition 2.5.4: Linear independence

A list $v_1, \ldots, v_m$ of vectors is **linearly independent** if and only if

$$a_1v_1 + \ldots + a_mv_m = \vec{0} \implies a_1, \ldots, a_m = 0$$

A list of vectors in V is **linearly dependent** if it is not linearly independent.

We can also say that a list in V is linearly dependent if there exist $a_1, \ldots, a_m$ not all zero such that $a_1v_1 + \ldots + a_mv_m = 0$.

Lemma 2.5.1

If $v_1, \ldots, v_m$ are linearly independent, then there is exactly one way to write $v \in \text{Span}(v_1, \ldots, v_m)$ as a linear combination of the vectors.

Theorem 2.5.2

Let V be a finite-dimensional vector space. If $v_1, \ldots, v_m$ is an arbitrary set of linearly independent vectors in V and $w_1, \ldots, w_k$ is an arbitrary set of vectors that span V, then $m \leq k$.

Lemma 2.5.2: Linear Dependence Lemma

Let $v_1, \ldots, v_m \in V$ be a set of linearly dependent vectors. Then there exists an index j such that

- $v_j \in \text{Span}(v_1, \ldots, v_{j-1})$
- $\text{Span}(v_1, \ldots, v_m) = \text{Span}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$

Lemma 2.5.3: Finite-dimensional subspaces

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

2.6 Bases

Definition 2.6.1: Basis

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

Lemma 2.6.1: Criterion for basis

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

$$v = a_1 v_1 + \ldots + a_n v_n$$

where $a_1, \ldots, a_n \in F$.

Theorem 2.6.1

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

Hence, every finite-dimensional vector space has a basis.

Corollary 2.6.1

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

Theorem 2.6.2: Soulmate space

Let $(V, F, +, \cdot)$ be a finite dimensional vector space. let $U \subset V$ be a subspace. Then there exists a $W \subset V$ such that

$$U \oplus W = V$$

Note that W is a finite-dimensional vector space.

2.7 Dimension

Definition 2.7.1: Dimension

The **dimension** of a finite-dimensional vector space is the length of any basis of the vector space.

Lemma 2.7.1: Dimension of a subspace

If V is finite-dimensional and U is a subspace of V, then

$$\text{dim} U \leq \text{dim} V.$$

Lemma 2.7.2

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

Dimension of a sum of two spaces

Proposition 2.7.1

Let $U_1, U_2 \in V$ be finite dimensional vector spaces over some field F.

Then

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Corollary 2.7.1

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

Chapter 3

Linear Transformations

Let V, W be vector spaces over a field F.

Definition 3.0.1: Linear Map

A **linear transformation** or **map** is a function $T : V \to W$ such that, for all $u, v \in V$:

$$T(u+v) = T(u) + T(v)$$
$$T(\lambda \cdot v) = \lambda \cdot T(v)$$

Notice that $T(\vec{0}_V) = \vec{0}_W$.

We denote by $\mathcal{L}(V, W) := \{T : V \to W \mid T \text{ is linear}\}$

Example 3.0.1: Examples of Linear Transformations

Some examples include

- Zero map: $0 : V \to W$ is defined as $v \mapsto \vec{0}_W$
- Identity: $\text{Id} : V \to V$ is defined as $v \mapsto v$
- Differentation: $D : P(\mathbb{R}) \mapsto P(\mathbb{R})$ is defined as $p \mapsto p'$
- Integration: $T : P(\mathbb{R}) \to \mathbb{R}$ is defined as $p \mapsto \int_0^1 p := \int_0^1 p(x)dx$
- Shift: $S : F^\infty \to F^\infty$ is defined as $(x_1, x_2, \ldots) \mapsto (x_2, x_3, \ldots)$
- $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined as $(x, y, z) \mapsto (5x - y + 2z, 7x + 37 - 19z)$

Corollary 3.0.1

The set $\mathcal{L}(\mathcal{V}, \mathcal{W})$ is itself a vector space over F.

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

Definition 3.1.1: Addition and scalar multiplication on linear maps

Suppose $S, T \in \mathcal{L}(V, W)$ and $\lambda \in F$. The **sum** $S + T$ and **product** λT are the linear maps from $V \to W$ defined by

$$(S+T)(v) = Sv + Tv$$
$$(\lambda T)(v) = \lambda(Tv)$$

for all $v \in V$.

Composition of maps is also a map; $S(T(V))$ is similar to "multiplying" maps.

Moreso, these compositions are associative: $(T_3 \circ (T_2 \circ T_1) = (T_3 \circ T_2) \circ T_1)$. It also has the identity: $Id_W \circ T = T \circ Id_V$.

Finally, it is also distributive; $T \circ (S_1 + S_2) = T \circ S_1 + T \circ S_2$

Say $T \in L(V, W)$, $T : V \to W$. Suppose that V is a finite-dimensional vector space. Then T is *determined* by what it does to a basis of V :

Say $v_1, \ldots, v_n$ is a basis for V. Suppose we know the output of $T(v_1), \ldots, T(v_n)$. Then let $v \in V$. Know $v = a_1 v_1 + \ldots + a_n v_n$. Then $T(v) = T(a_1 v_1 + \ldots + a_n v_n) = T(a_1 v_1) + \ldots + T(a_n v_n) = a_1 T(v_1) + \ldots + a_n T(v_n)$

Therefore, we can construct $T(v)$ for any v from knowledge of how T applies to the basis.

Theorem 3.1.1: Linear maps and basis of domain

Suppose that $w_1, \ldots, w_n \in W$ are arbitrary vectors and $v_1, \ldots, v_n$ is a basis of V. Then there is **exactly** one linear map $T : V \to W$ that sends v_i to w_i.

Note that $\{w_i\}$ are not necessarily linearly independent or spanning!

3.2 Kernels

Definition 3.2.1: Kernel

Let $T \in L(V, W)$, $T : V \to W$. The **kernel** of T is

$$\mathrm{ker} T := \left\{ v \in V \mid T(v) = \vec{0})W \right\}$$

Note: $T(\vec{0}_V) = \vec{0}_W$ and so $\vec{0}_V \in \ker(T)$ automatically. We also refer to the kernel by $\mathrm{Null}(T)$ sometimes.

