MATH 354 Lecture Notes

Taught by Dr. Berg

Shaquille Que

Fall 2018

These are my notes for Rice University's MATH 354: Honors Linear Algebra, taught by Dr. Jennifer Berg. This file was created in LaTeXand uses Evan Chen's `evan.sty` package. Any mistake herein is my own. Please let me know of any errors by emailing me at stq1@rice.edu.

Contents

§1 August 20, 2018

§1.1 Fields

Definition

A **field** is a set $\mathbb{F}$ with two operations, $+$ and $\times$, taking $\mathbb{F} \times \mathbb{F} \to \mathbb{F}$. ex. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$

Properties

For $a, b, c \in \mathbb{F}$, these operations satisfy

1. **Associativity**
   - $a + (b + c) = (a + b) + c$.
   - $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
2. **Commutativity**
   - $a + b = b + a$.
   - $a \cdot b = b \cdot a$.
3. **Identities**
   - $\exists 0 \in \mathbb{F}$, $a + 0 = a$.
   - $\exists 1 \in \mathbb{F}$, $a \cdot 1 = a$.
4. **Inverses**
   - $\exists - a \in F$ satisfying $a + (-a) = 0$.
   - $\exists a^{-1} \in F$ (for $a \neq 0$) satisfying $a \cdot (a^{-1} = 1$.
5. $\mathbf{0 \neq 1}$.
6. **Distributivity**
   $a \cdot (b + c) = a \cdot b + a \cdot c$

More Examples of Fields

- $\mathbb{F} = \{0, 1\}$, sometimes denoted $\mathbb{F}_2$.
- $\mathbb{F} = \{0, 1, \alpha, \alpha^2\}$ or $\mathbb{F}_4$.
  $\alpha \neq 1 \implies \alpha^2 + \alpha + 1 = 0$.

$+$	0	1	α	α^2
0	0	1	α	α^2
1	1	0	α^2	α
α	α	α^2	0	1
α^2	α^2	α	1	0

$\times$	0	1	α	α^2
0	0	0	0	0
1	0	1	α	α^2
α	0	α	α^2	1
α^2	0	α^2	1	α

Theorem 1.1 (Uniqueness of the Identity Element)

Proof. Let $\mathbb{F}$ be a field and suppose $0, 0' \in \mathbb{F}$ are additive identities. Then

$$0 = 0 + 0' = 0'. \quad \square$$

§2 August 22, 2018

§2.1 Some Field Results

Theorem 2.1

For any $a \in \mathbb{F}, a \cdot 0 = 0$

Proof. Let $a \in \mathbb{F}$. Let $b = a \cdot 0$ and $-b$ be the additive inverse of b.

$$\begin{aligned} 0 + 0 &= 0 \\ a \cdot (0 + 0) &= a \cdot 0 \\ a \cdot 0 + a \cdot 0 &= a \cdot 0 \\ a \cdot 0 + (a \cdot 0 + -b) &= a \cdot 0 + (-b) \\ a \cdot 0 + 0 &= 0. \end{aligned}$$

□

Proposition 2.2

For any $a \in \mathbb{F}, (-1) \cdot a = -a$.

§2.2 Vector Spaces

Introduction

Let $\mathbb{F}$ be a field. Consider a list of ordered elements of $\mathbb{F}$ of length n: $(x_1, \dots, x_n), x_i \in \mathbb{F}$. Define $\mathbb{F}^n \stackrel{\text{def}}{=} \{(x_1, \dots, x_n) \mid x_i \in \mathbb{F}\}$. Then $\mathbb{F}$ is a **vector space**.

Definition

A vector space over $\mathbb{F}$ is a set V together with two operations:

$$+ : V \times V \quad \cdot : \mathbb{F} \times V$$

and satisfies the following:

Properties

1. **Commutativity.** $u + v = v + u \quad \forall u, v \in V$.
2. **Associativity.**
   $(u + v) + w = u + (v + w) \quad \forall u, v, w \in V$.
   $(\lambda_1 \lambda_2) \cdot v = \lambda_1 (\lambda_2 \cdot v) \quad \forall v \in V$.
3. **Additive Identity.** $\exists \vec{0}$ such that $\vec{0} + v = v \quad \forall v \in V$.
4. **Multiplicative Identity.** $\exists 1_{\mathbb{F}} \in \mathbb{F}$ such that $1_{\mathbb{F}} \cdot v = v \quad \forall v \in V$.
5. **Additive Inverse.** $\exists w \in V$ such that $v + w = \vec{0} \quad \forall v \in V$.
6. **Distributivity.**
   $\lambda(u + v) = \lambda u + \lambda v \quad \forall \lambda \in \mathbb{F}, u, v \in V$
   $(\lambda_1 + \lambda_2) v = \lambda_1 v + \lambda_2 v \quad \forall v \in V$.

§3 August 24, 2018

§3.1 Examples of Vector Spaces

Example 3.1
$(\mathbb{C}^2, \mathbb{C}, +, \cdot)$ $(\mathbb{C}^2, \mathbb{R}, +, \cdot)$ $(\mathbb{F}^\infty, \mathbb{F}, +, \cdot)$

Example 3.2
$V :=$ { polynomials in x of deg ≤ 17 with coefficients in $\mathbb{R}$}, $\mathbb{F} = \mathbb{R}$
where addition is done polynomially and multiplication is done on coefficients.

§4 August 27, 2018

Another Example of a Vector Space

Definition 4.1. $\mathbb{F}^S \overset{\text{def}}{=} \{\text{functions from } S \text{ to } F\} = \{f : S \to \mathbb{F}\}$

Example 4.2
$\mathbb{F}$ a field, S a set

$$V = \mathbb{F}^S := \{\text{functions from } S \text{ to } F\} = \{f : S \to \mathbb{F}\}$$

Addition: $f, g : S \to \mathbb{F}$ $(f + g)(s) \overset{\text{def}}{=} f(s) + g(s)$ for all $s \in S$
Scalar Multiplication: $\lambda \in \mathbb{F}$, $f : S \to \mathbb{F}$ $(\lambda \cdot f)(s) \overset{\text{def}}{=} \lambda \cdot f(s)$

Proof that $V = \mathbb{F}^2$ is a vector space

Additive identity: the zero function
Additive inverse: if $f \in V$, then its additive inverse is $-f : S \to \mathbb{F}$.

§4.1 Consequences of the Axioms of a Vector Space

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

Proposition 4.3
The additive identity of V is unique.

Proposition 4.4
For each $v \in V$, the additive inverse of v is unique.

Proof. Suppose that w and w' are additive inverses of v.

$$\begin{aligned} w &= w + \vec{0} \\ &= w + (v + w') && \text{since } w' \text{ is additive inverse of v} \\ &= (w + v) + w' && \text{by associativity of addition} \\ &= \vec{0} + w' && \text{since } w \text{ is additive inverse of v} \\ &= w' && \square \end{aligned}$$

Notation: From now on, $v - w$ means $v + (-w)$.

Proposition 4.5

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

Proof.

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

Now let's compute $0 \cdot v - 0 \cdot v = (0 \cdot v + 0 \cdot v) - 0 \cdot v$

$$\vec{0} = 0 \cdot v + (0 \cdot v - 0 \cdot v)$$

$$\vec{0} = 0 \cdot v + \vec{0}$$

$$\vec{0} = 0 \cdot v \quad \square$$

Proposition 4.6

$$-v = (-1) \cdot v$$

Hint: Start with $1 + (-1) = 0$ in $\mathbb{F}$

§4.2 Subspaces

Definition

Let $(V, F, +, \cdot)$ be a vector space. A subset $U \subseteq V$ is a **subspace** if $(U, \mathbb{F}, +, \cdot)$ is a vector space in its own right.

Example 4.7

$\{(x_1, x_2, 0) \,|\, x_1, x_2 \in \mathbb{R}\} \subseteq \mathbb{R}^3$ is a subspace.

Criteria for Subspaces

Proposition 4.8 (Sufficient and Necessary Conditions)
Subspaces must satisfy the following:

1. $\vec{0} \in U$.
2. U is closed under addition.
3. U is closed under scalar multiplication.

These conditions ensure that U has a additive identity and that addition and multiplication make sense within U.
e.g. If $u \in U$, then (3) implies $(-1)u \in U$, i.e. $-u \in U$ so U has additive inverses.
Remaining axioms are inherited from V. e.g. associativity
If $u, v, w \in U$, then $(u+v)+w = u+(v+w)$ in V. But (2) $u+v$ and $v+w$ are both elements of U. So $(u+v)+w$ and $u+(v+w)$ are also in U by (2) again.
So equality holds in U!

§5 August 29, 2018

Let $V = \mathbb{R}^{(0,3)}$ and $U = \{f : (0,3) \to \mathbb{R} \mid f \text{ is differentiable and } f'(2) = 0\}$.
Let's see whether U is a subspace of V.

- The zero function $0 : (0,3) \to \mathbb{R}$, $x \to 0$ is differentiable and $0'(2) = 0$ so (1) holds.
- Let $f, g \in U$. Calculus tells us that $f+g$ is differentiable, and $(f+g)'(2) = f'(2) + g'(2) = 0 + 0 = 0$. So $f + g \in U$, so (2) holds.
- Let $\lambda \in \mathbb{R}$ and $f \in U$. Calculus says that $\lambda \cdot f$ is differentiable and $(\lambda f)'(2) = \lambda \cdot f'(2) = \lambda \cdot 0 = 0$. So $\lambda \cdot f \in U$ and so (3) holds.

Thus, U is a subspace of V.

Example 5.1 (More Examples of Subspaces)
$\{\vec{0}\} = \{(0,0,0)\}$ is a subspace.
A line through the origin is a subspace.
A plane through the origin is a subspace.

Theorem 5.2
The only proper subspaces of $\mathbb{R}^3$ are lines through the origin and planes through the origin.

Exercise 5.3. Is $U = \{(a,a) \mid a \in \mathbb{R}, a \geq 0\}$ a subspace of $\mathbb{R}^2$?

Solution. No. U is not closed under scalar multiplication by $\lambda < 0$. □

Exercise 5.4. Is $U = \{(a,a) \mid a \in \mathbb{R}\} \cup \{(-a,a) \mid a \in \mathbb{R}\}$ a subset of $\mathbb{R}^2$?

Solution. No. U is not closed under addition. □

§5.1 Forming New Subspaces

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

Definition 5.5. $U_1 + \ldots, U_m = \{u_1 + \ldots + u_m \mid u_i \in U_i \text{ for } i = 1, \ldots, m\}$.

Example 5.6

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

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

Definition 5.7. If $U_1, \ldots, U_m$ are subspaces of V, then $U_1, \ldots + U_m$ is called the **sum of the subspaces** $U_1, \ldots, U_m$.

Definition 5.8. S is the **smallest subspace** of V containing $U_1, \ldots, U_m$ if all subspaces W of V that contains $U_1, \ldots, U_m$ also contains S.

Proposition 5.9

If $U_1, \ldots, U_m$ are subspaces of V, then $U_1 + \ldots + U_m$ is the smallest subspace of V that contains each of $U_1, \ldots, U_m$.

Proof. First, we show that $U_1 + \ldots + U_m$ is a subspace.

i) $\vec{0} = \vec{0} + \ldots + \vec{0}$

ii) $(u_1 + \ldots + u_m) + (u'_1 + \ldots + u'_m) = (u_1 + u'_1) + \ldots + (u_m + u'_m)$ where $(u_1 + u'_1) \in U_1, \ldots, (u_m + u'_m) \in U_m$.

iii) Check scalar multiplication for yourself.

Now we show that $U_1 + \ldots + U_m$ contains each U_i for $i = 1, \ldots, m$.
e.g. $U_i \in U_1 + \ldots + U_m$. This holds because given any $u_i \in U_i$, we can write it as $u_i = \vec{0} \cdot u_1 + \ldots + u_i + \ldots + \vec{0} \cdot u_m$.

Finally, we show that for all subspaces W of V that contains $U_1, \ldots, U_m$, then W contains $U_1 + \ldots + U_m$.
By assumption, $U_i \subseteq W$ for each i, so $u_i \in W$ for each i. But W is a subspace, so by closure of addition, $u_1 + \ldots + u_m \in W$. Since $u_1 + \ldots + u_m$ was arbitrary, $U_1 + \ldots + U_m \subseteq W$. □

§6 August 31, 2018

§6.1 Direct Sums

Definition 6.1. Suppose $U_1, \ldots, U_m \subseteq V$ are subspaces. If each $v \in U_1 + \ldots + U_m$ ca be written in exactly one way as $v = u_1 + \ldots + u_m$ with each $u_i \in U_i$, then we say $U_1 + \ldots + U_m$ is a **direct sum** and we write $U_1 \oplus \ldots \oplus U_m$.

Proposition 6.2

$V = \mathbb{R}^3$, $\mathbb{F} = \mathbb{R}$. $U_1 = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ $\quad U_2 = \{(0, 0, z) \mid z \in \mathbb{R}\}$ Then

$$U_1 + U_2 \text{ is direct and } U_1 \oplus U_2 = \mathbb{R}^3.$$

Example 6.3

Let $U_3 = \{(0, y, y) \mid y \in \mathbb{R}\}$. Then $U_1 + U_2 + U_3 = \mathbb{R}^3$, but the sum is *not* direct.

$$(0, 0, 0) = (0, 0, 0) + (0, 0, 0) + (0, 0, 0)$$

$$(0, 0, 0) = (0, 1, 0) + (0, 0, 1) + (0, -1, -1)$$

Criterion for Direct Sums

Proposition 6.4 (A Necessary and Sufficient Condition for Direct Sums)

Suppose that $U_1, \ldots, U_m \subseteq V$ are subspaces. Then $U_1 + \ldots + U_m$ is a direct sum if and only if the only way of writing $\vec{0} = u_1 + \ldots + u_m$ with each $u_i \in U_i$ for each i is to take $\vec{0} = \vec{0} + \ldots + \vec{0}$.

Proof. Assume that $U_1 + \ldots + U_m$ is a direct sum.

It is always true that $\vec{0} = \vec{0} + \ldots + \vec{0}$. Since $U_1 + \ldots + U_m$ is direct, this must be the only way to express $\vec{0}$ as a sum of the form $u_1 + \ldots + u_m$ with $u_i \in U_i$. This proves the if condition.

Suppose that the only way to write $\vec{0} = u_1 + \ldots + u_m$ with $u_i \in U_i$ is to take $u_i = 0$ for all i.

Let $v \in U_1 + \ldots + U_m$ and suppose that there are two ways of writing v as $v = u_1 + \ldots + u_m$, $u_i \in U_i$ and $v = u'_1 + \ldots + u'_m$, $u'_i \in U_i$. Then we obtain $\vec{0} = (u_1 - u'_1) + \ldots + (u_m - u'_m)$. Note that $u_i - u'_i$

Note that $u_i - u'_i \in U_i$ since U_i is a subspace. Our hypothesis implies that each $u_i - u'_i = \vec{0}$. Hence, $u_i = u'_i$ for each i. ☐

Lemma 6.5

Let $U, W \subseteq V$ be subspaces. Then $U + W$ is a direct sum $\iff U \cap W = \{\vec{0}\}$.

Proof. Suppose that $U + W$ is a direct sum. Let $v \in U \cap W$. We want to show that $v = \vec{0}$.

$U \cap W$ is a subspace, since U, W are both subspaces. So $-1 \cdot v = -v$ is also in $U \cap W$. But then $\vec{0} = v + (-v)$. Since $U + W$ is direct, $v = \vec{0}$, $-v = \vec{0}$.

Now suppose $U \cap W = \{\vec{0}\}$. We will show that $\vec{0} = u + w$, $u \in U$, $w \in W$. It suffices to show that $u = w = \vec{0}$.

$\vec{0} = u + w$ tells us that $u = -w \in W$. Hence $u \in U \cap W = \vec{0}$. So $u = \vec{0}$. But then $\vec{0} = \vec{0} + w \implies w = \vec{0}$. ☐

Example 6.6

$U = \{(x, y, z) \mid x + y + z = 0\}$ (plane through origin)
$W = \{(x = y = 0\}$ ($z-$axis)
$U \cap W = \{(0, 0, 0)\} \qquad \therefore U + W$ is a direct sum.

§6.2 Span

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

Definition 6.7. A **linear combination** of $v_1, \ldots, v_m \in V$ is a vector of the form

$$a_1 v_1 + \ldots + a_m v_m, \qquad a_i \in \mathbb{F}.$$

Example 6.8

$(17, -4, 2) = 6(2, 1, -3) + 5(1, -2, 4)$ is a linear combination of $(2, 1, -3)$ and $(1, -2, 4)$.

Definition 6.9. The **span** of $v_1, \ldots, v_m \in V$ is the set of *all* linear combinations of $v_1, \ldots, v_m$. (i.e. $\text{span}(v_1, \ldots, v_m) \stackrel{\text{def}}{=} \{a_i v_i + \ldots + a_m v_m \mid a_i \in \mathbb{F}\}$).

By convention, $\text{span}() \stackrel{\text{def}}{=} \{\vec{0}\}$

Proposition 6.10

$\text{span}(v_1, \ldots, v_m)$ is the smallest subspace containing $v_1, \ldots, v_m$.

§7 September 5, 2018

Proof. We do this in three parts.