Example 3.2.1: Examples of kernels

- The zero map obviously has a kernel of the whole set
- The kernel of Id_V is the set $\left\{ \vec{0}_V \right\}$
- Differentiation has a kernel consisting of constants

Definition 3.2.2: Injective

A map is **injective** if $T(u) = T(v) \implies u = v$
In other words, different inputs give different outputs.

Proposition 3.2.1

Let $T \in L(V, W)$. Then T is injective if and only if $\ker(T) = \left\{ \vec{0}_V \right\}$

3.3 Range and Surjectivity

Definition 3.3.1: Range/Image

For T a function from $V \to W$, the **range** or **image** of T is the subset of W defined by:

$$\text{range}T = \mathrm{Im}T = \{Tv \mid v \in V\}.$$

Proposition 3.3.1: Surjectivity

A linear map is **surjective** if for every $w \in W$, there is a $v \in V$ such that $T(v) = w$. In other words:

$$\mathrm{Im}T = W$$

Theorem 3.3.1: Rank-Nullity theorem

Let V be a finite dimensional vector space and W a vector space. Let $T \in \mathcal{L}(V, W)$. Then $\mathrm{Im}T$ is a finite dimensional vector space and

$$\dim V = \dim(\mathrm{Ker}T) + \text{dim}(\mathrm{Im}T).$$

Notice that W is not necessarily finite-dimensional. This theorem carries many huge consequences.

Corollary 3.3.1

Let V, W be finite dimensional vector spaces over F. Suppose that $\text{dim}V > \text{dim}W$. Then any linear map $T : V \to W$ cannot be injective.

Corollary 3.3.2

Let V, W be finite dimensional vector spaces over F. Suppose that $\text{dim}V < \text{dim}W$. Then any linear map $T : V \to W$ cannot be surjective.

The Rank-Nullity Theorem is really powerful and allows us to prove many theorems more directly (versus a standard matrix elimination proof, which is often more verbose).

Proposition 3.3.2

Consider a system of linear equations given by

$$\sum_{j=1}^{n} A_{1,j}x_j = b_1$$
$$\vdots$$
$$\sum_{j=1}^{n} A_{m,j}x_j = b_m$$

where $A_{i,j} \in F$ constant, $x_i \in F$ undetermined, and $b_i \in F$ constant. We call the system **homogeneous** when $b_i = 0$ for all $1 \leq i \leq m$.

If the system of equations is homogeneous and $n > m$ (more variables than equations), then the system of equations has nontrivial solutions.

If instead, the system of equations is **inhomogeneous** (not all b_i are zero), and $m > n$ (more equations than variables), then for every choice of $A_{i,j}$ there exist b_i for which there are no solutions.

Proof. For the first part of the proposition, define $T : F^m \to F^n$ by

$$T(x_1, \dots, x_n) = \left(\sum_{j=1}^{n} A_{1,j}x_j, \dots, \sum_{j=1}^{n} A_{n,j}x_j \right)$$

Then the homogeneous system given in the proposition is exactly

$$T(x_1, \dots, x_n) = 0.$$

Then it suffices to show that $\mathrm{null}(T) \neq \{0\}$. By the Rank-Nullity Theorem, T cannot be injective, and so this holds. Notice that because $\mathrm{null}(T)$ is a vector space, we can find a basis for $\mathrm{null}(T)$ which then describes all solutions to the homogeneous system of equations.

For the second part of the proposition, we use the same T, but instead observe that because $m > n$, T cannot be surjective. This implies that there exist b_i for which the system has no solutions, regardless of choice of $A_{i,j}$.

■

3.4 Matrices

Representing a Linear Map by Matrices

Definition 3.4.1: Matrix

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

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

where j refers to the row number, and k refers to the column number.

Definition 3.4.2: Matrix of a linear map

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

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

If the bases are not clear from context, we use the notation $\mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))$ is used.

Example 3.4.1: Examples of matrices of linear maps

Consider $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $(x, y) \mapsto (x + 3y, 2x + 5y, 7x + 9y)$. By applying this to the standard basis for $\mathbb{R}^2$ and $\mathbb{R}^3$, we can determine the matrix:

$$\begin{aligned} T(v_1) &= T(1, 0) = (1, 2, 7) \\ T(v_2) &= T(0, 1) = (3, 5, 9) \\ \text{so then } \mathcal{M}(T) &= \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix} \end{aligned}$$

Another important example is the differentiation map $D : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ defined by $p \mapsto p'$.

Choose bases $P_3(\mathbb{R}) = \{1, x, x^2, x^3\}$ and $P_2(\mathbb{R}) = \{1, x, x^2\}$. Then

$$\begin{aligned} D(v_1) &= D(1) = 0 = 0 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 \\ D(v_2) &= D(x) = 1 = 1 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 \\ D(v_3) &= D(x^2) = 2x = 0 \cdot w_1 + 2 \cdot w_2 + 0 \cdot w_3 \\ D(v_4) &= D(x^3) = 3x^2 = 0 \cdot w_1 + 0 \cdot w_2 + 3 \cdot w_3 \\ \text{So therefore } \mathcal{M}(D) &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \end{aligned}$$

3.5 Addition and Scalar Multiplication of Matrices of Linear Maps

Corollary 3.5.1: Linearity of Matrices

Suppose $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.

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

Notation: The set of all m-by-n matrices with entries in F is denoted by $F^{m,n}$.

Corollary 3.5.2

If m, n are positive integers, then $F^{m,n}$ is a vector space with dimension mn.

3.6 Matrix Multiplication

Definition 3.6.1

Suppose A is an $m \times n$ matrix and C is an $n \times p$ matrix. Then AC is defined to be the $m \times p$ matrix whose entries are defined as follows:

$$(AC)_{jk} = \sum_{r=1}^{n} A_{jr} C_{rk}.$$

Matrix multiplication is not commutative.

Corollary 3.6.1

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

We denote $A_{j,\cdot}$ to be the $1 \times n$ matrix consisting of row j of A. Likewise, $A_{\cdot,k}$ denotes the $m \times 1$ matrix consisting of column k of A.