1. $\text{span}(v_1, \ldots, v_m)$ is a subspace.
   a) $\vec{0} = 0 \cdot v_1 + \ldots + 0 \cdot v_m. \quad \therefore \vec{0} \in \text{span}(v_1, \ldots, v_m)$.
   b) Closed under addition. $(a_1 v_1 + \ldots + a_m v_m) + (b_1 v_1 + \ldots + b_m v_m) = (a_1 + b_1) v_1 + \ldots + (a_m + b_m) v_m \in \text{span}(v_1, \ldots, v_m)$.
   c) Closed under multiplication. $\lambda(a_1 v_1 + \ldots + a_m v_m) = (\lambda a_1) v_! + \ldots + (\lambda a_m) v_m \in \text{span}(v_1, \ldots, v_m)$.
2. $v_1, \ldots, v_m \in \text{span}(v_1, \ldots, v_m)$.
   $v_i = 0 \cdot v_1 + \ldots + 1 \cdot v_i + \ldots 0 \cdot v_m \in \text{span}(v_1, \ldots, v_m$
3. Smallest subspace. Let $W \subseteq V$ be a subspace that contains $v_1, \ldots, v_m$. We need to show that $\text{span}(v_1, \ldots, v_m) \subseteq W$.
   Let $v \in \text{span}(v_1, \ldots, v_m)$. We'll show that $v \in W$.

$$v = a_1 v_1 + \ldots + a_m v_m \qquad a_i \in \mathbb{F}.$$

   By hypothesis, $v_i \in W$ for each i
   $\implies a_i v_i \in W$ for each i
   $\implies a_1 v_1 + \ldots + a_m v_m \in W$
   $\implies v \in W$.

□

Special Situation

When $\text{span}(v_1, \ldots, v_m) = V$, we say that $v_1, \ldots, v_m$ **span** V.

Example 7.1
$V = \mathbb{C}^4$, $\mathbb{F} = \mathbb{C}$.

$$(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1) \text{ span } V.$$

Example 7.2
$P_m(\mathbb{F}) \stackrel{\text{def}}{=} \{\text{polynomials in } x \text{ with degree at most } m \text{ with coefficients in } \mathbb{F}\}$
$= \text{span}(x^m, x^{m-1}, \ldots, x, 1)$
since each element of $P_m(\mathbb{F})$ has the form $a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ with $a_i \in \mathbb{F}$.

§7.1 Finite Dimension

Definition 7.3. V is **finite dimensional** if there is a finite list $v_1, \ldots, v_m \in V$ such that $\text{span}(v_1, \ldots, v_m) = V$.

Definition 7.4. V is **infinite dimensional** if it is not finite dimensional.

Theorem 7.5 (Vector Space of All Polynomials has Infinite Dimension)
$P(\mathbb{F}) \stackrel{\text{def}}{=} \{\text{all polynomials in } x \text{ with coefficients in } \mathbb{F}\}$ is not finite dimensional.

Proof. Suppose for the sake of contradiction that $P(\mathbb{F})$ is finite dimensional. Then by definition, $\exists$ a finite list $p_1, \ldots, p_m \in P(\mathbb{F})$ such that $\text{span}(p_1, \ldots, p_m) = P(\mathbb{F})$.

Let $N = \max\{\deg p_1, \ldots, \deg p_m\}$. Then $\deg(a_1 p_1 + \ldots + a_m p_m) \leq N$ and $x^{N+1} \notin \text{span}(p_1, \ldots, p_m)$, which is a contradiction. □

§7.2 Linear Dependence

Definition 7.6. A list $v_1, \ldots, v_m \in V$ is **linearly independent** if the only way to write

$$\vec{0} = a_1 v_1 + \ldots + a_m v_m$$

is with $a_1 = \ldots = a_m = 0$.

Example 7.7
In $\mathbb{R}^3$, $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ are linearly independent because $(0, 0, 0) = a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = (a_1, a_2, a_3)$. Thus, $a_1 = a_2 = a_3 = 0$.

Definition 7.8. A list of vectors is **linearly dependent** if it is not linearly independent.

Let $v \in \text{span}(v_1, \ldots, v_m)$. Then $v = a_1v_1 + \ldots + a_mv_m$.
Suppose $v = a'_1v_1 + \ldots + a'_mv_m$. Then $\vec{0} = (a_1 - a'_1)v_1 + \ldots + (a_m - a'_m)v_m$.
If $v_1, \ldots, v_m$ is linearly independent, then $a_1 = a'_1, \ldots, a_m = a'_m$
i.e. there is exactly one way to write $v \in \text{span}(v_1, \ldots, v_m)$ as a linear combination of $v_1, \ldots, v_m$.

Proposition 7.9 (Key Result)
The length of any linearly independent list is at most the length of any spanning list of a finite dimensional vector space.

We will present the proof of Proposition 7.9 later on in Section 8.

Example 7.10

$$\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \text{ spans } \mathbb{R}^3$$

Therefore, any four vectors cannot be linearly independent in this vector space.

§8 September 7, 2018

Lemma 8.1 (Linear Dependence Lemma)
Suppose that $v_1, v_2, \ldots, v_m$ is a linearly dependent list in V. Then there exists an index j in $\{1, 2, \ldots, m\}$ such that the following hold:

(a) $v_j \in \text{span}(v_1, v_2, \ldots, v_{j-1})$

(b) $\text{span}(v_1, v_2, \ldots, v_m) = \text{span}(v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$

Proof. By definition of linear dependence, $\exists\, a_1, a_2, \ldots, a_m \in \mathbb{F}$ not all zero such that

$$\vec{0} = a_1v_1 + a_2v_2 + \ldots + a_mv_m.$$

Let $j = \max\{i \mid a_i \neq 0\}$.
Then $a_1v_1 + a_2v_2 + \ldots + a_jv_j = \vec{0}$.
$\implies a_1v_2 + a_2v_2 + \ldots + a_{j-1}v_{j-1} = -a_jv_j$
$\implies -\frac{1}{a_j}(a_1v_1 + a_2v_2 + \ldots + a_{j-1}v_{j-1}) = v_j^{\dagger}$.
So $v_j \in \text{span}(v_1, v_2, \ldots, v_{j-1})$, proving (a).

It is clear that $\text{span}(v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m) \subseteq \text{span}(v_1, v_2, \ldots, v_m)$, so we will show that $\text{span}(v_1, v_2, \ldots, v_m) \subseteq \text{span}(v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$.

Let $v \in \text{span}(v_1, v_2, \ldots, v_m)$. Then
$v = c_1v_1 + c_2v_2 + \ldots + c_jv_j + \ldots + c_mv_m, \quad c_i \in \mathbb{F}$
$= c_1v_1 + c_2v_2 + \ldots + c_{j-1}v_{j-1} + c_j(-\frac{1}{a_j}(a_1v_1 + \ldots + a_{j-1}v_{j-1})) + c_{j+1}v_{j+1} + \ldots + c_mv_m$.
by replacing v_j using †. This shows that $v \in \text{span}(v_1, v_2, \ldots, v_m)$. □

Remark 8.2. If $j = 1$ in the lemma above, then (a) would say that $v_1 \in \text{span}() = \{\vec{0}\}$ i.e. $v_1 = \vec{0}$.

Remark 8.3. $\vec{0}$ can *never* be part of a linearly independent list of vectors.

Now we are ready to prove Proposition 7.9.

Proof of Key Result. First, we show that $v_1 \in \text{span}(u_1, \ldots, u_n) \implies v_1, u_1, \ldots, u_n$ is linearly dependent.

Homework!

Lemma 8.1 tells us that we can remove some u_{j_1} without changing the span, i.e.

$$\text{span}(v_1, u_1, \ldots, u_n) = \text{span}(v_1, u_1, \ldots, u_{j_1-1}, u_{j_1+1}, \ldots, u_n).$$

Note that Lemma 8.1 does not remove v_1. If it did, then $v_1 = \vec{0}$ (because then $v_1 \in \text{span}()$), so then $v_1, \ldots, v_m$ is not linearly independent.

Replace $v_1, u_1, \ldots, u_n$ with $v_1, u_1, \ldots, u_{j_1-1}, u_{j_1+1}, \ldots, u_n$.

Next, we note that $v_2 \in \text{span}(v_1, u_1, \ldots, u_{j_1-1}, u_{j_1+1}, \ldots, u_n) = V$
$\implies v_1, v_2, u_1, \ldots, u_{j_1-1}, u_{j_1+1}, \ldots, u_n$ is linearly dependent.

So by Lemma 8.1, we can remove u_{j_2} without changing the span
i.e. $\text{span}(v_1, v_2, u_1, \ldots, \cancel{u_{j1}}, \ldots, u_n) = \text{span}(v_1, v_2, u_1, \ldots, \cancel{u_{j1}}, \ldots, \cancel{u_{j2}}, \ldots, u_n)$

Again, we note that the lemma does not remove v_2, because if it did, then $v_2 \in \text{span}(v_1) \implies v_1, \ldots, v_m$ is not linearly independent.

By repeating this process, we will end up with $v_1, v_2, \ldots, v_m$, and maybe some $u's$ spanning V. This proves that $m \leq n$. □

§9 September 10, 2018

§9.1 Basis

Definition 9.1. A **basis** of V is a list $v_1, \ldots, v_m$ that spans V and is linearly independent.

Example 9.2

Some examples and non-examples of bases.

1. For $V = \mathbb{F}^n$, $v_1 = (1, 0, \ldots, 0), v_2 = (0, 1, 0, \ldots, 0), \ldots, v_n = (0, 0, \ldots, 0, 1)$ is called the **standard basis** of $\mathbb{F}^n$.
2. $(1, 2, 7), (3, 4, 1)$ is not a basis. They are linearly independent in $\mathbb{R}^3$ but they don't span.
3. $(1, 2), (2, 3)$ are a basis for $\mathbb{R}^3$.
4. $1, x, x^2, \ldots, x^n$ form a basis for $\mathcal{P}_n(F)$.

Proposition 9.3

A list $v_1, \ldots, v_m$ is a basis $\iff$ every $v \in V$ can be written $v = a_1 v_1 + \ldots + a_m v_m, \quad a_i \in \mathbb{F}$ in exactly one way.

Proof. Suppose $v_1, \ldots, v_m$ is a basis of V. Let $v \in V$. Since $v_1, \ldots, v_m$ spans V, $v = a_1 v_1 + \ldots + a_m v_m, \quad a_i \in \mathbb{F}$.

Since $v_1, \ldots, v_m$ is linearly independent, $a_1, \ldots, a_m \in \mathbb{F}$ are unique.

Conversely, suppose $v \in V$ can be written in exactly one way as $v = a_1 v_1 + \ldots + a_m v_m, \quad a_i \in \mathbb{F}$. Then every v is in $\text{span}(v_1, \ldots, v_m)$. Taking $v = \vec{0}$ implies that the zero vector can be expressed in exactly one way as $\vec{0} = a_1 v_1 + \ldots + a_m v_m$. Since taking $a_1, \ldots, a_m = 0, \ldots, 0$ works, this must be the only way to express the zero vector. This implies that the list is linearly independent. □

Theorem 9.4

Every spanning list can be trimmed to a basis.

Sketch of Proof. Let $v_1, \ldots, v_m$ be a spanning list for V, $B := \{v_1, \ldots, v_m\}$
For $j = 1, \ldots, m$, do:
(If $v_j \in \text{span}(\{v_1, \ldots, v_{j-1}\} \cap B)$, then delete v_j from B.)
When the loop finishes, B is a basis. □

Corollary 9.5

Every finite dimensional vector space has a basis.

Proof. Finite dimensional means that there exists a finite spanning list. By Theorem 9.4, we can trim the spanning list to a basis. □

We can also go the other way.

Theorem 9.6

Any linearly independent list in a finite dimensional vector space can be extended to a basis.

Proof. Suppose $v_1, \ldots, v_m \in V$ is linearly independent. Let $w_1, \ldots, w_n \in V$ be a basis (which exists by Corollary 9.5). Apply Theorem 9.4 on $v_1, v_2, \ldots, v_m, w_1, \ldots, w_n$ and note that no v_i's will be removed since $v_1, \ldots, v_m$ is linearly independent. □

Lemma 9.7

Let V be a finite dimensional vector space and $U \subseteq V$ be a subspace. Then U is also finite dimensional.

Proof. If $U = \{\vec{0}\}$, then we are done. Otherwise, $\exists\, v_1 \neq 0 \in U$. If $U = \text{span}(v_1)$, then we are also done. Otherwise, $\exists\, v_2 \in U$ with $v_2 \notin \text{span}(v_1)$. If $U = \text{span}(v_1, v_2)$, then we are done too. We repeat this process and claim that it will end at some point.

Indeed, the process above produces a growing list $v_1, \ldots, v_n$ of linearly independent vectors (by Lemma 8.1) in V. Since V is finite dimensional, V has a spanning list $W_1, \ldots, W_m$.

By Proposition 7.9, we have $n \leq m$, so the process must terminate. □

§10 September 12, 2018

Theorem 10.1

Let V be a finite dimensional vector space and let U be a subspace of V. Then there exists a subspace W such that $V = U \oplus W$.

Proof. Let $u_1, \ldots, u_n \in U$ be a basis of U.
We extend this to a basis for V: $u_1, \ldots, u_n, w_1, \ldots, w_m$.

For each $v \in V$, we know that

$$v = \underbrace{a_1u_1 + \ldots + a_nu_n}_{\in U} + \underbrace{b_1w_1 + \ldots + b_mw_m}_{\in W}.$$

Thus, $V \subseteq U + W$. Since $U + W \subseteq V$, we must have $V = U + W$.

Now we show that $U \cap W = \{\vec{0}\}$. Let $v \in U \cap W$. Then since $v \in U$, we can write $v = a_1u_1 + \ldots + a_nu_n$, and since $v \in W$, we can also write $v = b_1v_1 + \ldots + b_mw_m$. Then subtracting the two equations yields $\vec{0} = a_1u_1 + \ldots + a_nu_n - b_1w_1 - \ldots - b_mw_m$. But since $u_1, \ldots, u_n, w_1, \ldots, w_m$ is a basis of V, then $a_1 = \ldots = a_n = b_1 = \ldots = b_m = 0$, so $v = \vec{0}$. □

§10.1 Dimension

First, we need to prove the following theorem.

Theorem 10.2

Let $(V, \mathbb{F}, +, \cdot)$ be a finite dimensional vector space. Then any two bases of V must have the same length.

Proof. Let $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ be bases of V.

First, simply consider $u_1, \ldots, u_m$ as a linearly independent list and $v_1, \ldots, v_m$ as a spanning list. By Proposition 7.9, $m \leq n$.

Similarly, considering $u_1, \ldots, u_m$ as a spanning list and $v_1, \ldots, v_n$ as a linearly independent list, we have $n \leq m$.

This proves $m = n$. □

Now that we have shown that all bases of a vector space have the same length, we can formally define what a dimension is.

Definition 10.3. The **dimension** of V is the length of a basis of V, denoted by $\dim(V)$.

Example 10.4

We list the dimensions of vector spaces that we have previously encountered.

1. Since the standard basis of $V = \mathbb{F}^n$ has length n, then $\dim V = n$.
2. Since $1, x, x^2, \ldots, x^n$ is a basis of $V = \mathcal{P}_n(\mathbb{F})$, then $\dim(\mathcal{P}_n(\mathbb{F})) = n + 1$.

Proposition 10.5

If U is a subspace of V then $\dim U \leq \dim V$.

Proof. A basis of U is a linearly independent list in V while a basis of V spans V. So Proposition 7.9 $\implies \dim U \leq \dim V$. □

Proposition 10.6

If $\dim V = n$ and $v_1, \ldots, v_n$ is linearly independent, then $v_1, \ldots, v_n$ spans V.

Proof. Extend $v_1, \ldots, v_n$ to a basis. Since $\dim V = n$, then the new list must have length n, Hence, $v_1, \ldots, v_n$ was already a basis. □

Proposition 10.7

If $\dim V = n$ and $v_1, \ldots, v_n$ spans V, then $v_1, \ldots, v_n$ is linearly independent.

Proof. Boring and similar to above. □

Example 10.8

$$\text{Let } U = \{f(x) \in \mathcal{P}_3(\mathbb{R}) \mid f'(-2) = 0\} \subseteq \underbrace{\mathcal{P}_3(\mathbb{R})}_{\text{4 dimensional}},$$

Then $\dim U < 4$.

Proof. $\dim U \le \dim V = 4$. Suppose that $\dim U = 4 = \dim V$.

But this implies that $U = V$ since if $u_1, \ldots, u_4$ is a basis of U, then it is linearly independent in V. But $\dim V = 4$, so $u_1, \ldots, u_4$ is a basis of $V \implies U = V$.

But $f(x) = x$ is in $\mathcal{P}_3(\mathbb{R}) = V$, but $f'(-2) = 1 \ne 0$, hence $U \ne V$, which is a contradiction. □

Proposition 10.9 (Sum of Dimensions)

Let V be a finite dimensional vector space and let $U_1, U_2 \subseteq V$ be subspaces. Then

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

§11 September 14, 2018

Proof. Let $u_1, \ldots, u_m$ be a basis of $U_1 \cap U_2$. We'll extend this basis in two ways:

- to a basis of U_1: $u_1, \ldots, u_m, v_1, \ldots, v_n$; and
- to a basis of U_2: $u_1, \ldots, u_m, w_1, \ldots, w_s$.

We claim that $u_1, \ldots, u_m, v_1, \ldots, v_n, w_1, \ldots, w_s$ is a basis of $U_1 + U_2$.

First, we'll show linear independence. Suppose that $\vec{0} = \sum a_i u_1 + \sum b_j v_j + \sum c_k w_k$.

$$\text{Then } \underbrace{\sum a_i u_1 + \sum b_j v_j}_{\text{in } U_1} = -\sum c_k w_k$$

which means that $-\sum c_k w_k \in U_1 \cap U_2 \implies -\sum c_k w_k = \sum d_i u_i$ since $u_1, \ldots, u_m$ is a basis of $U_1 \cap U_2$. i.e.

$$\vec{0} = \sum c_k w_k + \sum d_i u_i.$$

$\implies$ c_k's and d_i's are all 0 since $u_1, \ldots, u_m, w_1, \ldots, w_s$ are a basis of U_2.
Since c_k's are all 0, we must have

$$\vec{0} = \sum a_i u_i + \sum b_j v_j.$$

But then a_i's and b_j's are 0, since $u_1, \ldots, u_m, v_j, \ldots, v_n$ are a basis of U_1. This proves the claim, and now we can compute the dimensions.
$\dim(U_1 + U_2) = m + n + s = (m+n) + (m+s) - m = \dim(U_1) + \dim(U_2) + \dim(U_1 \cap U_2)$ as desired. □

§11.1 Linear Transformations

Definition 11.1. Let V, W be vector spaces over $\mathbb{F}$. A **linear map** or **linear transformation** from V to W is a function $T : V \to W$ with the following properties:

- **Additivity:** $T(\underbrace{u+v}_{\text{in } V}) = \underbrace{T(u) + T(v)}_{\text{in } W}$.
- **Homogeneity:** $T(\underbrace{\lambda v}_{\text{in } V}) = \underbrace{\lambda \cdot T(v)}_{\text{in } W}$.

Remark 11.2. $T(\vec{0}_V) = \vec{0}_W$ since $T(\vec{0}_V) = T(\vec{0}_V + \vec{0}_V) = T(\vec{0}_V) + T(\vec{0}_V)$
$\implies T(\vec{0}_V) = \vec{0}_W$

Notation: $\mathscr{L}(V, W) = \{T : V \to W \mid T \text{ linear}\}$.

Example 11.3 (Common Linear Maps)
Some common linear maps are:

1. **Zero map:** $0 : V \to W$ mapping $v \mapsto \vec{0}_W$.
2. **Identity map:** $Id_v : V \to V$ mapping $v \mapsto v$ $Id_v \in \mathscr{L}(V, V)$
3. **Differentiation**
4. **Integration**
5. **Shift:** $T : \mathbb{F}^\infty \to \mathbb{F}^\infty$ mapping $(x_1, x_2, \ldots) \mapsto (x_2, x_3, \ldots)$.
6. $T : \mathbb{R}^3 \to \mathbb{R}^2$ mapping $(x, y, z) \mapsto (5x - 2y + z, 7x + y - z)$.

§12 September 17, 2018

Proposition 12.1 (Linear Maps are Vector Spaces)
$\mathscr{L}(V, W)$ is a vector space over $\mathbb{F}$.

Proof. Let $S, T \in \mathscr{L}(V, W)$. Then $(S+T)(v) := S(v) + T(v)$ for all $v \in V$ and for all $\lambda \in \mathbb{F}$, $(\lambda T)(v) := \lambda \cdot T(v)$, so that $\mathscr{L}(V, W)$ is closed under addition and scalar multiplication. Moreover, the additive identity is just the zero map.

$\mathscr{L}(V, W)$ in fact has extra structure: we can also "multiply" maps using composition: $U \xrightarrow{T} V \xrightarrow{S} W$, where T, S are linear maps. $ST : U \to W$ by $ST(u) \stackrel{\text{def}}{=} S(T(u))$.

The properties of this composition are:

- **Associativity.**
  $U \xrightarrow{T_1} V \xrightarrow{T_2} W \xrightarrow{T_3} X$, where T_1, T_2, T_3 are linear maps, then $T_3(T_2T_1) = (T_3T_2)T_1$.

- **Identity.**
  $\mathrm{Id}_w \cdot T = T \cdot \mathrm{Id}_v = T$

- **Distributivity.**
  $S_1, S_2 \in \mathscr{L}(V, W), T \in \mathscr{L}(W, X)$
  $T(S_1 + S_2) = T \cdot S_1 + T \cdot S_2$. □

Proposition 12.2
Linear maps are determined by what they do to a basis of a domain.

Let $T : V \to W$ is a linear map and suppose $v_1, \ldots, v_n$ is a basis of V. Suppose we know what the elements of $T(v_1), T(v_2), \ldots, T(v_n) \in W$ are.
Let $v \in V$. Then $v = a_1v_1 + \ldots + a_nv_n, a_i \in \mathbb{F}$ and $T(v)$ is already defined:

$$\begin{aligned} T(v) &= T(a_1v_1 + \ldots + a_nv_n) \\ &= T(a_1v_1) + \ldots + T(a_nv_n) \\ &= a_1T(v_1) + \ldots + a_nT(v_n). \end{aligned}$$

Proposition 12.3 (Axler 3.5)
Suppose $v_1, \ldots, v_n \in V$ is a basis of V and $w_1, \ldots, w_n$ is *any* list of vectors in W. Then there is *exactly* one linear map $T : V \to W$ sending $v_i \mapsto w_i$ for each i.

More precisely, for any $v = a_1v_1 + \ldots + a_nv_n \in V$, the linear map in Proposition 12.3 is the map $T(v) = a_1w_1 + \ldots + a_nw_n$.

§12.1 Kernels

Definition 12.4. Let $T \in \mathscr{L}(V, W)$. The **kernel**, or null space, of the map T is

$$\ker(T) \stackrel{\text{def}}{=} \{v \in V \mid T(v) = \vec{0}_W\}.$$

Remark 12.5. $T(\vec{0}_V) = \vec{0}_W \implies \vec{0}_V \in \ker(T)$ always.

Example 12.6 (Kernels of Common Linear Maps)
The kernels of some common linear maps:

- The zero map $0 \in \mathscr{L}(V, W)$ has $\ker(0) = V$.
- The identity map $\mathrm{Id}_V \in \mathscr{L}(V, W)$ has $\ker(\mathrm{Id}_v) = \{\vec{0}\}$.
- The differentiation map $D \in \mathscr{L}(V, W)$ has $\ker D = \{c \mid c \in \mathbb{R}\}$.
- The shift map shift $\in \mathscr{L}(F^\infty, \mathbb{F}^\infty)$ has $\ker(\text{shift}) = \{(a, 0, 0, \ldots,) \mid a \in \mathbb{F}\}$.

Lemma 12.7 (Kernel of a Linear Map is a Subspace)
If $T \in \mathscr{L}(V, W)$, then $\ker(T)$ is a subspace of V.

Proof. Check each of the properties. □

Definition 12.8. A map is **injective** if $T(u) = T(v) \implies u = v$.

§13 September 19, 2018

Proposition 13.1 (Injective Maps Have a Trivial Kernel)
A linear map is injective $\iff \ker(T) = \{\vec{0}_V\}$.

Proof. Suppose T is injective and let $v \in \ker T$. We know that $T(v) = \vec{0}_W = T(\vec{0}_V)$, but since T is injective, $v = \vec{0}_V$.

Conversely, suppose $\ker T = \{\vec{0}_V\}$. Suppose $\exists\, u, v \in V$ such that $T(u) = T(v)$. Then $\vec{0}_W = T(u) - T(v) = T(u - v)$, but since the kernel of T only contains $\vec{0}_V$, then $u - v = \vec{0}_V \implies u = v$, so T is injective. □

§13.1 Images

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

Definition 13.2. The **image** of T is $\operatorname{Im}(T) = \{w \in W \mid w = T(v), v \in V\}$.

Claim 13.3. $\operatorname{Im}(T) \subseteq W$ is a subspace, e.g. $\vec{0}_W \in \operatorname{Im}(T)$, since $T(\vec{0}_V) = \vec{0}_W$.

Example 13.4
Some examples:

1. $0 \in \mathscr{L}(V, W), \quad \operatorname{Im} 0 = \{\vec{0}_W\}$
2. $\operatorname{Id}_V \in \mathscr{L}(V, V), \quad \operatorname{Im}(\operatorname{Id}_V) = V$
3. Differentiation $D \in \mathscr{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R})), \quad \operatorname{Im}(D) = \mathcal{P}(\mathbb{R})$

Definition 13.5. A map $T \in \mathscr{L}(V, W)$ is **surjective** if for every $w \in W$, there exists $v \in V$ such that $T(v) = w$, i.e. $\operatorname{Im}(T) = W$.

Example 13.6
$D = \mathscr{L}(\mathcal{P}_5(\mathbb{R}), \mathcal{P}_5(\mathbb{R}))$ differentiation is *not* surjective since $x^5 \notin \operatorname{Im} D$.

Let V be a finite dimensional vector space over $\mathbb{F}$ and suppose $T \in \mathscr{L}(V, W)$.

Theorem 13.7 (Rank-Nullity Theorem)

$$\dim V = \dim(\ker T) + \dim(\operatorname{Im} T)$$

The Rank-Nullity Theorem also tells us that $\operatorname{Im} T$ is finite dimensional.

Proof. Let $u_1, \ldots, u_n$ be a basis for $\ker T$. Note that $\dim(\ker T) = n$.

We extend this to a basis for $V : u_1, \ldots, u_n, v_1, \ldots, v_m$, so $\dim V = n + m$.

We claim that $T(v_1), \ldots, T(v_m)$ is a basis for $\operatorname{Im}(T)$, in particular $\operatorname{Im}(T)$ is finite dimensional.

First, we show that the list spans $\operatorname{Im} T$. let $w \in \operatorname{Im} T$. Then $w = T(v)$ for some $v \in V$. We have $v = a_1 u_1 + \ldots + a_n u_n + b_1 v_1 + \ldots + b_m v_m$, so $w = T(v) = a_1 T(u_1) + \ldots + a_n T(u_n) + b_1(T(v_1) + \ldots + b_m T(v_m) = b_1 T(v_1) + \ldots + b_m T(v_m)$ since the u_i's are in $\ker T$. This means that $w \in \operatorname{span}(T(v_1), \ldots, T(v_m))$.

Now, we show that the list is linearly independent. Suppose $\vec{0}_W = c_1 T(v_1) + \ldots + c_m T(v_m) = T(c_1 v_1) + \ldots + T(c_m v_m) = T(c_1 v_1 + \ldots + c_m v_m)$, i.e. $c_1 v_1 + \ldots + c_m v_m \in \ker T$. Since $\ker T$ has the u_i's as a basis, we have $c_1 v_1 + \ldots + c_m v_m = d_1 u_1 + \ldots + d_n u_n$, then $\vec{0}_V = d_1 u_1 + \ldots + d_n u_n - c_1 v_1 - \ldots - c_m v_m$. Since $u_1, \ldots, u_n, v_1, \ldots, v_n$ are a basis of V by assumption, this forces $c_1 = \ldots = c_m = d_1 \ldots = d_n = 0$, so the only way to write $\vec{0}_W$ as a linear combination of the $T(v_i)$'s is when all coefficients are zero, i.e. $T(v_1), \ldots, T(v_m)$ is a basis of $\operatorname{Im} T$. □

§14 September 21, 2018

Corollary 14.1 (Linear Maps to Lower Dimensions Are Not Injective)

Suppose $\dim V > \dim W$. Then any linear map $T : V \to W$ cannot be injective.

Proof. By the Rank-Nullity Theorem,

$$\dim \ker T = \dim V - \dim \operatorname{Im} T \geq \dim V - \dim W > 0$$

So $\ker T \neq \{\vec{0}_V\}$. □

Corollary 14.1 can be applied to systems of linear equations:

$$\begin{matrix} a_{1,1}x_1 + \ldots + a_{1,n}x_n = 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n = 0 \end{matrix}$$

We observe that one solution of this is $x_1 = \ldots = x_n = 0$.

We can rephrase this as: $T : \mathbb{R}^n \to \mathbb{R}^m \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1}x_1 + \ldots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n \end{pmatrix}$

Is $\ker T = \{\vec{0}_V\}$?

If $n > m$, then $\ker T \neq \{\vec{0}_V\}$. More variables than the number of equations implies that we get solutions to the homogeneous systems!

§15 September 24, 2018

Corollary 15.1 (Linear Maps to Higher Dimensions Are Not Surjective)

Suppose $\dim V < \dim W$. Then any linear map $T : V \to W$ cannot be surjective.

Proof. $\dim(\operatorname{Im} T) = \dim V - \dim \ker T \leq \dim V < \dim W$. So $\operatorname{Im} T \neq W \implies T$ is not surjective. □

§15.1 Matrices

The Matrix of Linear Map

Let V, W be finite dimensional vector spaces and $T \in \mathscr{L}(V, W)$.
We choose bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W.Then
$T(v_1) \in W$ means $T(v_1) = a_{1,1}w_1 + a_{2,1}w_2 + \dots + a_{m,1}w_m$
$T(v_2) \in W$ means $T(v_2) = a_{1,2}w_1 + a_{2,2}w_2 + \dots + a_{m,2}w_m$
$\vdots$
$T(v_n) \in W$ means $T(v_n) = a_{1,n}w_1 + a_{2,n}w_2 + \dots + a_{m,n}w_m$
Recall that a linear map is determined by what it does to a basis. So the array of numbers

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

$$\begin{matrix} \uparrow & \uparrow & & \uparrow \\ T(v_1) & T(v_2) & \dots & T(v_n) \end{matrix}$$

encodes T. This is called the matrix of T with respect to $v_1, \dots, v_n$ and $w_1, \dots, w_m$.
Notation: $M(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ or $M(T)$.

Example 15.2 (Importance of the Choice of Bases)
$T : \mathbb{F}^2 \to \mathbb{F}^3$ mapping $(x, y) \mapsto (x + 3y, 2x + 5y, 7x + 9y)$ has

$$M(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}$$

if we choose the standard bases for $\mathbb{F}^2$ and $\mathbb{F}^3$. However, if instead we choose the basis $v_1' = (1, 1)$ and $v_2 = (3, -1)$ for V but keep the standard basis for W, we would get $T(v_1') = (4, 7, 16), T(v_2') = (0, 1, 12)$ so

$$M(T) = \begin{pmatrix} 4 & 0 \\ 7 & 1 \\ 16 & 12 \end{pmatrix}$$

Example 15.3 (Differentiation)
For $D \in \mathscr{L}(\mathcal{P}_3(\mathbb{R}), \mathcal{P}_2(\mathbb{R}))$, if we choose the basis $\mathcal{P}_3(\mathbb{R}) : \underbrace{1}_{v_1}, \underbrace{x}_{v_2}, \underbrace{x^2}_{v_3}, \underbrace{x^3}_{v_4}$, $P_2(\mathbb{R}) : \underbrace{1}_{w_1}, \underbrace{x}_{w_2}, \underbrace{x^2}_{w_3}$,

$$\begin{aligned} D(1) &= 0 &&= 0 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 \\ D(x) &= 1 &&= 1 \cdot w_1 + 0 \cdot w_2 + 0 \cdot w_3 \\ D(x^2) &= 2x &&= 0 \cdot w_1 + 2 \cdot w_2 + 0 \cdot w_3 \\ D(x^3) &= 3x^2 &&= 0 \cdot w_1 + 0 \cdot w_2 + 3 \cdot w_3 \end{aligned}$$

so

$$M(T) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

§16 September 26, 2018

We know that linear maps can be encoded by matrices, and we also know that $LL(V, W)$ is itself a vector space over $\mathbb{F}$. This means that linear maps can be added, scaled, and composed. Below we examine what happens to the corresponding matrices under these operations.

1. Addition: $S, T \in \mathscr{L}(V, W) \implies (S + T)(v) = S(v) + T(v)$.
   We choose bases $v_1, \ldots, v_n$ of V and $w_1, \ldots, w_m$ of W.
   Looking at the k^{th} column, if we have $S(v_k) = a_{1,k}w_1 + \ldots + a_{m,k}w_m$ and $T(v_k) = b_{1,k}w_1 + \ldots + b_{m,k}w_m$, then

$$M(S) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} \qquad M(T) = \begin{pmatrix} b_{1,1} & \ldots & b_{1,n} \\ \vdots & \vdots & \vdots \\ b_{m,1} & \ldots & b_{m,n} \end{pmatrix}$$

$$\begin{matrix} \uparrow & & \uparrow & & \uparrow & & \uparrow \\ S(v_1) & \ldots & S(v_n) & & T(v_1) & \ldots & T(v_n) \end{matrix}$$

   $(S + T)(v_k) = S(v_k) + T(v_k)$
   $= (a_{1,k} + b_{1,k})w_1 + \ldots + (a_{m,k} + b_{m,k})w_m$, so

$$M(S + T) = \begin{pmatrix} a_{1,1} + b_{1,1} & \ldots & a_{1,n} + b_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} + b_{m,1} & \ldots & a_{m,n} + b_{m,n} \end{pmatrix}$$

   i.e. we can define addition of matrices entry by entry $M(S) + M(T) \stackrel{\text{def}}{=} M(S + T)$

2. Scalar Multiplication: Then $(\lambda T)(v_k) = \lambda T(v_k) = \lambda(a_{1,k}w_1 + \ldots + a_{m,k}w_k) = \lambda a_{1,k}w_1 + \ldots + \lambda a_{m,k}w_k$

$$M(\lambda T) = \begin{pmatrix} \lambda a_{1,1} & \ldots & \lambda a_{1,n} \\ \vdots & \vdots & \vdots \\ \lambda a_{m,1} & \ldots & \lambda a_{m,n} \end{pmatrix}$$

$\lambda M(T) \stackrel{\text{def}}{=} M(\lambda T)$.