Corollary 3.6.2

Suppose A is an $m \times n$ matrix and C is an $n \times p$ matrix. Then

$$(AC)_{jk} = A_{j\cdot} C_{\cdot k}$$

Corollary 3.6.3

Suppose A is $m \times n$ matrix and C is $n \times p$ matrix. Then

$$(AC)_{k} = AC_{k}$$

Corollary 3.6.4

Suppose A is $m \times n$ matrix and $c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ is an $n \times 1$ matrix. Then

$$Ac = c_1 A_{\cdot 1} + \ldots + c_n A_{\cdot n}.$$

3.7 Invertibility and Isomorphic Vector Spaces

Invertible Linear Maps

Definition 3.7.1

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 that S that satisfies this is called an **inverse** of T. we denote the inverse by T^{-1}.

Lemma 3.7.1

An invertible linear map has a unique inverse.

Theorem 3.7.1

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

We will not show it here, but one can see that the inverse of a linear map corresponds to the matrix inverse.

Isomorphic Vector Spaces

Definition 3.7.2: Isomorphism and Isomorphic

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

Lemma 3.7.2: Equal dimension implies isomorphic

Two finite-dimensional vector spaces over F are isomorphic if and only if they have the same dimension.

Corollary 3.7.1

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

Corollary 3.7.2: Dimension of Space of Linear Maps

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

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

Linear Maps Thought of as Matrix Multiplication

Definition 3.7.3: Matrix of a vector

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

$$\mathcal{M}(v) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

where $c_1, \ldots, c_n$ are the scalars such that

$$v = c_1 v_1 + \ldots + c_n v_n.$$

Corollary 3.7.3

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \ldots, v_n$ is a basis of V, $w_1, \ldots, w_m$ a basis of W. Then the k-th column of $\mathcal{M}(T)$ equals $\mathcal{M}(v_k)$.

Lemma 3.7.3: Linear maps act like matrix multiplication

Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Suppose v, w are bases of W. Then

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

Operators

Definition 3.7.4: Operators

A linear map from a vector space to itself is called an **operator** or **endomorphism**. We notate $\mathcal{L}(V)$ as the set of all operators on V.

Theorem 3.7.2: Injectivity is equivalent to surjectivity in finite dimensions

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

- T is invertible
- T is injective
- T is surjective

Chapter 4

Special Vector Spaces

4.1 Products and Direct Sums

Definition 4.1.1: Product of Vector Spaces

Suppose $V_1, \ldots, V_m$ are vector spaces over F.

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

$$V_1 \times \ldots \times V_m = \{(v_1, \ldots, v_m) \mid v_1 \in V_1, \ldots, v_m \in V_m\}.$$

- Addition on $V_1 \times \ldots \times V_m$ is defined by

$$(u_1, \ldots, u_m) + (v_1, \ldots, v_m) = (u_1 + v_1, \ldots, u_m + v_m).$$

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

$$\lambda(v_1, \ldots, v_m) = (\lambda v_1, \ldots, \lambda v_m).$$

Lemma 4.1.1

Suppose $V_1, \ldots, V_m$ are vector spaces over F. Then $V_1 \times \ldots \times V_m$ is a vector space over F.

Lemma 4.1.2: Dimension of a product is sum of dimensions

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

$$\dim(V_1 \times \ldots \times V_m) = \dim V_1 + \ldots + \dim V_m.$$

Lemma 4.1.3: Products and direct sums

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

$$\Gamma(u_1, \ldots, u_m) = u_1 + \ldots + u_m.$$

Then $U_1 + \ldots + U_m$ is a direct sum if and only if Γ is injective (and thus invertible).

Lemma 4.1.4

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

$$\dim(U_1 + \ldots + U_m) = \dim U_1 + \ldots + \dim U_m.$$

4.2 Quotients of Vector Spaces

Definition 4.2.1: v + U

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

$$v + U = \{v + u \mid u \in U\}.$$

Definition 4.2.2: 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 4.2.3: Quotient space

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,

$$V/U = \{v + U \mid v \in V\}.$$

Lemma 4.2.1

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

- $v - w \in U$
- $v + U = w + U$
- $(v + U) \cap (w + U) \neq \emptyset$

Definition 4.2.4: addition and scalar multiplication on V/U

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

$$(v + U) + (w + U) = (v + w) + U$$
$$\lambda(v + U) = (\lambda v) + U$$

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

Theorem 4.2.1

Suppose U is a subspace of V. Then V/U is a vector space.

Definition 4.2.5: 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

$$\pi(v) = v + U$$

for $v \in V$.

Lemma 4.2.2

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

$$\dim V/U = \dim V - \dim U.$$

Each linear map T on V induces a linear map $\hat{T} on V/(\text{null} T)$, which we define:

Definition 4.2.6

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

$$\hat{T}(v + \text{null} T) = Tv.$$

Lemma 4.2.3: Null space and range of $\hat{T}$

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

- $\hat{T}$ is a linear map from $V/(\text{null} T)$ to W
- $\hat{T}$ is injective
- $\text{range } \hat{T}$ = $\text{range } T$
- $V/(\text{null} T)$ is isomorphic to $\text{range} T$.

4.3 Dual Space and Dual Map

Definition 4.3.1: Linear functional

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

Definition 4.3.2: Dual Space

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

Corollary 4.3.1

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

Definition 4.3.3: Dual Basis

If $v_1, \ldots, v_n$ is a basis of V, then the **dual basis** of $v_1, \ldots, v_n$ is the list $\varphi_1, \ldots, \varphi_n$ of elements of V', where each φ_j is the linear functional on V such that

$$\varphi_j(v_j) = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}$$

Lemma 4.3.1

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

Definition 4.3.4: dual map

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 $T'(\varphi) = \varphi \circ T$ for $\varphi \in W'$.

Lemma 4.3.2: Algebraic properties of dual maps

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

Definition 4.3.5: Transpose

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

$$(A^T)_{k,j} = A_{j,k}$$

Corollary 4.3.2

If A is $m \times n$, and C is $n \times p$, then $(AC)^T = C^T A^T$

These properties extend to their matrices.

Theorem 4.3.1: The matrix of T' is the transpose of the matrix of T

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

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

Definition 4.4.1: Annihilator

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

$$U^0 = \{\varphi \in V' \mid \varphi(u) = 0 \; \forall u \in U\}.$$

Example 4.4.1

Suppose U is the subspace of $P(\mathbb{R})$ consisting of all polynomial multiples of x^2. If φ is the linear functional on $P(\mathbb{R})$ defined by $\varphi(p) = p'(0)$, then $\varphi \in U^0$.

Corollary 4.4.1

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

Theorem 4.4.1: Dimension of Annihilator

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

$$\dim U + \dim U^0 = \dim V.$$

Theorem 4.4.2: Null Space of T'

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

- $\text{null} T' = (\text{range} T)^0$
- $\dim \text{null}(T') = \dim \text{null}(T) + \dim(W) - \dim(V)$

Corollary 4.4.2

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

Theorem 4.4.3: Range of T'

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

- $\dim \text{range } T' = \dim \text{range } T$
- $\text{range } T' = (\text{null} T)^0$

Corollary 4.4.3

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

Rank

We will define row and column ranks initially, but keep in mind that this terminology will soon prove to be superfluous.

Definition 4.4.2: Row and Column Rank

Suppose A is an m-by-n matrix with entries in F. The **row rank** of A is the dimension of the span of the rows of A in $F^{1,n}$, and the **column rank** of A is the dimension of the span of the columns of A in $F^{m,1}$.

One can then prove that the column rank of $\mathcal{M}(T)$ is equal to $\dim \text{Im}(T)$.

Lemma 4.4.1

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

This is what allows us to get rid of the superfluous terminology and create a simpler definition:

Definition 4.4.3: rank

The **rank** of a matrix $A \in F^{m,n}$ is the column rank of A. That is, the dimension of the span of the columns of A in $F^{m,1}$, or the dimension of $\mathrm{Im}(A)$.

Chapter 5

Eigenvalues, Eigenvectors, and Eigenspaces

Definition 5.0.1: Invariant subspace

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

5.1 Eigenvalues and Eigenvectors

Definition 5.1.1: Eigenvalue

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

Lemma 5.1.1: Equivalent conditions

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

- λ is an eigenvalue of T
- $T - \lambda I$ is not injective
- $T - \lambda I$ is not surjective
- $T - \lambda I$ is not invertible

Definition 5.1.2: Eigenvector

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

Lemma 5.1.2: Linearly independent eigenvectors

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

5.2 Existence of Eigenvalues

Theorem 5.2.1

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

Definition 5.2.1: Upper-triangular Matrix

A matrix ix called **upper-triangular** if all the entries below the diagonal equal zero.

Theorem 5.2.2

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 basis of V.

Theorem 5.2.3

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.

Theorem 5.2.4

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: Eigenspace

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

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

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

Corollary 5.3.1: Sum of eigenspaces is a direct sum

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

$$E(\lambda_1, T) + \ldots + E(\lambda_m, T)$$

is a direct sum. Furthermore,

$$\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T) \leq \dim V.$$

Definition 5.3.2: 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 $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:

- T is diagonalizable
- V has a basis consisting of eigenvectors of T
- There exist one-dimensional subspaces $U_1, \ldots, U_n$ of V, each invaraint under T, such that

$$V = U_1 \bigoplus \ldots \bigoplus U_n$$

- $V = E(\lambda_1, T) \bigoplus \ldots \bigoplus E(\lambda_m, T)$
- $\dim V = \dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T)$.

Corollary 5.3.2

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

5.4 Trace and Determinant

The trace and determinant of linear maps are a vital tool to characterizing linear maps, and now that we have defined eigenvalues, we can work with the concepts well enough to get a deep understanding.

Definition 5.4.1: Trace

Suppose $T \in \mathcal{L}(V)$. The **trace** of T is the sum of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity.

Note that if our underlying field is R, we have to include complex eigenvalues as well.

This might seem different than the traditional definition used in other linear algebra notes, but we will soon see that it corresponds exactly to the other definitions. In particular, the trace has a connection to characteristic polynomials.

Proposition 5.4.1

Suppose $T \in \mathcal{L}(V)$ and let $n = \dim(V)$. Then $\mathrm{Tr}(T)$ is equal to the negative of the coefficient of x^{n-1} in the characteristic polynomial of T.

Now we will make the connection to entries of a matrix.

Theorem 5.4.1

Let $T \in \mathcal{L}(V)$. Then the trace of $\mathcal{M}(T)$ is the sum of the entries of the diagonal, and exactly equals $\mathrm{Tr}(T)$.

Note that this also tells us that the trace is invariant under change of basis, of course.

Corollary 5.4.1

If A and B are linear maps of the same size, then

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

The trace also has a useful additivity property.

Proposition 5.4.2

Suppose $S, T \in \mathcal{L}(V)$. Then $\mathrm{Tr}(S+T) = \mathrm{Tr}(S) + \mathrm{Tr}(T)$.

The trace can be useful in proving some rather strong properties. For example

Proposition 5.4.3

There do not exist operators $S, T \in \mathcal{L}(V)$ such that

$$ST - TS = I.$$

Proof. To see this, observe that

$$\begin{aligned}
\mathrm{Tr}(ST - TS) &= \mathrm{Tr}(ST) - \mathrm{Tr}(TS) \\
&= \mathrm{Tr}(ST) - \mathrm{Tr}(ST) \\
&= 0
\end{aligned}$$

but $\mathrm{Tr}(I) = \dim(V)$, and so equality can never hold. ■

Determinant

Once again, we will define the determinant in an abstract manner, and show it corresponds to our traditional notion of a determinant.

Definition 5.4.2: Determinant

Let $T \in \mathcal{L}(V)$. The **determinant** of T is the product of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity.

If our underlying field is $\mathbb{R}$, then we need to include the complex eigenvalues as well.

Once again, the determinant has many deep connections, including to the characteristic polynomial.

Proposition 5.4.4

Suppose $T \in \mathcal{L}(V)$ and let $n = \dim(V)$. Then $\det(T)$ equals $(-1)^n$ times the constant term of the characteristic polynomial of T.