3. Composition:

$$U \xrightarrow{T} V \xrightarrow{S} W$$

We choose bases $(u_1, \ldots, u_r), (v_1, \ldots, v_n), (w_1, \ldots, w_m)$ of U, V, W respectively.

$$T(u_j) = b_{1,j}v_1 + \ldots + b_{n,j}v_n \qquad S(v_k) = a_{1,k}w_1 + \ldots + a_{m,k}w_m$$

$$M(T) = \begin{pmatrix} b_{1,1} & \ldots & b_{1,n} \\ \vdots & \vdots & \vdots \\ b_{m,1} & \ldots & b_{m,n} \end{pmatrix} \quad M(S) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$$

$$\begin{aligned} &S \circ T(u_j) = S(T(u_j)) \\ &= S(b_{1,j}v_1 + \ldots + b_{n,j}v_n) \\ &= S(b_{1,j}v_1) + \ldots + S(b_{n,j}v_n) \\ &= b_{1,j}S(v_1) + \ldots + b_{n,j}S(v_n) \\ &= b_{1,j}(a_{1,1}w_1 + \ldots + a_{m,1}w_m) + \ldots + b_{n,j}(a_{m,1}w_1 + \ldots + a_{m,n}w_m) \\ &= (b_{1,j}a_{1,1} + b_{2,j}a_{1,2} + \ldots + b_{n,j}a_{1,n})w_1 + \ldots \\ &= \left(\sum_{k=1}^{n} a_{1,k}b_{k,j}\right) w_1 + \ldots + \left(\sum_{k=1}^{n} a_{m,k}b_{k,j}\right) w_m. \end{aligned}$$

Definition 16.1. Define **matrix multiplication** by

$$\begin{pmatrix} \ddots & & \\ & a_{i,j} & \\ & & \ddots \end{pmatrix} \cdot \begin{pmatrix} \ddots & & \\ & b_{i,j} & \\ & & \ddots \end{pmatrix} = \begin{pmatrix} \ddots & & \\ & c_{i,j} & \\ & & \ddots \end{pmatrix}$$

with $c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$

Definition 16.2. $\mathbb{F}^{m,n} \stackrel{\text{def}}{=}$ matrices with m rows, n columns and its entries in $\mathbb{F}$. Define addition and scalar multiplication as we just did. Then $\mathbb{F}^{m,n}$ is a vector space over $\mathbb{F}$.

§17 September 28, 2018

Suppose A is an $m \times n$ matrix over $\mathbb{R}$.

$$A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}.$$

Then $A = M(T_A)$, where $T_A : \mathbb{R}^n \to \mathbb{R}^m$ with standard bases $(0, \ldots, \underbrace{1}_{k^{\text{th}} \text{ entry}}, \ldots, 0) \mapsto$ $(a_{1,k}, a_{2,k}, \ldots, a_{m,k})$

Proposition 17.1 (Matrix Multiplication is Associative)
If A, B, C are matrices, then $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Proof. We'll view each matrix in terms of linear maps.
$A \longleftrightarrow T_A : \mathbb{R}^n \to \mathbb{R}^m$
$B \longleftrightarrow T_B : \mathbb{R}^p \to \mathbb{R}^n$
$C \longleftrightarrow T_C : \mathbb{R}^r \to \mathbb{R}^p$, so

$$\begin{aligned}(A \cdot B) \cdot C &= (M(T_A) \cdot M(T_B)) \cdot M(T_C) \\ &= (M(T_A \circ T_B)) \cdot M(T_C) \\ &= M((T_A \circ T_B) \circ T_C) \\ &= M(T_A \circ (T_B \circ T_C)) \\ &= M(T_A) \cdot M(T_B \circ T_C) \\ &= M(T_A) \cdot (M(T_B) \cdot M(T_C)) \\ &= A \cdot (B \cdot C).\end{aligned}$$

□

§17.1 Invertible Linear Maps

Definition 17.2. $T \in \mathscr{L}(V, W)$ is **invertible** if $\exists\, S \in \mathscr{L}(W, V)$ such that $S \circ T = Id_V$ and $T \circ S = Id_W$.

Definition 17.3. If S is the inverse of T, then $S \stackrel{\text{def}}{=} T^{-1}$.

Remark 17.4. Inverses, when they exist, are unique.

Proof. Say $S_1, S_2 \in \mathscr{L}(W, V)$ are both inverses of $T \in \mathscr{L}(V, W)$. Then

$$S_1 = S_1 \circ (T \circ S_2) = (S_1 \circ T) \cdot S_2 = S_2.$$

□

Theorem 17.5 (Invertible Maps are Both Injective and Surjective)
$T \in \mathscr{L}(V, W)$ is invertible $\iff$ T is injective and surjective.

Proof. Suppose T is invertible, and let T^{-1} be its inverse.

Then T is injective. Indeed, if $\exists\, u, v \in V$ such that $T(u) = T(v)$, then

$$u = T^{-1}(T(u)) = T^{-1}(T(v)) = v$$

so T is injective.

T is also surjective. Indeed, let $w \in W$. Then

$$w = T(T^{-1}(w))$$

$\implies w \in \operatorname{Im}(T)$. Since w was arbitrary, T is surjective. □

Now suppose T is both injective and surjective. We'll define an inverse for T.

Let $S : W \to V$ mapping $w \mapsto$ the unique $v \in V$ such that $T(v) = w$. We claim that S is linear.

Indeed, $T(S(w_1) + S(w_2)) = T(S(w_1)) + T(S(w_2)) = w_1 + w_2 = T(S(w_1 + w_2))$ so that $S(w_1) + S(w_2) = S(w_1 + w_2)$, showing that S is additive.

Moreover, $T(\lambda S(w)) = \lambda T(S(w)) = \lambda w$ so $\lambda S(w) = S(T(\lambda S(w))) = S(\lambda w)$, so S is homogenous. Hence, S is linear.

$T \circ S = Id_W$ by definition of S and $S \circ T = Id_V$. Indeed, let $v \in V$. Then $T \circ (S \circ T)(v) = (T \circ S) \circ T(v) = T(v)$. T injective means $(S \circ T)(v) = v$. □

§18 October 1, 2018

§18.1 Isomorphism

Definition 18.1. An invertible map $T \in \mathscr{L}(V, W)$ is called an **isomorphism**.

This is denoted by $V \simeq W$, i.e. $\exists$ an isomorphism $T \in \mathscr{L}(V, W)$.

Let V, W be finite dimensional vector spaces. Then we have the following proposition.

Proposition 18.2 (Isomorphism Means Equal Dimension)
$V \simeq W \iff \dim V = \dim W$.

Proof. Note that $V \simeq W$ means $\exists$ invertible $T \in \mathscr{L}(V, W)$. We apply Theorem 13.7: $\dim V = \dim(\ker T) + \dim(\operatorname{Im} T) = \underbrace{0}_{\text{T is injective}} + \underbrace{\dim W}_{\text{T is surjective}} = \dim W.$ □

Conversely, if $\dim V = \dim W$, then we can construct an invertible $T \in \mathscr{L}(V, W)$. Indeed, let $v_1, \ldots, v_n$ be a basis for V and $w_1, \ldots, w_n$ a basis for W. By Axler 3.5, there exists a unique linear map $T : V \to W$ such that $T(v_i) = w_i$ for $1 \leq i \leq n$.

We show that T is surjective. Indeed, let $w \in W$ such that $w = a_1w_1 + \ldots + a_nw_n$ for $a_i \in \mathbb{F}$. We note that $T(a_1v_1 + \ldots + a_nv_n) = a_1w_1 + \ldots + a_nw_n = w$. Hence, $w \in \operatorname{Im} T \implies W \subseteq \operatorname{Im} T$, and since $\operatorname{Im} T \subseteq W$ by definition, T is surjective.

T is also injective. Indeed, we apply rank-nullity.
$\dim V = \dim(\ker T) + \dim(\operatorname{Im} T) = \dim(\ker T) + \dim W \implies \dim \ker T = 0 \implies \ker T = \{\vec{0}_v\}$. This concludes the proof. □

Example 18.3 ($V \simeq W \iff \dim V = \dim W$)
$\mathcal{P}_3(\mathbb{C})$ and $\mathbb{C}^4$ are isomorphic.

Proof. We choose bases for each vector space and follow the recipe from the previous proof.

$$\begin{aligned} \exists! T : \mathcal{P}_3(\mathbb{C}) \to& \mathbb{C}^4 \\ 1 \mapsto& (1, 0, 0, 0) \\ x \mapsto& (0, 1, 0, 0) \\ x^2 \mapsto& (0, 0, 1, 0) \\ x^3 \mapsto& (0, 0, 0, 1) \end{aligned}$$

and this maps $a_0 + a_1x + a_2x^2 + a_3x^3 \mapsto (a_1, a_2, a_3, a_4)$. □

Example 18.4

Let V, W are finite dimensional vector spaces over $\mathbb{F}$. We choose bases $v_1, \ldots, v_n$ of V and $w_1, \ldots, w_m$ of W. Then

$$\begin{aligned} M : \mathscr{L}(V, W) &\to \mathbb{F}^{m,n} \\ T &\mapsto M(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m)) \\ T_A &\leftarrow\!\shortmid A \end{aligned}$$

is an isomorphism.

Corollary 18.5

$\dim \mathscr{L}(V, W) = \dim \mathbb{F}^{m,n} = m \cdot n$.

Definition 18.6. A linear map $T : V \to V$ is called an **endomorphism** (or an **operator** of V).

Notation: $\mathscr{L}(V) \overset{\text{def}}{=} \mathscr{L}(V, W)$.

Example 18.7

We have the following

1. $T \in \mathscr{L}(\mathcal{P}(\mathbb{R}))$

$$\begin{array}{rcl} T : \mathcal{P}(\mathbb{R}) & \to & \mathcal{P}(\mathbb{R}) \\ f & \mapsto & x^2 f \end{array}$$

   is injective, but not surjective

2. $S : \mathbb{C}^\infty \to \mathbb{C}^\infty$ with $(x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, \ldots)$ is surjective but not injective.

Theorem 18.8 (Injectivity, Surjectivity, and Invertibility of Endomorphisms)

Let V be a finite dimensional vector space. Let $T \in \mathscr{L}(V)$. Then the following are equivalent:

1. T is injective.

2. T is surjective.

3. T is invertible.

Proof. It suffices to show that $(1) \implies (2), (2) \implies (3), (3) \implies (1)$, but the last two have actually been shown in Theorem 17.5.

$(1) \implies (2)$: T is injective means that $\ker T = \{\vec{0}_V\}$. By Rank-Nullity, $\dim V = \dim \ker T + \dim \operatorname{Im} T$ since $T : V \to V$, $\operatorname{Im} T \subseteq V$ and says $\dim V = \dim \operatorname{Im} T$ so $V = \operatorname{Im} T$, i.e. T is surjective. $\square$

§19 October 3, 2018

Example 19.1

Show that for $q \in \mathcal{P}(\mathbb{R})$, there exists $p \in \mathcal{P}(\mathbb{R})$ such that $q = [(x^2 + 2x + 3) \cdot p)]''$

Proof. We first restrict ourselves to the finite dimensional setting. Let $T : \mathcal{P}_m(\mathbb{R}) \to \mathcal{P}_m(\mathbb{R})$ mapping $p \mapsto [(x^2 + 2x + 3) \cdot p]''$. Check that this is a linear map. We need to show that T is surjective, but by Theorem 18.8, it suffices to show that T is injective.

We look at $\ker T$. The only polynomials that have second derivatives equal to 0 are of the form $ax + b$, so $p \in \ker T$ if $p \cdot (x^2 + 2x + 3) = ax + b$. This forces p to be the zero polynomial. Thus, $\ker T = \{\vec{0}\} \implies T$ is injective. □

Remark 19.2. The proof of Example 19.1 actually shows that p is unique.

§19.1 Linear Maps as Matrix Multiplication

Let V be a finite dimensional vector space with basis $v_1, \ldots, v_n$ and $T \in \mathscr{L}(V, W)$. Let $v \in V$. Then $v = c_1 v_1 + \ldots + c_n v_n$ for $c_i \in \mathbb{F}$.

Define $M(v) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. e.g., $V = \mathcal{P}_3(\mathbb{R}), v_1 = 1, v_2 = x, v_3 = x^3, v_4 = x^3$, so $v = 2 - 7x + 5x^3 \implies M(v) = \begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \end{pmatrix}$.

Claim 19.3. $M(T(v)) = M(T) \cdot M(v)$.

Proof. $T(v) = T(c_1 v_1 + \ldots + c_n v_n) = c_1 T(v_1) + \ldots + c_n T(v_n)$, so we have that $M(T(v)) = c_1 \cdot M(T(v_1)) + \ldots + c_n M(T(v_n))$.

$$M(T) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{matrix} T(v_1) & T(v_2) & & T(v_n) \end{matrix}$$

Note that $M(T(v_k)) = k^{th}$ column $= \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix}$. Thus, $M(T(v)) = c_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix} = \begin{pmatrix} c_1 a_{1,1} + \cdots + c_n a_{1,n} \\ \vdots \\ c_1 a_{m,1} + \cdots + c_n a_{m,n} \end{pmatrix} = M(T) \cdot M(v)$. □

§19.2 Product of Vector Spaces

Let $V_1, \ldots, V_m$ be vector spaces over $\mathbb{F}$.

Definition 19.4. $V_1 \times \ldots \times V_m \stackrel{\text{def}}{=} \{(v_1, \ldots, v_n) \mid v_i \in V_i\}$.

Thus, elements of $\mathbb{R}^2 \times \mathbb{R}^3$ look like $((x_1, x_2), (y_1, y_2, y_3))$ and elements of $\mathcal{P}_2(\mathbb{R}) \times \mathbb{R}^2$ look like $(a_0 + a_1 x + a_2 x^2, (z_1, z_2))$.

Define addition and scalar multiplication component-wise, i.e.

$$(v_1, \dots, v_m) + (v_1', \dots, v_m') = (v_1 + v_1', v_2 + v_2', \dots v_m + v_m').$$

This makes the product into a vector space over $\mathbb{F}$.

Proposition 19.5

Let $V_1, \dots, V_m$ be finite dimensional vector spaces. Then $V_1 \times \dots \times V_m$ is also a finite dimensional vector space and

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

Proof. Let $v_{i,j}, 1 \le j \le n$ be a basis for V_i. Then $(0, 0, \dots, \underbrace{v_{i,j_k}}_{i^{\text{th}} \text{ slot}}, 0, 0, \dots, 0)$ is a basis. ☐

§20 October 5, 2018

Connection to Direct Sums

Say $U_1, \dots, U_m \subseteq V$ are all subspaces. Define $\Gamma : U_1 \times \dots \times U_m \to U_1 + \dots + U_m$ mapping $(u_1, \dots, u_m) \mapsto u_1 + \dots + u_m$. Check that this is linear.

Remark 20.1. Γ is injective $\iff$ the only way to write $\vec{0} = u_1 + \dots + u_m$ is to take $u_i = \vec{0}$ for all i, i.e. $U_1 + \dots + U_m$ is a direct sum.

Remark 20.2. Γ is surjective.

Remark 20.3. If V if finite dimensional, then $U_1, \dots, U_m$ are also finite dimensional $\implies U_1 \times \dots \times U_m$ is finite dimensional.

Proposition 20.4 (Additivity of Dimensions in Direct Sums)

If $U_1, \dots, U_m \subseteq V$ are subspaces of a finite dimensional vector space, then $U_1 + \dots + U_m = U_1 \oplus \dots \oplus U_m \iff \dim(U_1 + \dots + U_m) = \dim(U_1) + \dots + \dim(U_m)$.

Proof. The sum is direct $\iff$ Γ is injective $\iff$ $\ker \Gamma = \{\vec{0}\}$. By Rank-Nullity,

$$\dim(U_1 \times \dots \times U_m) = \dim \ker \Gamma + \dim \operatorname{Im} \Gamma$$

$$\iff \dim U_1 + \dots + \dim U_m = 0 + \dim(U_1 + \dots + U_m).$$

☐

§20.1 Affine Subsets

Let V be a vector space over $\mathbb{F}$, $U \subseteq V$ be a subspace and $v \in V$.

Definition 20.5. $v + U \stackrel{\text{def}}{=} \{v + u \mid u \in U\}$ is called the **affine subset** parallel to U.

Example 20.6

$V = \mathbb{R}^2, U = \{(x, 2x) \mid x \in \mathbb{R}\}$ with $v_1 = (3, 1), v_2 = (4, 3)$ has

$$v_1 + U := \{(3 + x, 1 + 2x) \mid x \in \mathbb{R}\}$$
$$v_2 + U := \{(4 + x, 3 + 2x) \mid x \in \mathbb{R}\}.$$

Note that $v_1 + U = v_2 + U$ even if $v_1 \neq v_2$.

Remark 20.7. In the above example, the affine subsets are no longer subspaces because they do not contain $(0, 0)$. In general, affine subsets parallel to U need not be subspaces.

Lemma 20.8 (Parallel Affine Subsets Are Equal or Disjoint)

The following statements are equivalent:

(i) $v_1 + U = v_2 + U$

(ii) $v_1 - v_2 \in U$

(iii) $(v_1 + U) \cap (v_2 + U) \neq 0$

Proof. (ii) $\implies$ (i): Suppose $v_1 - v_2 \in U$. Let $v \in v_1 + U$. So $v = v_1 + u$ for some $u \in U$, i.e. $v = \underbrace{v_2 + (v_1 = v_2) + u}_{\in U} \in v_2 + U$. Thus, $v_1 + U \subseteq v_2 + U$. Similarly, we'll have $v_2 + U \subseteq v_1 + U$. □

(i) $\implies$ (iii): Clear

(iii) $\implies$ (ii): Suppose $w \in (v_1 + U) \cap (v_2 + U)$. Then $w = v_1 + u$ and $w = v_2 + u'$ with $u, u' \in U$. So $\vec{0} = v_1 + u - (v_2 + u') \implies \vec{0} = (v_1 - v_2) + (u - u') \implies v_1 - v_2 = u' - u \in U$. □

§20.2 Quotient Vector Spaces

Definition 20.9. The **quotient** V/U is the set of all affine subsets parallel to U.

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

Example 20.10

$V = \mathbb{R}^2, U = \{(x, 2x) \mid x \in \mathbb{R}\}$ has V/U as the set of all lines parallel to U, i.e. all lines in $\mathbb{R}^2$ with slope 2.