This alone actually gives us some powerful consequences already.

Proposition 5.4.5

An operator on V is invertible if and only if its determinant is nonzero.

Proof. Recall that the operator T is invertible if and only if 0 is not an eigenvalue of T. Of course, this only occurs if and only if the product of the eigenvalues of T is nonzero, and hence we have the proposition. ■

Some take the proposition below to be the definition, but we will have it follow as a consequence.

Proposition 5.4.6

Let $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T equals $\det(xI - T)$.

All of this corresponds to the traditional definition of determinants of matrices. We will omit the details to show this, however.

Definition 5.4.3: Determinant of a matrix

Let A be an n-by-n matrix given by

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}.$$

The **determinant** of A is defined by

$$\det(A) = \sum_{(\sigma_1,\ldots,\sigma_n)\in \mathrm{perm}(n)} (\mathrm{sign}(\sigma_1,\ldots,\sigma_n))\, A_{m_1,1}\ldots A_{\sigma_n,n}.$$

Proposition 5.4.7

Suppose A is a square matrix and B is obtained from A by interchanging two columns. Then

$$\det(A) = -\det(B).$$

We can also see that if a square matrix has two equal columns, then $\det(A) = 0$.

Proposition 5.4.8

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

$$\det(ST) = \det(TS) = \det(S)\det(T)$$

Many of our special operators have unique traces and determinants.

Proposition 5.4.9

Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is an isometry. Then $|\det(S)| = 1$.

Volume

Proposition 5.4.10

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

$$|\det(T)| = \det\left(\sqrt{T^*T}\right).$$

We define a linear map on a set by

$$T(\Omega) := \{Tx \mid x \in \Omega\}.$$

Proposition 5.4.11

Let $T \in \mathcal{L}(\mathbb{R}^n)$ and $\Omega \subset \mathbb{R}^n$. Then

$$\mathrm{Vol}(T(\Omega)) = |\det(T)|\, \mathrm{Vol}(\Omega)$$

This also implies that isometries don't change volume.

Chapter 6

Inner Products and Norms

6.1 Inner Products

Definition 6.1.1: Dot Product

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

$$x \cdot y = x_1 y_1 + \ldots + x_n y_n$$

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

The dot product takes in two vectors and returns a scalar. Moreover, $x \cdot x = \|x\|^2$.

Corollary 6.1.1: Properties of dot products

- $x \cdot x \geq 0$ for all $x \in \mathbb{R}^n$
- $x \cdot x = 0$ if and only if $x = 0$
- for $y \in \mathbb{R}^n$ fixed, the map from $\mathbb{R}^n$ to $\mathbb{R}$ that sends $x \in \mathbb{R}^n$ to $x \cdot y$ is linear
- $x \cdot y = y \cdot x$ for all $x, y \in \mathbb{R}^n$

This is almost quite right for complex numbers, but for that we must generalize a little further.

Definition 6.1.2: Inner Product

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

- $\langle v, v \rangle \geq 0$ for all $v \in V$ **(positivity)**
- $\langle v, v \rangle = 0$ if and only if $v = 0$ **(definiteness)**
- $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$ **(additivity in first slot)**
- $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$ for all $\lambda \in F$ and all $u, v \in V$ **(homogeneity in first slot)**
- $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in V$ **(conjugate symmetry)**.

Definition 6.1.3: Inner Product Space

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

Notation: For the rest of the chapter, V denotes a inner product space over F.

Corollary 6.1.2: Basic properties of an inner product

- For each fixed $u \in V$, the function that takes v to $\langle v, u\rangle$ is a linear map from V to F
- $\langle 0, u\rangle = 0$ for every $u \in V$
- $\langle u, 0\rangle = 0$ for every $u \in V$
- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$
- $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ for all $\lambda \in F$ and $u, v \in V$

6.2 Norms

Definition 6.2.1: Norm

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

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Corollary 6.2.1: Basic properties of the norm

Suppose $v \in V$.

- $\|v\| = 0$ if and only if $v = 0$.
- $\|\lambda v\| = |\lambda|\,\|v\|$ for all $\lambda \in F$.

Definition 6.2.2: Orthogonal

Two vectors $u, v \in V$ are called **orthogonal** if $\langle u, v\rangle = 0$.

Corollary 6.2.2

- 0 is orthogonal to every vector in V
- 0 is the only vector in V that is orthogonal to itself

Theorem 6.2.1: Pythagorean Theorem

Suppose u, v are orthogonal vectors in V. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Corollary 6.2.3: Orthogonal Decomposition

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

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

Theorem 6.2.2: Cauchy-Schwarz Inequality

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\|v\|.$$

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

Theorem 6.2.3: Triangle Inequality

Suppose $u, v \in V$. Then

$$\|u+v\| \leq \|u\| + \|v\|.$$

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

Theorem 6.2.4: Parallelogram Equality

Suppose $u, v \in V$. Then

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right).$$

6.3 Orthonormal Bases

Definition 6.3.1: Orthonormal

A list of vectors is called **orthonormal** if each vector in the list has norm 1 and is orthogonal to all other vectors in the list.

Corollary 6.3.1

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

Lemma 6.3.1

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

$$v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \ldots + |\langle v, e_n\rangle|^2.$$

Theorem 6.3.1: Gram-Schmidt Procedure

Suppose $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = \frac{v_1}{\|v_1\|}$. For $j = 2, \ldots, m$, define e_j inductively by

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

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

$$span(v_1, \ldots, v_j) = span(e_1, \ldots, e_j)$$

Corollary 6.3.2

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

Corollary 6.3.3

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

Lemma 6.3.2

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 6.3.2: 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.

Theorem 6.3.3: Riesz Representation Theorem

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

$$\phi(v) = \langle v, u \rangle$$

for every $v \in V$.

To construct this vector, first we write

$$\begin{aligned}\varphi(v) &= \varphi\left(\langle v, e_1 \rangle e_1 + \ldots + \langle v, e_n \rangle e_n\right) \\ &= \langle v, e_1 \rangle \varphi(e_1) + \ldots + \langle v, e_n \rangle \varphi(e_n) \\ &= \langle v, \overline{\varphi(e_1)} e_1 + \ldots + \overline{\varphi(e_n)} e_n \rangle\end{aligned}$$

for every $v \in V$. Thus, we can simply set

$$u = \overline{\varphi(e_1)} e_1 + \ldots + \overline{\varphi(e_n)} e_n.$$

Lemma 6.3.3

Suppose there exists $u_1, u_2 \in V$ such that

$$\langle v, u_1 \rangle = \langle v, u_2 \rangle \text{ for all } v \in V.$$

Then $u_1 = u_2$.

6.4 Orthogonal Complements

Definition 6.4.1: Orthogonal Complement

Let $U \subset V$. Then we denote by $U^\perp$ the **orthogonal complement** of U, where

$$U^\perp := \{v \in V \mid \langle v, u \rangle = 0 \quad \forall u \in U\}$$

Be cautious with this definition– it doesn't quite correspond to our normal intuition for orthogonal spaces. For example, a line could be orthogonal to a plane in $\mathbb{R}^3$, and indeed there exists isomorphisms so that the line is an orthogonal complement of said plane, but one must first ensure that the line and plane map to vector spaces accordingly.

Proposition 6.4.1: Properties of Orthogonal Complement

- If $U \subset V$, then $U^\perp \subset V$.
- $\{0\}^\perp = V$.
- $V^\perp = \{0\}$.
- If $U \subset V$, then $U \cap U^\perp = \{0\}$.
- If $U, W \subset V$ and $U \subset W$, then $W^\perp \subset U^\perp$.

Recall the soulmate theorem, which stated that if $U \subset V$, then there existed a vector space $W \subset V$ such that

$$U \oplus W = V.$$

It turns out this theorem is further improved by our new definitions.

Theorem 6.4.1: Soulmate Theorem Revisited

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

$$U \oplus U^\perp = V.$$

By the Rank-Nullity theorem, this gives us that $\dim(U^\perp) = \dim(V) - \dim(U)$.

Theorem 6.4.2

Let U be a finite-dimensional subspace of V. Then

$$U = (U^\perp)^\perp.$$

Note that this fails when U is infinite-dimensional.

Orthogonal Projection

Definition 6.4.2: Orthogonal Projection

Let $U \subset V$ be a finite-dimensional subspace. The **orthogonal projection** of V onto U is the operator P_U on V defined by:

$$P_U v = u$$
$$v = u + w \quad u \in U,\ w \in U^\perp$$

This operator has quite a few very nice properties.

Proposition 6.4.2: Properties of Orthogonal Projection

Let $U \subset V$ be a finite-dimensional subspace, and $v \in V$. Then

- $P_U u = u$ for all $u \in U$
- $P_U w = 0$ for all $w \in U^\perp$
- $\mathrm{Im} P_U = U$
- $\mathrm{Ker} P_U = U^\perp$
- $v - P_U v \in U^\perp$
- $P_U^2 = P_U$
- $\|P_U v\| \leq \|v\|$
- For every orthonormal basis $e_1, \ldots, e_m$ of U,

$$P_U v = \langle v, e_1 \rangle e_1 + \ldots + \langle v, e_m \rangle e_m.$$

Chapter 7

Adjoints and Operators

7.1 Adjoints

Definition 7.1.1: adjoint

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

$$\langle Tv, w \rangle = \langle v, T^* w \rangle$$

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

Corollary 7.1.1

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

Corollary 7.1.2: Properties of the adjoint

- $(S + T)^* = S^* + T^*$ for all $S, T \in \mathcal{L}(V, W)$
- $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in F$ and $T \in \mathcal{L}(V, W)$
- $(T^*)^* = T$ for all $T \in \mathcal{L}(V, W)$
- $I^* = I$, where I is the identity operator on V
- $(ST)^* = T^* S^*$ for all $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$

We can derive some nice properties of the adjoint map by looking at orthogonal complements.

Proposition 7.1.1: Kernel and Image of the Adjoint

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

- $\mathrm{Ker} T^* = (\mathrm{Im} T)^\perp$
- $\mathrm{Im} T^* = (\mathrm{Ker} T)^\perp$
- $\mathrm{Ker} T = (\mathrm{Im} T^*)^\perp$
- $\mathrm{Im} T = (\mathrm{Ker} T^*)^\perp$

Proposition 7.1.2: The matrix of T^*

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

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

is the conjugate transpose of

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

7.2 Self-Adjoint Operators

Definition 7.2.1: self-adjoint

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

$$\langle Tv, w\rangle = \langle v, Tw\rangle$$

for all $v, w \in V$.

Corollary 7.2.1

Every eigenvalue of a self-adjoint operator is real.

Corollary 7.2.2

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

$$\langle Tv, v\rangle = 0$$

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

7.3 Normal Operators

Definition 7.3.1: Normal

An operator on an inner product space is called **normal** if it commutes with its adjoint. In other words, $T \in \mathcal{L}(V)$ is normal if

$$TT^* = T^*T.$$

Proposition 7.3.1

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

$$\|Tv\| = \|T^*v\|$$

for all $v \in V$.

Corollary 7.3.1

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 $\overline{\lambda}$.

Corollary 7.3.2

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

Proposition 7.3.2: Normal operators and invariant subspaces

Suppose V is an inner product space, $T \in \mathcal{L}(V)$ normal, and $U \subset V$ a subspace that is invariant under T. Then

- $U^\perp$ is invariant under T
- U is invariant under T^*
- $(T|_U)^* = (T^*)|_U$
- $T|_U \in \mathcal{L}(U)$ and $T|_U^\perp \in \mathcal{L}(U^\perp)$ are normal operators

7.4 The Spectral Theorem

The Complex Spectral Theorem

Theorem 7.4.1: Complex Spectral Theorem

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

- T is normal
- V has an orthonormal basis consisting of eigenvectors of T.
- T has a diagonal matrix with respect to some orthonormal basis of V.

The Real Spectral Theorem

Lemma 7.4.1

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

$$T^2 + bT + cI$$

is invertible.