Note that an element of $\mathbb{R}^2/U$ is a whole line in $\mathbb{R}^2$.

Example 20.11

Let $V = \mathbb{R}^3$ and$U$ be a plane through the origin. Then $\mathbb{R}^3/U$ is the set of all planes parallel to U.

V/U is a vector space.

Definition 20.12. Addition: $(v_1 + U) + (w + U) \stackrel{\text{def}}{=} (v + w) + U$.

Definition 20.13. Scalar Multiplication: $\lambda \cdot (v + U) \stackrel{\text{def}}{=} \lambda v + U$.

§21 October 10, 2018

We check that the operations are well-defined (independent of a choice), i.e. if $v + U = v' + U$, and $w + U = w' + U$, then $(v + w) + U = (v' + w') + U$.
$v + U = v' + U \iff v - v' \in U$ and $w + U = w' + U \iff w - w' \in U$ by Lemma 20.8.
Since U is a subspace, $v - v' + w - w' \in U \iff (v + w) - (v' + w') \in U \iff (v + w) + U = (v' + w') + U$.

We also check for scalar multiplication.
$v - v' \in U \iff \lambda(v - v') \in U \iff \lambda v - \lambda v' \in U \iff \lambda v + U = \lambda v' + U$.

So the operations are indeed well-defined.

§21.1 Quotient Map

$$\begin{aligned} \pi : & V \to V/U \\ & v \to v + U \end{aligned}$$

Suppose that V is finite dimensional, then so is U.
Rank-nullity theorem states that $\dim V = \dim \ker \pi + \dim \operatorname{Im} \pi$.

Proposition 21.1
π is surjective $\implies \operatorname{Im} \pi = V/U$.

Proof. $\ker \pi = \{v \in V \mid v + U = \vec{0} + U\} = U$. □

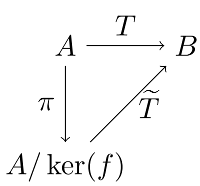
This tells us something important.

Corollary 21.2
$\dim V = \dim U + \dim(V/U)$

§21.2 First Isomorphism Theorem

Let $T \in \mathscr{L}(V, W)$. Define

$$\begin{aligned} \widetilde{T} : & V/\ker T \to \operatorname{Im} T \\ & v + \ker T \mapsto T(v) \end{aligned}$$



We'll prove the following claims:

1. $\widetilde{T}$ is well-defined.

   Proof. If $v_1 + \ker T = v_2 + \ker T$, then $T(v_1) = T(v_2)$ because $v_1 - v_2 \in \ker T \implies T(v_1 - v_2) = \vec{0} \implies T(v_1) = T(v_2)$. □

2. $\widetilde{T}$ is injective.

Proof. Suppose $\widetilde{T}(v + \ker T) = \vec{0}$. Then $T(v) - \vec{0} \implies v \in \ker T \iff v + \ker T = \vec{0} + \ker T$. □

3. $\widetilde{T}$ is surjective.

   Proof. Let $w \in \operatorname{Im} T$. Then $T(v) = w$ for some $v \in V \implies \widetilde{T}(v + \ker T) = T(v) = w$. □

This proves that $V/\ker T$ is isomorphic to $\operatorname{Im} T$.

Proposition 21.3

$V/\ker T \simeq \operatorname{Im} T$

§21.3 Duality

Let V be a vector space over $\mathbb{F}$.

Definition 21.4. $V' \overset{\text{def}}{=} \mathscr{L}(V, \mathbb{F})$ is called the **dual space** of V.

If V is finite dimensional, then $\dim V = \dim \mathscr{L}(V, \mathbb{F})$.

Definition 21.5. $\phi \in \mathscr{L}(V, \mathbb{F})$ is called a **linear functional**.

Example 21.6

Some linear functionals:

1. $V = \mathbb{R}^3, \mathbb{F} = \mathbb{R}, \phi : \mathbb{R}^3 \to \mathbb{R}$.

§22 October 12, 2018

§22.1 Dual Bases

Suppose V is a finite dimensional vector space. Choose a basis $v_1, \ldots, v_n$. Define

$$\begin{aligned} \phi_i : \quad V &\to \quad \mathbb{F} \\ v_j &\mapsto \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Definition 22.1. The shorthand $\phi_i(v_j) = \delta_{i,j}$ is the **Kronecker delta** function.

The list $\phi_1, \ldots, \phi_n$ is called **dual basis** of V' with respect to $v_1, \ldots, v_m$.

Example 22.2

For $V = \mathbb{R}^3, \mathbb{F} = \mathbb{R}$ and standard bases $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$,

$$\begin{aligned} &\phi_1((x_1, x_2, x_3)) \\ =&\phi_1(x_1 e_1 + x_2 e_2 + x_3 e_3) \\ =&x_1 \phi_1(e_1) + x_2 \phi_1(e_2) + \phi_1(e_3) \\ =&x_1 \end{aligned}$$

In general, $\phi_i((x_1, x_2, x_3)) = x_i$.

Theorem 22.3 (Dual Basis is a Basis)
$\phi_1, \ldots, \phi_n$ are a basis of V'.

Proof. Linear independence: Suppose that $a_1\phi_1 + \ldots + a_n\phi_n = 0$ in $\mathscr{L}(V, \mathbb{F})$. Then $a_1\phi_1(v_i) + \ldots + a_n\phi_n(v_i) = 0 \implies a_i = 0$. So the list is linearly independent.
Span: $\dim V' = n$ and $\phi_1, \ldots, \phi_n$ are linearly independent, so they span the vector space. □

§22.2 Dual Maps

Let $T' \in \mathscr{L}(V, W)$.

$$\begin{aligned} \text{Construct } T' : \quad W' &\to V'(T' \in \mathscr{L}(W', V')) \\ \phi &: W \to F \mapsto \phi \circ T : V \to \mathbb{F} \end{aligned}$$

That is $T'(\phi) = \phi \circ T$ as elements of V'

Example 22.4
$D \in \mathscr{L}(\mathcal{P}_5(\mathbb{R}), \mathcal{P}_4(\mathbb{R}))$ and $D' \in \mathscr{L}(\mathcal{P}_4(\mathbb{R})', \mathcal{P}_5(\mathbb{R})')$.
Let's take an actual $\phi \in \mathcal{P}_4(\mathbb{R})'$, e.g.

$$\begin{aligned} \phi : \quad \mathcal{P}_4(\mathbb{R}) &\to \mathbb{R} \\ p &\mapsto \int_0^1 p \end{aligned}$$

Then

$$\begin{aligned} D'(\phi) : \quad \mathcal{P}_3(\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto (\phi \circ D)(f) = \int_0^1 f' = f(1) - f(0) \end{aligned}$$

Claim 22.5. T' is linear.

Proof. $\phi, \psi \in W'$. Then

$$\begin{aligned} T'(\phi + \psi) &= (\phi + \psi) \circ T \\ &= \phi \circ T + \psi \circ T \\ &= T'(\phi) + T'(\psi) \end{aligned}$$

$$\begin{aligned} T'(\lambda \cdot \phi) &= (\lambda \cdot \phi) \circ T \\ &= \lambda \cdot (\phi \circ T) \\ &= \lambda \cdot T'(\phi). \end{aligned}$$

□

Properties of T'

For $T, S \in \mathscr{L}(V, W)$

1. $(T + S)' = T' + S'$

2. $(\lambda \cdot T)' = \lambda \cdot T'$

3. $(S \circ T)' = T' \circ S'$. $U \xrightarrow{T} V \xrightarrow{S} W \implies W' \xrightarrow{S'} V' \xrightarrow{T'} U'$

Proof of (3).

$$\begin{aligned}(S \circ T)'(\phi) &= \phi \circ (S \circ T) \\ &= (\phi \circ S) \circ T \\ &= T'(\phi \circ S) \\ &= T'(S'(\phi)) \qquad\qquad = (T' \circ S')(\phi)\end{aligned}$$ □

We want to describe $\ker T'$ and $\operatorname{Im} T'$ in terms of $\ker T$ and $\operatorname{Im} T$.
Let $U \subseteq V$ be a subspace

Definition 22.6. The **annihilator** $U^0 \subseteq V'$ is

$$U^0 \stackrel{\text{def}}{=} \{\phi \in V' \mid \phi(u) = 0 \forall u \in U\}.$$

Uposhot: $\ker T' = (\operatorname{Im} T)^0$ and $\operatorname{Im} T' = (\ker T)^0$.

Example 22.7

$V = \mathbb{R}^5$ with standard basis $e_1, \ldots, e_5$. V' has dual basis $\phi_1, \ldots, \phi_5$ with $\phi_i(e_j) = \delta_{i,j}$
$U = \operatorname{span}(e_1, e_2) \subseteq V$. Then $U^0 = \operatorname{span}(\phi_3, \phi_4, \phi_5)$.

§23 October 15, 2018

Proof. We show that $U^0 \subseteq \operatorname{span}(\phi_3, \phi_4\phi_5)$. Let $\phi \in U^0$. Then $\phi = a_1\phi_1 + \ldots + a_5\phi_5$ and $\phi(e_1) = 0 = \phi(e_2)$.

Thus $0 = \phi(e_1) = a_1\phi(e_1) + a_2\phi_2(e_1) + \ldots + a_5\phi_1(e_5) = a_1$. Similarly, $a_2 = 0$, so that $\phi = a_3\phi_3 + a_4\phi_4 + a_5\phi_5$

We now show that $\operatorname{span}(\phi_3, \phi_4, \phi_5) \subseteq U^0$.

Let $\phi = a_3\phi_3 + a_4\phi_4 + a_5\phi_5$ be in $\operatorname{span}(\phi_3, \phi_4, \phi_5)$ and let $u \in U$, so $u = a_1e_1 + a_2e_2$.

So $\phi(u) = a_3\phi_3(a_1e_1+a_2e_2)+\ldots+a_5\phi_5(a_1e_1+a_2e_2) = a_3a_1\phi_3(e_1)+a_3a_2\phi_3(e_2)+\ldots = 0$.
So $\phi \in U^0$. □

Exercise 23.1. $U^0 = \{\phi \in V' \mid \phi(u) = 0 \quad \forall u \in U\}$ is a subspace of V'.

Proposition 23.2

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

Proof. Let $u_1, \ldots, u_n$ be a basis of U. Extend this to a basis of V: $u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+m}$. Let $\phi_1, \ldots, \phi_{n+m}$ be a dual basis for V', i.e. $\phi_i(u_j) = \delta_{i,j}$. We claim that $\phi_{n+1}, \ldots, \phi_{n+m}$ is a basis of U^0.

Linear independence: Suppose $a_{n+1}\phi_{n+1} + \ldots + a_{n+m}\phi_{n+m} = 0$ in V'. Plug in $u_{n+j} \implies a_{n+j} = 0$.

span: Let $\phi \in U^0 \in V'$, $\phi = a_1\phi_1 + \ldots + a_{n+m}\phi_{n+m}$ Then by definition of annihilator of U, $\phi(u_1) = \phi(u_2) = \ldots = \phi(u_n) = 0 \implies a_1 = a_2 = \ldots = a_n = 0$. Hence $\phi = a_{n+1}\phi_{n+1} + \ldots + a_{n+m}\phi_{n+m}$, so $U^0 \subseteq \operatorname{span}(\phi_{n+1}, \ldots, \phi_{n+m})$. The other direction follows from the definition of ϕ_j's. Finally, $\dim U^0 = m$. □

Remark 23.3. U^0 is *not* a subspace of V.

Theorem 23.4

Let V, W be finite dimensional vector spaces and $T \in \mathscr{L}(V, W)$, so $T' \in \mathscr{L}(W', V')$. Then

1. $\ker T' = (\operatorname{Im} T)^0$
2. $\operatorname{Im} T' = (\ker T)^0$.

Proof.
1. $\phi \in \ker T' \subseteq W'$
$\iff \phi \circ T = 0$ as maps in V'
$\iff (\phi \circ T)(v) = 0 \quad \forall v \in V$
$\iff \phi(T(v)) = 0 \quad \forall v \in V$
$\iff \phi(w) = 0 \quad \forall w \in \operatorname{Im} T$
$\iff \phi \in (\operatorname{Im} T)^0$.

2. Let $\phi \in \operatorname{Im} T'$. Then $\phi = T'(\psi) = \psi \circ T$ for some $\psi \in W'$.
Let $v \in \ker T$. Then $\phi(v) = \psi(T(v)) = \psi(\vec{0}) = 0 \implies \phi \in (\ker T)^0$.
To finish, it suffices to show that $\dim \operatorname{Im} T' = \dim(\ker T)^0$.

$$\begin{aligned}\dim \operatorname{Im} T' &= \dim W' - \dim \ker T' \\ &= \dim W' - \dim(\operatorname{Im} T)^0 \\ &= \dim W - \dim(\operatorname{Im} T)^0 \\ &= \dim \operatorname{Im} T \\ &= \dim V - \dim \ker T \\ &= \dim(\ker T)^0\end{aligned}$$

□

Remark 23.5. This proof also shows that $\dim \operatorname{Im} T = \dim \operatorname{Im} T'$.

Corollary 23.6

$\dim \ker T' = \dim \ker T + \dim W - \dim V$

Proof.

$$\begin{aligned}\dim \ker T' &= \dim(\operatorname{Im} T)^0 \\ &= \dim W - \dim \operatorname{Im} T \\ &= \dim W - (\dim V - \dim \ker T)\end{aligned}$$

□

Remark 23.7. If $\dim V = \dim W$, then $\dim \ker T = \dim \ker T'$.

Corollary 23.8

Let V, W be finite dimensional vector spaces and $T \in \mathscr{L}(V, W)$.

(1) T is injective $\iff$ T' is surjective.

(2) T is surjective $\iff$ T' is injective.

Proof of (1). T is injective $\iff$ $\ker T = \{\vec{0}\}$ $\iff$ $(\ker T)^0 = V'$ $\iff$ $\operatorname{Im} T' = V'$ $\iff$ T' is surjective. □

§24 The Matrix of the Dual Map

Consider $T : V \to W$ with finite dimensional vector spaces V, W having bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ respectively.

T had dual map $T' : W' \to V'$ with W', V' having the dual bases $\psi_1, \ldots, \psi_m$ such that $\psi_i(w_j) = \delta_{i,j}$ and $\phi_1, \ldots, \phi_n$ such that $\phi_i(v_j) = \delta_{i,j}$ as their bases respectively.

Claim 24.1. $\mathcal{M}(T', (\psi_1, \ldots, \psi_m), (\phi_1, \ldots, \phi_n)) = \mathcal{M}(T, (v_1, \ldots, v_n), (w_1, \ldots, w_m))^t$

Recall: The transpose of a matrix A is A^t where $(A^t)_{i,j} = A_{j,i}$.

e.g. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ $\quad A^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$

Proof of Claim. Say $\mathcal{M}(T) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$ and $\mathcal{M}(T') = \begin{pmatrix} c_{1,1} & \ldots & c_{1,m} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \ldots & c_{n,m} \end{pmatrix}$.

Note that this means $T'(\psi_j) = c_{1,j}\phi_1 + \ldots + c_{n,j}\phi_n$.

Then $(\psi_j \circ T)(v_i) = \psi_j(T(v_i)) = \psi_j(a_{1,i}w_1 + \ldots + a_{m,i}w_m) = a_{1,i}\psi_j(w_1) + \ldots + a_{m,i}\psi_j(w_m) = a_{j,i}$

We also have that $c_{1,j}\phi_1(v_i) + \ldots + c_{n,j}\phi_n(v_i) = 0 + \ldots + c_{i,j} \cdot 1 + \ldots + 0 = c_{i,j}$.

Together, this shows that $c_{i,j} = a_{j,i}$. □

> **Corollary 24.2**
> row rank = column rank

$m \times n$ matrix $A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix} \in \mathbb{F}^{m,n}$.

Definition 24.3. The **column span** of A is $CS(A)\,\mathrm{span}((a_{1,1}, \ldots, a_{m,1}), \ldots, (a_{1,n}, \ldots, a_{m,n})) \subseteq \mathbb{F}^m$.

Definition 24.4. The **row span** is $RS(A)\,\mathrm{span}((a_{1,1}, \ldots, a_{1,n}), \ldots, (a_{m,1}, \ldots, a_{m,n})) \subseteq \mathbb{F}^n$.

§25 October 19, 2018

> **Theorem 25.1** (Rank of a Matrix)
> The column rank of A is equivalent to the row rank of A.

Proof. $A = \mathcal{M}(T_A)$ for $T_A \cdot \mathbb{F}^n \to \mathbb{F}^m$ with respect to the standard bases. $A = \begin{pmatrix} \vdots & \vdots & & \vdots \\ T_A(e_1) & T_A(e_2) & \ldots & T_A(e_n) \\ \vdots & \vdots & & \vdots \end{pmatrix}$, so $CS(A) = \mathrm{span}(T_A(e_1), \ldots, T_A(e_n)) = \operatorname{Im} T_A$ and $\dim CS(A) = \dim \operatorname{Im} T_A$. But since $\dim \operatorname{Im} T_A = \dim \operatorname{Im} T'_A$, and $\dim \operatorname{Im} T'_A = \dim CS(A^t) = \dim RS(A)$. □

§25.1 Invariant Subspaces

Our goal is to understand linear operators $T \in \mathscr{L}(V)$ for a finite dimensional vector space V. We do this by breaking up the problem into smaller pieces, e.g. $V = U_1 \oplus \ldots \oplus U_n$.

So we can first restrict our attention by considering $T|_{U_i} : U_i \to V$. The issue with this method is that $\operatorname{Im} T|_{U_i}$ might lie in U_i. This is where the notion of an invariant subspace comes in.

Definition 25.2. $U \subseteq V$ is an **invariant subspace** under T if $T|_U \in \mathscr{L}(U)$, i.e. $\operatorname{Im} T|_U \subseteq U$ $(\forall\, u \in U, T(u) \in U)$.

Example 25.3

The following are examples of invariant subspaces:

- $U = \{\vec{0}\}$ is trivially an invariant subspace: $T(\vec{0}) = \vec{0}$.
- $U = V$, $\operatorname{Im} T \subseteq V$.
- $U = \ker T$, $u \in \ker T \implies T(u) = \vec{0} \in \ker T$.

A natural question to ask is what invariant subspaces of dimension 1 look like.

To see this, we let $U = \operatorname{span}(v)$, $v \neq 0$. To say that $T(v) \in U$ means that $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$. Conversely, if $T(v) = \lambda v$ for some $v \neq \vec{0}$, then $U = \operatorname{span}(v)$ is a 1-dimensional invariant subspace. When this happens, λ is called an **eigenvalue** of T. If $v \neq \vec{0}$, then v is called an **eigenvector** of T corresponding to λ.

Proposition 25.4

The following are equivalent:

1. λ is an eigenvalue of T.
2. $T - \lambda \cdot \operatorname{Id}_V$ is not injective.
3. $T - \lambda \cdot \operatorname{Id}_V$ is not surjective.
4. $T - \lambda \cdot \operatorname{Id}_V$ is not invertible.

Proof. We'll prove that (1) $\iff$ (2).

$T(v) = \lambda v \iff T(v) - \lambda v = \vec{0} \iff (T - \lambda \operatorname{Id}_V)(v) = 0 \iff T - \lambda \operatorname{Id}_V$ is not injective.

(2) $\iff$ (3) $\iff$ (4) since $(T - \lambda \operatorname{Id}_V) \in \mathscr{L}(V)$. □

Eigenvectors for distinct eigenvalues are linearly i

Proposition 25.5

Let $T \in \mathscr{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues with eigenvectors $v_1, \ldots, v_m$. Then $v_1, \ldots, v_m$ are linearly independent.

§26 October 22, 2018

Proof. Suppose $v_1, \ldots, v_m$ is linearly dependent. Then by the linear dependence lemma, there is a smallest k such that $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$, where $v_1, \ldots, v_{k-1}$ are linearly independent. Then $v_k = a_1 v_1 + \ldots + a_{k-1} v_{k-1} \implies T(v_k) = a_1 \lambda_1 v_1 + \ldots + a_{k-1} \lambda_{k-1} v_{k-1}$. So we have

$$\begin{aligned} 0 = \lambda_k v_k - T(v_k) &= (a_1 \lambda_k - a_1 \lambda_1) v_1 + \ldots + (a_{k-1} \lambda_k v_{k-1}) v_{k-1} \\ &= a_1(\lambda_k - \lambda_1) v_1 + \ldots + a_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1} \end{aligned}$$

but since $v_1, \ldots, v_{k-1}$ are linearly independent, then $a_i(\lambda_k - \lambda_i) = 0$ for $1 \le i \le k-1$. This forces $a_i = 0$ since $\lambda_k \ne \lambda_i$ by assumption. This is a contradiction. ☐

Corollary 26.1 (Number of Eigenvalues $\le$ Dimension)
Let $\dim V = n$. Then $T \in \mathscr{L}(V)$ has at most n distinct eigenvalues.

Proof. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues with eigenvectors $v_1, \ldots, v_m$ Then the proposition says that $v_1, \ldots, v_m$ is a linearly independent list of length m, so $m \le n$ by Proposition 7.9. ☐

§26.1 Polynomials Applied to Linear Operators

Let $T \in \mathscr{L}(V)$. Define:

- $T^m \overset{\text{def}}{=} \underbrace{T \circ \cdots \circ T}_{m \text{ times}} \in \mathscr{L}(V)$
- $T^0 \overset{\text{def}}{=} \mathrm{Id}_V \in \mathscr{L}(V)$
- If T is invertible, then $T^{-m} \overset{\text{def}}{=} \underbrace{T^{-1} \circ \cdots \circ T^{-1}}_{m \text{ times}}$

Let $p(x) \in \mathcal{P}(\mathbb{F})$ be polynomial with $p(x) = a_0 + a_1 x + \ldots + a_n x^n$.

Definition 26.2. $P(T) \overset{\text{def}}{=} a_0 \mathrm{Id}_V + a_1 T + \ldots + a_n T^n \in \mathscr{L}(V)$.

Example 26.3
$V = P(\mathbb{F})$ and $D \in \mathscr{L}(V)$ (differentiation)
Let $p(x) = 7 - 3x + 5x^2$
Then $p(D) = 7\mathrm{Id}_V - 3D + 5D^2 \in \mathscr{L}(V)$.

§26.2 Existence of Eigenvalues

Eigenvalues don't always exist! Their existence usually depends on the field.

Example 26.4 (Eigenvalues Existing in $\mathbb{C}$ but not in $\mathbb{R}$)

$$\begin{aligned} T : \mathbb{F}^2 &\to \mathbb{F}^2 \\ (x, y) &\mapsto (-y, x) \end{aligned}$$

Then an eigenvector corresponding to an eigenvalue λ has the form $T(v) = \lambda v$, i.e. $(-y, x) = \lambda(x, y)$. This means that $\lambda x = -y$ and $\lambda y = x$ so $(\lambda^2 + 1)y = 0$.
Then either $y = 0$ (which forces $x = 0$, i.e. v is not an eigenvector) or $\lambda^2 + 1 = 0$.
If $\mathbb{F} = \mathbb{R}$, there are no eigenvalues.
If $\mathbb{F} = \mathbb{C}$, then $\lambda = i, -i$ are eigenvalues.

Theorem 26.5 (Maps in Vector Spaces over $\mathbb{C}$ have Eigenvalues)
Let $V \neq \{\vec{0}\}$ be a finite dimensional vector space over $\mathbb{C}$. Let $T \in \mathscr{L}(V)$. Then T has at least one eigenvalue.

Proof. Pick $v \neq \vec{0} \in V$. Let $n = \dim V$. Then the $n+1$ vectors $v, T(v), T^2(v), \ldots, T^n(v)$ must be linearly dependent. So there exist $a_0, \ldots, a_n$ not all zero such that

$$a_0 v + a_1 T(v) + \ldots + a_n T^n(v) = \vec{0}.$$

Let $p(x) = a_0 + a_1 x + \ldots + a_n x^n \in \mathcal{P}(\mathbb{C})$. Since $\mathbb{F} = \mathbb{C}$, then $p(x) = a_n(x - \lambda_1)\cdots(x - \lambda_n)$ $\implies a_n(T - \lambda_1 \mathrm{Id}_V)\cdots(T - \lambda_n \mathrm{Id}_V)(v) = \vec{0}$.
So at least one of $T - \lambda_j \mathrm{Id}_V$ must not be injective. This is equivalent to T having an eigenvalue λ_j. □

§27 October 24, 2018

Let V be a finite dimensional vector space and $T \in \mathscr{L}(V)$. Choose a basis $v_1, \ldots, v_n$ of V and use the same basis for the domain and codomain of T:

$$\mathcal{M}(T, (v_1, \ldots, v_n)) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \ldots & a_{n,n} \end{pmatrix}$$

Example 27.1
Find the matrix of $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(x, y) = (-x + y, -6x + 4y)$ using the standard basis.

Solution. $\mathcal{M}(T) = \begin{pmatrix} -1 & 1 \\ -6 & 4 \end{pmatrix}$. □

Example 27.2
Find the matrix of $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(x, y) = (-x + y, -6x + 4y)$ using the basis $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Solution. $\mathcal{M}(T) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. □

Definition 27.3. A matrix is called **upper-triangular** if the entries below the diagonal are zero.

Let $T \in \mathscr{L}(V)$ and choose a basis $v_1, \dots, v_n$

Proposition 27.4

The following are equivalent:

(1) $\mathcal{M}(T)$ is upper triangular

(2) $T(v_j) \in \operatorname{span}(v_1, \dots, v_j)$ for each $j = 1, \dots, n$

(3) $\operatorname{span}(v_1, \dots, v_j)$ is invariant under T for $j = 1, \dots, n$.

Proof. (1) $\iff$ (2): The matrix is upper triangular $\iff$ all coefficients below $a_{j,j}$ is 0, i.e. $T(v_j) = a_{1,j}v_1 + \dots + a_{j,j}v_j + a_{j+1,j}v_{j+1} + \dots + a_{n,j}v_n = a_{1,j}v_1 + \dots + a_{j,j}v_j$.

(3) $\implies$ (2) follows from definition: $x \in \operatorname{span}(v_1, \dots, v_j)$, then so is $T(x)$. In particular, $T(v_j) \in \operatorname{span}(v_1, \dots, v_j)$

(2) $\implies$ (3): Let $x \in \operatorname{span}(v_1, \dots, v_j)$ Then $x = b_1v_1 + \dots + b_jv_j$ and so $Tx = b_1Tv_1 + \dots + b_jTv_j$. We want to show that $Tx \in \operatorname{span}(v_1, \dots, v_j)$, so it suffices to show that $Tv_1, \dots, Tv_j \in \operatorname{span}(v_1, \dots, v_j)$. However, we know that $Tv_i \in \operatorname{span}(v_1, \dots, v_i) \subseteq \operatorname{span}(v_1, \dots, v_j)$. □

Theorem 27.5 (Operators over Complex Space Have Upper Triangular Matrices)

Let V be a finite dimensional vector space over $\mathbb{C}$ and $T \in \mathscr{L}(V)$ Then there exists a basis of V for which $\mathcal{M}(T)$ is upper triangular.

Proof. Induction on $n = \dim V$. For the base case, every 1×1 matrix is trivially upper triangular.

For the inductive step, let $n = \dim V$. Assume that the statement holds for all finite dimensional vector space W over $\mathbb{C}$ such that $\dim W < n = \dim V$. Now let λ be an eigenvalue of T (which exists since $\mathbb{F} = \mathbb{C}$). Let $W = \operatorname{Im}(T - \lambda \mathrm{Id}_V) \subseteq V$.
Claim 1: $\dim W < n = \dim V$.

Proof of Claim 1: $T - \lambda \mathrm{Id}_V$ is not surjective, so $W \subsetneq V$.
Claim 2: W is invariant under T.

Proof of Claim 2: Let $w \in W$. Then $T(w) = T(w) - \lambda w + \lambda w = (T - \mathrm{Id}_V)(w) + \lambda w$ which is in W

Claims 1 and 2 imply that we can restrict T, $T|_W : W \to W$. Since $\dim W < n$, then by the inductive hypothesis, there is a basis $w_1, \dots, w_m$ of W such that $\mathcal{M}(T|_W)$ is upper triangular.

Extend this to a basis of V $w_1, \dots, w_m, v_1, \dots, v_{n-m}$. Now for each $k = 1, \dots, n-m$, $T(v_k) = (T - \mathrm{Id}_V)(v_k) + \lambda v_k$, where $(T - \lambda \mathrm{Id}_V)(v_k) \in W = \operatorname{span}(w_1, \dots, w_m)$ and $v_k \in \operatorname{span}(w_1, \dots, w_m, v_1, \dots, v_k)$. Thus, $T(v_k) \in \operatorname{span}(w_1, \dots, w_m, v_1, \dots, v_k)$.

By Proposition 27.4, $\mathcal{M}(T)$ is upper triangular with respect to the basis $w_1, \dots, w_m, v_1, \dots, v_k$. □

Let V be a finite dimensional vector space over $\mathbb{F}$ and$T \in \mathscr{L}(V)$

Proposition 27.6

Suppose there exists a basis for V for which $\mathcal{M}(T)$ is upper triangular. Then T is invertible $\iff$ 0 is not on the diagonal of $\mathcal{M}(T)$.

Proof. Suppose $\mathcal{M}(T, (v_1, \ldots, v_n)) = \begin{pmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$ with $\lambda_i \neq 0$ for all i.

Then $T(v_1) = \lambda_1 v_1 \implies v_1 = \frac{1}{\lambda_1} T(v_1) = T(\frac{v_1}{\lambda_1}$, so $v_1 \in \operatorname{Im} T$.

$T(v_2) = a_{1,2} v_1 + \lambda_2 v_2 \implies v_2 = T(\frac{v_2}{\lambda_2}) - \frac{a_{1,2}}{\lambda_2} v_1$, so v_2 is also in $\operatorname{Im} T$. Similarly, $v_1, \ldots, v_n \in \operatorname{Im} T$ so T is surjective, and thus invertible.

Conversely, suppose that T is invertible. Note that $\mathcal{M}(T)$ is upper triangular $\implies$ $T(v_1) = \lambda_1 v_1 \neq 0$ since otherwise, $v_1 \in \ker T$, which is impossible since T is injective. Thus $\lambda_1 \neq 0$.

Suppose that $\lambda_j = 0$ for some $2 \leq j \leq n$. Then $T(v_j) = a_{1,j} v_1 + a_{2,j} v_2 + \ldots + a_{j-1,j} v_{j-1} + \cancel{\lambda_j v_j}$ so that $T(v_j) \in \operatorname{span}(v_1, \ldots, v_{j-1})$

Since T is upper triangular, $T(\operatorname{span}(v_1, \ldots, v_j)) \subseteq \operatorname{span}(v_1, \ldots, v_{j-1})$, where the LHS has dimension j and RHS has dimension $j-1$, so T is not injective and thus not invertible, contradiction. □

§28 October 29, 2018

Proposition 28.1

If $\mathcal{M}(T)$ is upper triangular, then the eigenvalues of T are the diagonal entries.

Proof. Suppose $\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & * & * & * & * \\ 0 & \lambda_2 & * & * & * \\ 0 & 0 & \lambda_3 & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$. Let $\lambda \in \mathbb{F}$. Then $\mathcal{M}(T - \lambda \mathrm{Id}_V) = \begin{pmatrix} \lambda_1 - \lambda & * & * & * & * \\ 0 & \lambda_2 - \lambda & * & * & * \\ 0 & 0 & \lambda_3 - \lambda & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_n - \lambda \end{pmatrix}$ λ is an eigenvalue of $T \iff T - \lambda \mathrm{Id}_V$ is not invertible $\iff \lambda - \lambda_i = 0$ for some i. □

§28.1 Eigenspaces

Let V be a finite dimensional vector space over $\mathbb{F}$ and let $T \in \mathscr{L}(V), \lambda \in \mathbb{F}$.

Definition 28.2. The **eigenspace** of T corresponding to λ is $E(\lambda, T) \stackrel{\text{def}}{=} \ker(T - \lambda \mathrm{Id}_V)$.

i.e. it is the set of all eigenvectors of T corresponding to *lambda*, together with $\vec{0}$.

Remark 28.3. λ is an eigenvalue of $T \iff E(\lambda, T) \neq \{\vec{0}\}$.

Example 28.4

$T \in \mathscr{L}(\mathbb{R}^3), \mathcal{M}(T) = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 19 & 0 \\ 0 & 0 & 19 \end{pmatrix}$ then

$$E(5, T) = \text{span}(v_1) \qquad E(19, T) = \text{span}(v_2, v_3).$$

Let V be a finite dimensional vector space over $\mathbb{F}$ with $T \in \mathscr{L}(V)$ and let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of T.

Proposition 28.5

$E(\lambda_1, T) + \ldots + E(\lambda_n, T)$ is a direct sum and $\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_n, T) \leq \dim V$.

Proof. Let $u_i \in E(\lambda_i, T)$ for $i = 1, \ldots, n$and suppose $u_1 + \ldots + u_n = \vec{0}$. By Proposition 25.5, $u_i = \vec{0}$ so the sum is direct, so $\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_n, T) = \dim(E(\lambda_1, T) + \ldots + E(\lambda_n, T)) \leq \dim V$. □

§29 October 31, 2018

Definition 29.1. An operator $T \in \mathscr{L}(V)$ is **diagonalizable** if there exists a basis of V for which $\mathcal{M}(T)$ is diagonal.

Example 29.2

$T \in \mathscr{L}(\mathbb{R}^2)$ defined by $T(x, y) = (41x + 7y, -20x + 74y)$. Then $\mathcal{M}(T,\ \text{standard basis}) = \begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix}$. But T is actually diagonalizable. Take basis $(1, 4), (7, 5)$ so that

$$\begin{aligned} T(v_1) &= (69, 276) &&= 69 \cdot v_1 + 0 \cdot v_2 \\ T(v_2) &= (322, 230) &&= 0 \cdot v_1 + 46 \cdot v_2 \end{aligned}$$

so $\mathcal{M}(T, (v_1, v_2)) = \begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$.

Theorem 29.3 (Conditions for Diagonalizability)

Suppose that V is a finite dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote all of the distinct eigenvalues of T. Then the following are equivalent:

1. T is diagonalizable.
2. V has a basis of eigenvectors of T.
3. There exist 1-dimensional subspaces $U_1, \ldots, U_n$ of V each invariant under T such that $V = U_1 \oplus \cdots \oplus U_n$.
4. $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.
5. $\dim V = \dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T)$.