Lemma 7.4.2

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

Lemma 7.4.3

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

- $U^{\perp}$ is invariant under T
- $T\mid_U \in \mathcal{L}(U)$ is self-adjoint
- $T\mid_{U^{\perp}} \in \mathcal{L}(U^{\perp})$ is self-adjoint.

Theorem 7.4.2: Real Spectral Theorem

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

- T is self-adjoint
- V has an orthonormal basis consisting of eigenvectors of T.
- T has a diagonal matrix with respect to some orthonormal basis of V.

7.5 Positive Operators and Isometries

Definition 7.5.1: Positive Operator

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

$$\langle Tv, v\rangle \geq 0$$

for all $v \in V$.

If V is a complex vector space, then T does not need to be self-adjoint (if an inner product is real for all values, then the operator must be self-adjoint; positivity implies real).

Definition 7.5.2: Square Root

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

Lemma 7.5.1: Properties of Positive Operators

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

- T is positive
- T is self-adjoint and all the eigenvalues of T are nonnegative
- T has a positive square root
- T has a self-adjoint square root
- there exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$.

Corollary 7.5.1

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

Using more tools, one can further state that if V is a complex vector space and T an invertible operator, then T has a square root.

Isometries

We can impose one further condition on isomorphisms to strengthen the relation.

Definition 7.5.3: Isometry

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

$$\|Av\| = \|v\|$$

for all $v \in V$.

Note that we haven't explicitly stated that isometries are isomorphisms. It turns out that this follows directly from the definition above.

Proposition 7.5.1: Characterization of Isometries

Let $A \in \mathcal{L}(V)$. The following are equivalent:

- A is an isometry
- $\langle Au, Av \rangle = \langle u, v \rangle$ for all $u, v \in V$
- $Ae_1, \ldots, Ae_n$ is orthonormal for all choices of orthonormal vectors $e_1, \ldots, e_n \in V$
- There exists an orthonormal basis $e_1, \ldots, e_n \in V$ such that $Se_1, \ldots, Se_n$ is orthonormal
- $S^*S = SS^* = I$
- S^* is an isometry
- S is invertible and $S^{-1} = S^*$

Of course, this gives us that isometries are normal. This in fact gives us one more vital characterzation:

Proposition 7.5.2

Let V be a complex inner product space and $A \in \mathcal{L}(V)$. Then there is an orthonormal basis of V consisting of eigenvalues of A with corresponding eigenvalues equal to $|1|$ if and only if A is an isometry.

7.6 Singular Value Decomposition

Let $V, \langle \cdot, \cdot \rangle$ be a finite-dimensional inner product space.

Definition 7.6.1: Singular Values

The **singular values** $s_1, \ldots, s_n$ of T are the eigenvalues of $\sqrt{T^*T}$, each one repeated $\dim E(s_i, \sqrt{T^*T})$ many times.

We can order the singular values $s_1, \ldots, s_n$ of T such that

$$s_1 \geq s_2 \geq \ldots \geq s_n \geq 0.$$

Theorem 7.6.1: Singular Value Decomposition

Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{R}$ of dimension n. Let $T \in \mathcal{L}(V)$ be a linear operator with singular values $s_1, \ldots, s_n$. There exists distinct orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V such that

$$\mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)) = \begin{bmatrix} s_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & s_n \end{bmatrix}$$

Furthermore,

$$T(v) = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n \quad \forall v \in V$$

This is a decomposition. If $h_1, \ldots, h_n$ is the standard basis of $\mathbb{R}^n$, then using the standard inner product, let $A = \mathcal{M}(T, (h_1, \ldots, h_n)$. The SVD theorem is equivalent to saying there are matrices U, Σ, V such that

$$A = U\Sigma V^T$$

This is basically a series of changes of bases.

The matrix V represents the matrix of the operator $T_V : \mathbb{R}^n \to \mathbb{R}^n$ that turns the standard basis vectors $h_1, \ldots, h_n$ to the basis $e_1, \ldots, e_n$. Therefore the columns of the matrix consist of the coefficients of the e_i when expressed in terms of the standard basis $h_1, \ldots, h_n$. This matrix is orthogonal, and thus $V^{-1} = V^T$

The matrix of U has as columns the coefficients of the f_i when they are expressed in terms of the standard basis $h_1, \ldots, h_n$. This matrix is also orthogonal.

Computing an SVD

Computing V

1. Compute $A^T A$
2. Find an orthonormal basis of eigenvectors for $A^T A$.
3. Form V by using these eigenvectors as its columns.

This also gives us $s_1^2, \ldots, s_n^2$.

Computing Σ

1. Compute V. When finding the eigenvectors, keep the eigenvalues (compute eigenvectors of $A^T A$).
2. The eigenvalues are $s_1^2, \ldots, s_n^2$, so take the square root of these (in order) to get the singular values
3. Place these in order along the diagonal

Computing U

1. Calculate Av_i for each $v_i \in V$
2. Divide each Av_i by s_i
3. This yields u_i, which are the columns of U.

Chapter 8

Generalized Eigenvectors and Nilpotent Operators

8.1 Null Spaces of Powers of an Operator

Proposition 8.1.1

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

$$\{0\} = \mathrm{null} T^0 \subset \mathrm{null} T^1 \subset \ldots \subset \mathrm{null} T^k \subset \mathrm{null} T^{k+1} \subset \ldots.$$

Proposition 8.1.2

Suppose $T \in \mathcal{L}(V)$. Suppose m is a nonnegative integer such that $\mathrm{null} T^m = \mathrm{null} T^{m+1}$. Then

$$\mathrm{null} T^m = \mathrm{null} T^{m+1} = \ldots.$$

Proposition 8.1.3

Suppose $T \in \mathcal{L}(V)$, where V is a finite-dimensional vector space. Let $n = \dim V$. Then

$$\mathrm{null} T^n = \mathrm{null} T^{n+1} = \ldots.$$

Proposition 8.1.4

Suppose $T \in \mathcal{L}(V)$, where V is a finite-dimensional vector space.. Let $n = \dim V$. Then

$$V = \mathrm{null} T^n \oplus \mathrm{range} T^n.$$

8.2 Generalized Eigenvectors

Some operators do not have enough eigenvectors to completely describe it. Thus we introduce a generalized eigenvectors which will allow us to describe operators more fully (no inner product means no guarantee of spectral theorem, hence this expands it).