Proof. (1) $\iff$ (2): $\mathcal{M}(T)$ is diagonal with respect to the basis $v_1, \ldots, v_n \iff T(v_i) = \lambda_i \cdot v_i \quad \forall\, i = 1, \ldots, n$.
(2) $\implies$ (3): Let $v_1, \ldots, v_n$ be a basis of V consisting of eigenvectors. $T(v_i) = \lambda_i v_i \implies U_i := \operatorname{span}(v_i)$ is invariant under T and is 1-dimensional. Moreover, $V = U_1 \oplus \cdots \oplus U_n$ since $v_1, \ldots, v_n$ is a basis.
(3) $\implies$ (2): $U_i = \operatorname{span}(v_i)$ for some $v_i \neq 0 \in V$. Since U_i is invariant under T, then $T(v_i) = \lambda_i v_i, \lambda_i \in \mathbb{F}$. Since $V = U_1 \oplus \cdots \oplus U_n$, then $v_1, \ldots, v_n$ is a basis.
(2) $\implies$ (4): Any $v \in V$ is a linear combination of eigenvectors $\implies V \subseteq E(\lambda_i, T) + \ldots + E(\lambda_m, T)$. Since $\lambda_1, \ldots, \lambda_m$ are distinct, Proposition 28.5 implies that the sum is direct.
(4) $\implies$ (5) Done last week.
(5) $\implies$ (2): Let $\dim V = n$. Choose a basis for each $E(\lambda_i, T)$ and concatenate to get list $v_1, \ldots, v_n$ which has length n.
Claim: $v_1, \ldots, v_n$ is a basis.
Proof: Suppose $a_1 v_1 + \ldots + a_n v_n = \vec{0}$. Rewrite this as $u_1 + \ldots + u_m = \vec{0}$ with $u_i \in E(\lambda_i, T)$, $u_i = \sum_k a_k v_k$ with $v_k \in E(\lambda_i, T)$. Then $u_1, \ldots, u_m$ are linearly independent since each $u_i = E(\lambda_i, T)$ so it's an eigenvector, and eigenvectors corresponding to distinct eigenvalues are linearly independent, so $u_i = \vec{0}$ so $u_k = \sum_k a_k v_k - \vec{0} \implies a_k = 0$ since v_k's form a basis of $E(\lambda_i, T)$. Apply to each U_i. Finally, $v_1, \ldots, v_n$ is a list of linearly independent vectors of length $n = \dim V$, and must therefore be a basis. □

Example 29.4

$T : \mathbb{C}^2 \to \mathbb{C}^2$ mapping $(w, z) \mapsto (z, 0)$. $\mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Eigenvalues: $0 \in \mathbb{C}$ and $E(0, T) = \{v \in \mathbb{C}^2 \mid T(v) = \vec{0}\} = \{(w, 0) \in \mathbb{C}^2\}$ so $\dim E(0, T) = 1$, so T is not diagonalizable.

Corollary 29.5

Suppose $\dim V = n$. If $T \in \mathscr{L}(v)$ has n distinct eigenvalues, then T is diagonalizable.

Proof. $\lambda_1, \ldots, \lambda_n$ distinct eigenvalues with corresponding eigenvectors $v_1, \ldots, v_n$. They are linearly independent, so they must be a basis. By (2) $\implies$ (1) of the theorem, we are done. □

§30 November 5, 2018

§30.1 Inner Product Spaces

Let $V = \mathbb{R}^n$ and let $\vec{x} = (x_1, \ldots, x_n)$.

Definition 30.1. The **norm** of $\vec{x}$ is

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

Definition 30.2. For $\vec{x}, \vec{y} \in \mathbb{R}^n$, the **dot product** is

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Remark 30.3. $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$.

Definition 30.4. The **angle** θ between $\vec{x}$ and $\vec{y} \in \mathbb{R}^n$ is

$$\|\vec{x}\| \cdot \|\vec{y}\| \cos\theta = \vec{x} \cdot \vec{y}.$$

Properties

- $\vec{x} \cdot \vec{x} \geq 0$ and $\vec{x} \cdot \vec{x} = 0 \iff \|\vec{x}\| = 0 \iff \vec{x} = \vec{0}$.
- $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$.
- Fix $\vec{y} \in \mathbb{R}^n$. The map $D_{\vec{y}} : \mathbb{R}^n \to \mathbb{R}$ taking $\vec{x} \mapsto \vec{x} \cdot \vec{y}$ is linear.

Now let V be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Definition 30.5. An **inner product** on V takes in an ordered pair of vectors $(\vec{u}, \vec{v}) \in V^2$ and outputs $\langle u, v\rangle \in \mathbb{F}$ satisfying:

1. $\langle \vec{v}, \vec{v}\rangle \geq 0$ for all $v \in V$ (in particular $\langle v, v\rangle$).
2. $\langle \vec{v}, \vec{v}\rangle = 0 \iff \vec{v} = \vec{0}$.
3. $\langle \vec{u} + \vec{v}, \vec{w}\rangle = \langle \vec{u}, \vec{w}\rangle + \langle \vec{v}, \vec{w}\rangle$.
4. $\langle \lambda \cdot \vec{v}, \vec{w}\rangle = \lambda \cdot \langle \vec{v}, \vec{w}\rangle$
5. $\langle \vec{u}, \vec{v}\rangle = \overline{\langle \vec{v}, \vec{u}\rangle}$

Example 30.6
Examples of inner products:

1. $V = \mathbb{R}^n, \mathbb{F} = \mathbb{R}$ and $\langle \, , \rangle =$ dot product.
$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = x_1 y_1 + \ldots + x_n y_n \in \mathbb{R}.$$
2. $V = \mathbb{C}^n, \mathbb{F} = \mathbb{C}$ and $\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = z_1 \overline{w_1} + z_2 \overline{w_2} + \ldots + z_n \overline{w_n}$.
$$\langle \vec{z}, \vec{z} \rangle = z_1 \overline{z_1} + z_2 \overline{z_2} + \ldots + z_n \overline{z_n} = |z_1|^2 + |z_2|^2 + \ldots + |z_n|^2 \geq 0.$$
3. $V = \mathcal{P}(\mathbb{R}), \mathbb{F} = \mathbb{R}$ and $\langle p, q \rangle = \displaystyle\int_0^\infty p(x) q(x) e^{-x} \, dx$.
4. $V = \{ f : [-1, 1] \to \mathbb{R} \mid f \text{ is continuous} \}, \mathbb{F} = \mathbb{R}$ and $\langle f, g \rangle = \int_{-1}^1 f(x) g(x) \, dx$.

Definition 30.7. An **inner product space** is a vector space with an inner product.

These axioms have some straightforward consequences:

1. Fix $\vec{u} \in V$. Define $T_{\vec{u}} : V \to \mathbb{F}$ with $\vec{v} \mapsto \langle \vec{v}, \vec{u} \rangle$. Then $T_{\vec{u}}$ is linear.
2. $\langle \vec{0}, \vec{u} \rangle = 0$ for all $\vec{u} \in V$.
3. $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$.

   Proof. $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \overline{\langle \vec{v} + \vec{w}, \vec{u} \rangle} = \overline{\langle \vec{v}, \vec{u} \rangle} + \overline{\langle \vec{w}, \vec{u} \rangle} = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$. ☐
4. $\langle \vec{u}, \vec{0} \rangle = 0$ for all $\vec{u} \in V$.
5. $\langle \vec{u}, \lambda \vec{v} \rangle = \overline{\lambda} \langle \vec{u}, \vec{v} \rangle$.

   Proof. $\langle \vec{u}, \lambda \vec{v} \rangle = \overline{\langle \lambda \vec{v}, \vec{u} \rangle} = \overline{\lambda \langle \vec{v}, \vec{u} \rangle}$. ☐

§31 November 7, 2018

Definition 31.1. The **norm** $\|\vec{v}\| \overset{\text{def}}{=} \sqrt{\langle v, v \rangle}$ satsifies

- $\|\vec{v}\| = 0 \iff \langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = \vec{0}$.
- $\|\lambda \vec{v}\| = \sqrt{\langle \lambda \vec{v}, \lambda \vec{v} \rangle} = |\lambda| \cdot \|\vec{v}\|$.

Definition 31.2. Two vectors $\vec{u}, \vec{v} \in V$ are **orthogonal** if $\langle \vec{u} \vec{v} \rangle = 0$.

Example 31.3 (Orthogonality in $\mathbb{R}^2$)
$V = \mathbb{R}^2, \mathbb{F} = \mathbb{R} \langle \, , \rangle =$ dot product.
Then $\langle (x_1, y_1), (x_2, y_2) \rangle = 0 \iff x_1 x_2 + y_1 y_2 = 0$.
$\iff y_1 y_2 = -x_1 x_2$
$\iff \frac{y_1}{x_1} = -\frac{x_2}{y_2} = -\frac{1}{\frac{y_2}{x_2}}$
$\iff$ vectors are perpendicular.

Theorem 31.4 (Pythagorean Theorem)
Suppose $\vec{u}, \vec{v} \in V$ are orthogonal. Then

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$$

Proof.

$$\begin{aligned} \|\vec{u} + \vec{v}\|^2 &= \langle \vec{u} + \vec{v}, \vec{u} + \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} + \vec{v} \rangle + \langle \vec{v}, \vec{u} + \vec{v} \rangle \\ &= \langle \vec{u}, \vec{u} \rangle + \cancel{\langle \vec{u}, \vec{v} \rangle} + \cancel{\langle \vec{v}, \vec{u} \rangle} + \langle \vec{v}, \vec{v} \rangle \\ &= \|\vec{u}\|^2 + \|\vec{v}\|^2 . \end{aligned}$$

□

§31.1 Orthogonal Projection

Idea: Given $\vec{u}, \vec{v} \in V$, find $c \in \mathbb{F}$ and $\vec{w} \in V$ such that $\vec{u} = c\vec{v} + \vec{w}$, where $\vec{w}$ is orthogonal to $\vec{v}$.

Then $\vec{u} = c\vec{v} + \underbrace{(\vec{u} - c\vec{v})}_{\vec{w}}$. We want $\langle \vec{w}, \vec{v} \rangle = 0$.

$$\begin{aligned} \langle \vec{w}, \vec{v} \rangle = 0 &\iff \langle \vec{u} - c\vec{v}, \vec{v} \rangle = 0 \\ &\iff \langle \vec{u}, \vec{v} \rangle + \langle -c\vec{v}, \vec{v} \rangle = 0 \\ &\iff \langle \vec{u}, \vec{v} \rangle - c\langle \vec{v}, \vec{v} \rangle = 0 \\ &\iff \langle \vec{u}, \vec{v} \rangle = c\langle \vec{v}, \vec{v} \rangle = c\|\vec{v}\|^2 \end{aligned}$$

This shows that

$$c = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2} \qquad \vec{w} = \vec{u} - \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v}.$$

§31.2 Cauchy-Schwartz Inequality

Suppose $\vec{u}, \vec{v} \in V$. Then

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\| .$$

Moreover, equality holds $\iff$ one of $\vec{u}, \vec{v}$ is a scalar multiple of the other.

Proof. If $\vec{v} = \vec{0}$, then both sides of the desired inequality are 0, and we are done. Assume $\vec{v} \neq \vec{0}$.

Consider the orthogonal decomposition $\vec{u} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v} + \vec{w}$ where $\vec{w}$ is orthogonal to $\vec{v}$. By the Pythagorean Theorem,

$$\begin{aligned} \|\vec{u}\|^2 &= \left\| \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v} \right\|^2 + \|\vec{w}\|^2 \\ &= \frac{|\langle \vec{u}, \vec{v} \rangle|^2 \cdot \|\vec{v}\|^2}{\|\vec{v}\|^4} + \|\vec{w}\|^2 \\ &= \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\|\vec{v}\|^2} + \|\vec{w}\|^2 \\ &\geq \frac{|\langle \vec{u}, \vec{v} \rangle|^2}{\|\vec{v}\|^2}. \end{aligned}$$

and the result follows. Equality holds $\iff \|\vec{w}\|^2 \cdot \|\vec{v}\|^2 = 0 \iff \|\vec{w}\|^2 = \vec{0} \iff \vec{v}$ is a multiple of $\vec{u}$. □

Example 31.5

Let $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ with the dot product. Then Cauchy-Schwartz tells us that

$$|x_1y_1 + \ldots + x_ny_n|^2 \leq (x_1^2 + \ldots + x_n^2)(y_1^2 + \ldots + y_n^2).$$

§32 November 9, 2018

Theorem 32.1 (Triangle Inequality)

Suppose $u, v \in V$. Then $\|u + v\| \leq \|u\| + \|v\|$.

Proof.

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\
&= \langle u,u\rangle + \langle v,v\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} \\
&= \|u\|^2 + \|v\|^2 + 2\Re\langle u,v\rangle \\
&\leq \|u\|^2 + \|v\|^2 + 2\left|\langle u,v\rangle\right| \\
&\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \quad \text{by Cauchy-Schwartz} \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

□

Definition 32.2. 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. i.e. $e_1, \ldots, e_m$ is orthonormal if $\langle e_i, e_j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

Example 32.3

Orthonormal lists

- the standard basis in $\mathbb{R}^n$ or $\mathbb{C}^n$ with respect to the dot product
- $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ in $\mathbb{R}^3$ or $\mathbb{C}^3$.
- $\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ added to the list above

Lemma 32.4

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

$$\|a_1e_1 + \ldots + a_me_m\|^2 = |a_1|^2 + \ldots + |a_m|^2$$

Proof. Since $\|e_j\| = 1$ for each j, we can apply the Pythagorean Theorem repeatedly to obtain the result. □

Proposition 32.5

Every orthonormal list of vectors in V is linearly independent.

Proof. Suppose that $e_1, \ldots, e_m$ is an orthonormal list and $a_1, \ldots, a_m \in \mathbb{F}$ are such that $a_1e_1 + \ldots + a_me_m = \vec{0}$.

Then take the norm of both sides and apply the lemma to obtain

$$|a_1|^2 + \ldots + |a_m|^2 = 0 \implies a_i = 0$$

so $e_1, \ldots, e_m$ is linearly independent. □

Definition 32.6. An **orthonormal basis** of V is an orthonormal list of vectors in V that is also a basis.

Example 32.7

$V = \mathbb{R}^4$.

$$v_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \qquad v_2 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$$
$$v_3 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right) \qquad v_4 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$$

Proof. To see that this is a basis, we'll check that it is an orthonormal list.

$$\|v_i\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = 1.$$

Check that $v_i \cdot v_j = 0$ for $i \neq j$ so the orthonormal and linearly independent list of length 4 must be a basis. □

Proposition 32.8

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

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

Proof. Since $e_1, \ldots, e_n$ is a basis of V, there are scalars $a_1, \ldots, a_n \in \mathbb{F}$ such that $v = a_1e_1 + \ldots + a_ne_n$.

Since $e_1, \ldots, e_n$ is orthonormal, consider the inner product

$$\begin{aligned}\langle v, e_j\rangle &= \langle a_1e_1 + \ldots + a_ne_n, e_j\rangle \\ &= \langle a_1e_1, e_j\rangle + \ldots + \langle a_ne_n, e_j\rangle \\ &= a_j\langle e_j, e_j\rangle \\ &= a_j\end{aligned}$$

□

§33 November 12, 2018

§33.1 Gram-Schmidt Procedure

Let V be a finite dimensional vector space.
Input: A linearly independent list $v_1, \ldots, v_m$ in V.
Output: An orthonormal list $e_1, \ldots, e_m$ satisfying $\text{span}(v_1, \ldots, v_j) = \text{span}(e_1, \ldots, e_j)$ for $j = 1, \ldots, m$.

Step 1: $e_1 = \dfrac{v_1}{\|v_1\|}$. Note $\langle e_1, e_1 \rangle = 1$.
Step j: $e_j = \dfrac{v_j - \langle v_j, e_1 \rangle e_1 - \ldots - \langle v_j, e_{j-1} \rangle e_{j-1}}{\|v_j - \langle v_j, e_1 \rangle e_1 - \ldots - \langle v_j, e_{j-1} \rangle e_{j-1}\|}$. Verify that $\langle e_j, e_k \rangle = 0$ if $j \neq k$.
Note that $e_j \in \text{span}(v_1, \ldots, v_j)$
e.g. $e_1 \in \text{span}(v_1)$ and $e_2 = \dfrac{v_2 - \langle v_2, e_1 \rangle e_1}{\|v_2 - \langle v_2, e_1 \rangle e_1\|} \in \text{span}(v_1, v_2)$.
This shows that $\text{span}(e_1, \ldots, e_j) \subseteq \text{span}(v_1, \ldots, v_j)$. Since both lists are linearly independent, they both have dimension j, so $\text{span}(e_1, \ldots, e_j) = \text{span}(v_1, \ldots, v_j)$.

Example 33.1
$V = \mathbb{R}^3$ with the dot product as inner product.

$$v_1 = (1, 1, 0) \quad v_2 = (2, 2, 3) \quad v_3 = (0, 1, -1)$$

Solution. Applying Gram Schmidt to v_1, v_2, v_3,
Step 1: $c_1 = \dfrac{(1,1,0)}{\|(1,1,0)\|} = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right)$.
Step 2: $e_2 = \dfrac{v_2 - \langle v_2, e_1 \rangle e_1}{\|v_2 - \langle v_2, e_1 \rangle e_1\|} = \dfrac{(2,2,3) - \left((2,2,3) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)\right)\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)}{\|\text{numerator}\|} = \dfrac{(2,2,3) - (2,2,0)}{\|\text{numerator}\|} = (0, 0, 1)$.
Step 3: The numerator of e_3 is $v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2 = (0, 1, -1) - (0, 1, -1) \cdot \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}, 0\right) e_1 - (0, 1, -1) \cdot (0, 0, 1) e_2 = (0, 1, -1) - \dfrac{1}{\sqrt{2}} e_1 + e_2 = \left(-\dfrac{1}{2}, \dfrac{1}{2}, 0\right)$. So $e_3 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$. □

Consequences of the Gram-Schmidt

1. Every finite dimensional inner product space has an orthonormal basis.
2. Any orthonormal list can be extended to an orthonormal basis
3. Suppose $T \in \mathscr{L}(V)$ and there exists a basis $v_1, \ldots, v_m$ of V such that $\mathcal{M}(T, (v_1, \ldots, v_m))$ is upper triangular. Then there exists an orthonormal basis $e_1, \ldots, e_m$ such that $\mathcal{M}(T, (e_1, \ldots, e_m))$ is upper triangular.

Proof of (3). Apply the Gram-Schmidt procedure to $v_1, \ldots, v_m$ to get an orthonormal list $e_1, \ldots, e_m$. Since $\text{span}(v_1, \ldots, v_j) = \text{span}(e_1, \ldots, e_j)$ for each $j = 1, \ldots, m$, $\mathcal{M}(T)$ is upper triangular means $T(\text{span}(v_1, \ldots, v_j)) \subseteq \text{span}(v_1, \ldots, v_j)$. This means that $\text{span}(v_1, \ldots, v_j)$ is T-invariant, which means that $\text{span}(e_1, \ldots, e_j)$ is also T-invariant, so that $\mathcal{M}(T, (e_1, \ldots, e_m))$ is upper triangular. □

§34 November 14, 2018

Linear Functionals on an Inner Product Space

Let V be a finite dimensional inner product space and $\phi : V \to \mathbb{F}$ be a linear functional. Let $e_1, \ldots, e_n$ be an orthonormal basis of V. If $v \in V$, then by Proposition 32.8, we can write v as $v = \langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n$. Then

$$\begin{aligned}\phi(v) &= \langle v, e_1\rangle\phi(e_1) + \ldots + \langle v, e_n\rangle\phi(e_n)\\ &= \langle v, e_1\overline{\phi(e_1)}\rangle + \ldots + \langle v, e_n\overline{\phi(e_n)}\rangle\\ &= \langle v, \underbrace{e_1\overline{\phi(e_1)} + \ldots + e_n\overline{\phi(e_n)}}_{\text{call this vector } u}\rangle\\ &= \langle v, u\rangle\end{aligned}$$

Note that u does not depend on v, so $\phi(v) = \langle v, u\rangle$ for all $v \in V$. This motivates the following theorem:

Theorem 34.1 (Riesz Representation Theorem)
For every $\phi \in V$, there exists a unique $u \in V$ such that

$$\phi(v) = \langle v, u\rangle \quad \forall v \in V.$$

Proof. Since existence has already been shown above, it suffices to show that this representation is unique. Suppose there exist u_1, u_2 such that $\langle v, u_1\rangle = \langle v, u_2\rangle \quad \forall v \in V$.
$\implies \langle v, u_1 - u_2\rangle = 0$ for all $v \in V$.
Choose $v = u_1 - u_2$, then $0 = \langle u_1 - u_2, u_1 - u_2\rangle = \|u_1 - u_2\| \implies u_1 - u_2 = 0 \implies u_1 = u_2$. □

Example 34.2
$V = \mathcal{P}_2(\mathbb{R})$ with $\langle p, q\rangle = \displaystyle\int_{-1}^{1} p(x)q(x)\,dx$.
Then Gram-Schmidt produces the orthonormal basis

$$e_1 = \frac{1}{\sqrt{2}}, \quad e_2 = \sqrt{\frac{3}{2}}x, \quad e_3 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$$

Let $\phi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$ sending $p \mapsto \displaystyle\int_{-1}^{1} p(x)\cos\pi x\,dx$.

We want to find $q(x) \in \mathcal{P}_2(x)$ such that $\phi(p) = \langle p, q\rangle$, i.e. $\displaystyle\int_{-1}^{1} p(x)\cos\pi x\,dx = \int_{-1}^{1} p(x)q(x)\,dx$. Note that we can't take $q(x)$ to be $\cos\pi x$ since $\cos\pi x$ is not in $\mathcal{P}_2(\mathbb{R})$.

Riesz Representation Theorem tells us that $q(x)$ is equal to

$$\frac{1}{\sqrt{2}}\overline{\left(\int_{-1}^{1}\frac{1}{\sqrt{2}}\cos\pi x\,dx\right)} + \sqrt{\frac{3}{2}}x\overline{\left(\int_{-1}^{1}\sqrt{\frac{3}{2}}x\cos\pi x\,dx\right)} + \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)\overline{\left(\int_{-1}^{1} e_3\cos\pi x\,dx\right)}$$

$$= -\frac{45}{2\pi^2}\left(x^2 - \frac{1}{3}\right).$$

§35 November 16, 2018

§35.1 Adjoint of an Operator

Let $(V, \langle\,,\rangle_V), W, \langle\,,\rangle_W$ be two inner product spaces. Let $T : V \to W$ be a linear map.

Definition 35.1. The **adjoint** of T is the function $T^* : W \to V$ such that

$$\langle T(v), w\rangle_W \overset{\text{def}}{=} \langle v, T^*(w)\rangle_V$$

for all $v \in V, w \in W$.

To show that this definition makes sense, fix $w \in W$ and consider $\phi : V \to \mathbb{F}$ defined by $\phi(v) := \langle T(v), w\rangle_W$. By Riesz Representation Theorem, there is a unique $u \in V$ such that $\phi(v) = \langle v, u\rangle_V$. Thus, we just call/define u as $T(w)$.

Example 35.2
Find the adjoint of $T : \mathbb{R}^3 \to \mathbb{R}^2$ with the dot product sending $(x_1, x_2, x_3) \mapsto (x_1 + x_2, 2x_2 + x_3)$.

Solution. $T^* : \mathbb{R}^2 \to \mathbb{R}^2$.

$$\begin{aligned}
\langle (x_1, x_2, x_3), T^*(y_1, y_2)\rangle_{\mathbb{R}^3} &= \langle T(x_1, x_2, x_3), (y_1, y_2)\rangle_{\mathbb{R}^2} \\
&= \langle (x_1 + x_2, 2x_2 + x_3), (y_1, y_2)\rangle_{\mathbb{R}^2} \\
&= y_1(x_1 + x_2) + y_2(2x_2 + x_3) \\
&= x_1 y_1 + x_2(y_1 + 2y_2) + x_3 y_2 \\
&= \langle (x_1, x_2, x_3), (y_1, y_1 + 2y_2, y_2)\rangle_{\mathbb{R}^3}
\end{aligned}$$

so $T^*(y_1, y_2) = (y_1, y_1 + 2y_2, y_2)$. □

Proposition 35.3 (The Adjoint is a Linear Map)
If $T \in \mathscr{L}(V, W)$, then $T^* : W \to V$ is linear.

Proof. Fix $w_1, w_2 \in W$ and let $v \in V$.

$$\begin{aligned}
\langle v, T^*(w_1 + w_2)\rangle_V &= \langle T(v), w_1 + w_2\rangle_W \\
&= \langle T(v), w_1\rangle_W + \langle T(v), w_2\rangle_W \\
&= \langle v_1, T^*(w_1)\rangle_V + \langle v, T^*(w_2)\rangle_V \\
&= \langle v, T^*(w_1) + T^*(w_2)\rangle_V
\end{aligned}$$

so $T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2)$.
Now let $\lambda \in \mathbb{F}$ and $w \in W$. If $v \in V$ then

$$\begin{aligned}
v, T^*(\lambda w)\rangle_V &= \langle T(v), \lambda_W\rangle_W \\
&= \overline{\lambda}\langle T(v), w\rangle_W \\
&= \overline{\lambda}\langle v, T^*(w)\rangle_V \\
&= \langle v, \lambda T^*(w)\rangle_V
\end{aligned}$$

so $T^*(\lambda w) = \lambda T^*(w)$. □

Properties of the Adjoint

Let $S, T \in \mathscr{L}(V, W)$.

1. $(S+T)^* = S^* + T^*$.

   Proof. $\langle v, (S+T)^*(w)\rangle_V = \langle (S+T)(v), w\rangle_W = \langle S(v), w\rangle_W + \langle T(v), w\rangle_W = \langle v, S^*(w)\rangle_V + \langle v, T^*(w)\rangle_V = \langle v, S^*(w) + T^*(w)\rangle_V$. □

2. $(\lambda T)^* = \lambda T^*$.

3. $(T^*)^* = T$.

   Proof. $\langle w, (T^*)^*\rangle_W = \langle T^*(w), v\rangle_V = \overline{\langle v, T^*(w)\rangle_V} = \overline{\langle T(v), w\rangle_W} = \langle w, T(v)\rangle_W$ □

4. $I^* = I$.

5. $(ST)^* = T^*S^*$.

§36 November 19, 2018

Proposition 36.1 (The Matrix of T^*)
Let $T \in \mathscr{L}(V, W)$, $e_1, \ldots, e_n$ be an orthonormal basis of V, and $f_1, \ldots, f_m$ be an orthonormal basis of W. Then

$$\mathcal{M}(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_m)) = \overline{\mathcal{M}(T)}^t.$$

Proof. Let $\mathcal{M}(T) = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$.

$T(e_k) = a_{1,k}f_1 + \ldots + a_{m,k}f_m = \langle T(e_k), f_1\rangle f_1 + \ldots + \langle T(e_k), f_m\rangle f_m$, so $a_{i,k} = \langle T(e_k), f_i\rangle$.

Now, also $M(T^*) = \begin{pmatrix} b_{1,1} & \ldots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \ldots & b_{n,m} \end{pmatrix}$.

$T^*(f_k) = b_{1,k}e_1 + \ldots + b_{n,k}e_n = \langle T^*(f_k), e_1\rangle e_1 + \ldots + \langle T^*(f_k), e_n\rangle e_n$, so $b_{i,k} = \langle T^*(f_k), e_i\rangle$. Hence,

$$\begin{aligned} b_{i,k} &= \langle T^*(f_k), e_i\rangle \\ &= \langle f_k, (T^*)^*(e_i)\rangle \\ &= \langle f_k, T(e_i)\rangle \\ &= \overline{\langle T(e_i), f_k\rangle} \\ &= \overline{a_{k,i}} \end{aligned}$$ □

Now we take the case when $V = W$, i.e. $T \in \mathscr{L}(V)$.

Definition 36.2. T is self-adjoint if $T^* = T$, i.e.

$$\langle T(v_i), v_2\rangle = \langle v_1, T(v_2)\rangle \qquad v_1, v_2 \in V.$$

Definition 36.3. T is normal if $T \circ T^* = T^* \circ T$.

Note that self-adjoint $\implies$ normal.

Remark 36.4. If T is self-adjoint, then $\mathcal{M}(T) = \overline{\mathcal{M}(T)}^t$ with respect to an orthonormal basis of V.

Example 36.5

$T : \mathbb{C}^3 \to \mathbb{C}^3$ with respect to the standard basis. Then

$$\mathcal{M}(T) = \begin{pmatrix} 3 & i & 2+i \\ -i & 7 & 0 \\ 2-i & 0 & 9 \end{pmatrix}.$$

Remark 36.6. Normal $\not\Rightarrow$ self-adjoint.

Example 36.7

$T : \mathbb{R}^2 \to \mathbb{R}^2$ with respect to the standard basis. Then $\mathcal{M}(T) = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$ is not self-adjoint. But $T^*T - TT^* = \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} - \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, so T is normal.

Remark 36.8. $S := T^*T - TT^*$ is self-adjoint.

Proposition 36.9

Self-adjoint operators have real eigenvalues.

Proof. Suppose T is self-adjoint and $Tv = \lambda v$, $v \neq 0$.
$\lambda \|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\|v\|^2 \implies \lambda = \overline{\lambda}$ since $\|v\|^2 \neq 0$. □

Proposition 36.10

If T is self-adjoint, then

$$\langle T(v), v\rangle = 0 \quad \forall\, v \in V \iff T = 0.$$

Corollary 36.11

T is normal $\iff \|T(v)\| = \|T^*(v)\| \quad \forall\, v \in V$.

§37 November 26, 2018

Proof. Define $S := T^*T - TT^*$.

$$\begin{aligned} T \text{ is normal} &\iff S = 0 \\ &\iff S \text{ is self-adjoint} \\ &\iff \langle S(v), v\rangle = 0 \quad \forall\, v \in V \\ &\iff \langle (T^*T - TT^*)(v), v\rangle = 0 \\ &\iff \langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle \\ &\iff \langle T(v), T(v)\rangle = \langle T^*(v), T^*(v)\rangle \\ &\iff \|T(v)\| = \|T^*(v)\| \quad \forall\, v \in V. \end{aligned}$$

□

Proposition 37.1

T is normal $\iff$ $T - \lambda \mathrm{Id}$ is also normal.

Proof. See Axler. □

Corollary 37.2

T normal $\implies$ eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof. Suppose $T(v) = \lambda v$ and $T(w) = \beta w$ where $\lambda \neq \beta$ and $v, w \neq \vec{0}$. Then

$$\begin{aligned} 0 &= \langle T(v), w\rangle - \langle v, T^*(w)\rangle \\ &= \langle \lambda v, w\rangle - \langle v, \overline{\beta} w\rangle \\ &= \lambda\langle v, w\rangle - \beta\langle v, w\rangle \\ &= (\lambda - \beta)\langle v, w\rangle. \end{aligned}$$

Since $\lambda \neq \beta$, then $\langle v, w\rangle = 0$. □

§37.1 Spectral Theorem over $\mathbb{C}$

Theorem 37.3 (Complex Spectral Theorem)

Let V be a finite dimensional vector space with $\mathbb{F} = \mathbb{C}$. Let $T \in \mathscr{L}(V)$. Then the following are equivalent:

1. T has a diagonal matrix with respect to some orthonormal basis of V.
2. V has an orthonormal basis of eigenvectors of T.
3. T is normal.

Proof. (3) $\implies$ (1): Let $n = \dim V$. By Theorem 27.5, we know that V has a basis $e_1, \ldots, e_n$ such that $\mathcal{M}(T)$ is upper-triangular, i.e. $\mathcal{M}(T) = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n,n} \end{pmatrix}$. Applying Gram-Schmidt to this list, we can assume that $e_1, \ldots, e_n$ is an orthonormal basis.

We claim that $\mathcal{M}(T)$ is already diagonal. To see this, consider $\mathcal{M}(T^*) = \begin{pmatrix} \overline{a_{1,1}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ \overline{a_{1,n}} & \dots & \overline{a_{n,n}} \end{pmatrix}$.

$T^*(e_1) = \overline{a_{1,1}}e_1 + \ldots + \overline{a_{1,n}}e_n$. So T normal $\implies \|T^*(e_1)\| = \|T(e_1)\| \implies |a_{1,1}|^2 + \ldots + |a_{1,n}|^2 = |a_{1,1}|^2 \implies a_{1,2} = \ldots = a_{1,n} = 0$. Repeating this argument for the other rows of $\mathcal{M}(T)$ yields the desired result. $\square$

(1) $\implies$ (3): Suppose there exists an orthonormal basis $e_1, \ldots, e_n$ such that $\mathcal{M}(T, (e_1, \ldots, e_n)) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$. Then $\mathcal{M}(T^*) = \begin{pmatrix} \overline{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \overline{\lambda_n} \end{pmatrix}$.

To show that $TT^* = T^*T$, it suffices to check that $\mathcal{M}(T)\mathcal{M}(T^*) = \mathcal{M}(T^*)\mathcal{M}(T)$, which holds because they are both diagonal matrices. $\square$

(1) $\iff$ (2): This holds by Theorem 29.3. $\square$

§37.2 Spectral Theorem over $\mathbb{R}$

Theorem 37.4 (Real Spectral Theorem)

Let V be a finite dimensional vector space with $\mathbb{F} = \mathbb{R}$. Let $T \in \mathscr{L}(V)$. Then the following are equivalent:

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

§38 November 28, 2018

§38.1 Positive Operators

Definition 38.1. $T \in \mathscr{L}(V)$ is **positive** or semi-definite if it is self adjoint and $\langle Tv, v\rangle \geq 0$ for all $v \in V$.

Remark 38.2. Positive operators have nonnegative eigenvalues.

Proof. $0 \leq \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle = \lambda \|v\|^2$. $\square$

Definition 38.3. A **square root** of $T \in \mathscr{L}(V)$ is an $R \in \mathscr{L}(V)$ such that $T = \mathbb{R}^2$.

Proposition 38.4

Positive operators have positive square roots.

Proof. Use the Spectral Theorem to obtain a diagonal matrix of nonnegative elements. Take the square root of each element to obtain matrix R. $\square$

Proposition 38.5

If $T \in \mathscr{L}(V)$, then $S := T^*T$ is positive.

Proof. S is self adjoint because $S^* = (T^*T)^* = T^*(T^*)^* = T^*T = S$.
Positivity condition: $\langle Sv, v\rangle = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \geq 0$. □

§38.2 Singular Value Decomposition

Let V have dimension n and $T \in \mathscr{L}(V)$. T^*T is positive $\implies$ it has a square root $\sqrt{T^*T}$.

Definition 38.6. The **singular values** $s_1, \ldots, s_n$ of T are the eigenvalues of $\sqrt{T^*T}$, each one repeated $\dim E(s_i, \sqrt{T^*T})$ times.

Order the singular values so that $s_1 \geq s_2 \geq \ldots \geq s_n \geq 0$.

Theorem 38.7

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

$$\mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)) = \begin{pmatrix} s_1 & 0 & \ldots & 0 \\ 0 & s_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & s_n \end{pmatrix}$$

In particular, $T(e_i) = s_i f_i$.
Since $v = \langle v, e_1\rangle + \ldots + \langle v, e_n\rangle$,

$$T(v) = s_1\langle v, e_1\rangle f_1 + \ldots + s_n\langle v, e_n\rangle f_n \quad \text{for all } v \in V$$

Example 38.8

Let A be the matrix of T with respect to the standard basis of T.
The SVD Theorem is equivalent to saying there are unitary matrices U, V such that

$$A = U \cdot \Sigma \cdot V^t$$

where Σ is the diagonal matrix whose diagonals contain the singular values of T.