Definition 8.2.1: Generalized Eigenvector

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

$$(T - \lambda I)^j v = 0$$

for some positive integer j.

Definition 8.2.2: Generalized Eigenspace

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in F$. The **generalized eigenspace** of T corresponding to λ, denoted $G(\lambda, T)$, is defined to be the set of all generalized eigenvectors of T corresponding to λ, along with the 0 vector.

Proposition 8.2.1

Suppose $T \in \mathcal{L}(V)$ where V is a finite-dimensional vector space and $\lambda \in F$. Then $G(\lambda, T) = \mathrm{null}(T - \lambda I)^{\dim V}$.

Proposition 8.2.2

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

8.3 Nilpotent Operators

Definition 8.3.1

An operator is called **nilpotent** if some power of it equals 0.

Proposition 8.3.1

Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then $N^{\dim V} = 0$.

Proposition 8.3.2

Suppose N is a nilpotent operator on V. Then there is a basis of V with respect to which the matrix of N has the form

$$\begin{bmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{bmatrix}$$

or in other words, all entries on and below the diagonal are 0.

We can actually construct a basis that corresponds to a nilpotent operator. This will be important when we get to Jordan form.

Proposition 8.3.3

Suppose $N \in \mathcal{L}(V)$ is nilpotent. Then there exist vectors $v_1, \ldots, v_n \in V$ and nonnegative integers $m_1, \ldots, m_n$ such that

- $\left\{N^{m_1} v_1, \ldots, N^2, N v_1, v_1, N^{m_2} v_2, \ldots, N^{m_n} v_n, \ldots, v_n\right\}$ is a basis of V
- $N^{m_1+1} v_1 = \ldots = N^{m_n+1} v_n = 0$.

8.4 Decomposition of an Operator

Description of Operators on Complex Vector Spaces

Proposition 8.4.1

Suppose $T \in \mathcal{L}(V)$ and $p \in P(F)$. Then $\text{null}\, p(T)$ and $\text{range}\, p(T)$ are invariant under T.

Theorem 8.4.1: Description of operators on complex vector spaces

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. Then

- $V = G(\lambda_1, T) \oplus \ldots \oplus G(\lambda_m, T)$
- each $G(\lambda_j, T)$ is invariant under T
- each $(T - \lambda_j I)\,|_{G(\lambda_j, T)}$ is nilpotent

Proposition 8.4.2

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then there is a basis of V consisting of generalized eigenvectors of T.

Multiplicity of an Eigenvalue

Definition 8.4.1: Multiplicity

Suppose $T \in \mathcal{L}(V)$. The **multiplicity** of an eigenvalue λ of T is defined to be the dimension of the corresponding generalized eigenspace $G(\lambda, T)$.

In other words, the multipicity of an eigenvalue λ of T equals $\dim \text{null}(T - \lambda I)^{\dim V}$.

Proposition 8.4.3

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Then the sum of the multiplicities of all the eigenvalues of T equals $\dim V$.

Block Diagonal Matrices

Proposition 8.4.4: First approximation to Jordan Form

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T, with multiplicities $d_1, \ldots, d_m$. Then there is a basis of V with respect to which T has a block diagonal matrix of the form

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}$$

where each A_j is a d_j-by-d_j upper-triangular matrix of the form

$$\begin{bmatrix} \lambda_j & & * \\ & \ddots & \\ 0 & & \lambda_j \end{bmatrix}$$

8.5 Jordan Form

Definition 8.5.1: Jordan Basis

Suppose $T \in \mathcal{L}(V)$, where V is a finite-dimensional vector space. A basis of V is called a **Jordan basis**for T if with respect to this basis, T has a block diagonal matrix

$$\begin{bmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_p \end{bmatrix}$$

where each A_j is an upper-triangular matrix of the form

$$A_j = \begin{bmatrix} \lambda_j & 1 & \dots & 0 \\ \vdots & \lambda_j & \ddots & \vdots \\ & & \ddots & 1 \\ 0 & \dots & & \lambda_j \end{bmatrix}$$

The matrix is said to be in **Jordan form**, and the A_j are called **Jordan blocks**.

Note: Some define the Jordan blocks with the 1's on the subdiagonal. This corresponds to reversing the order of the basis elements in each Jordan block.

8.6 Characteristic and Minimal Polynomials

Cayley-Hamilton Theorem

Definition 8.6.1: Characteristic Polynomial

Let V be a vector space and T an operator on V. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T with multiplities $d_1, \dots, d_m$. Then we denote the **characteristic polynomial** of T by

$$(x-\lambda_1)^{d_1} \dots (x-\lambda_m)^{d_m}.$$

The characteristic polynomial turns out to be a convenient characterization and a powerful tool.

Proposition 8.6.1

Let V be a vector space, T an operator on V, and q the characteristic polynomial of T. Then q has degree $\dim(V)$ and the zeroes of q are exactly the eigenvalues of T.

This trivially holds by the definition. Note that now we need to make an important distinction– if V is a real-valued vector space instead of a complex vector space, then the eigenvalues of T are instead the real zeroes of q. We need to make a couple adjustments along the way for these kinds of cases.

Theorem 8.6.1: Cayley-Hamilton Theorem

Let V be a vector space, T an operator on V, and q the characteristic polynomial of T. Then $q(T) = 0$.

Cayley-Hamilton. ■

Minimal Polynomial

Definition 8.6.2: Minimal Polynomial

Suppose $T \in \mathcal{L}(V)$. Then there is a unique monic polynomial p such that $p(T) = 0$. We refer to p as the **minimal polynomial**.

If there is a non-monic polynomial q such that $q(T) = 0$, then $q = ps$ for some polynomial s. Of course, the reverse holds as well.

Proposition 8.6.2

Let V be a vector space and T an operator over V. Then the characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.

Furthermore, we have that the zeroes of the minimal polynomial of T are precisely the eigenvalues of T. This makes it clear that our minimal polynomial looks similar to our characteristic polynomial of T, but with smaller exponents.