MATH405: Linear Algebra

Michael Li

Contents

Goals of this course are to discuss

- Vector spaces
- Linear transformations between vector spaces
- Other operations on vector spaces

1 Vector Space

1.1 Definitions

Definition - Field: A set of numbers containing $0, 1$ that can be added, subtracted, multiplied, and divided (except cannot divide by 0) that satisfy the following **Field Axioms**

1. $a, b \in K \implies a + b, ab \in K$
2. $+, \times$ are commutative so $a + b = b + a$ and $ab = ba$
3. $+, \times$ are associative so $(a + b) + c = a + (b + c)$ and $a(bc) = (ab)c$
4. Distributive Law: $a(b + c) = ab + ac$
5. Additive Identity: $a + 0 = 0 + a = a$
6. Multiplicative Identity: $a \cdot 1 = 1 \cdot a = a$
7. Additive Inverse: $\forall a \in K, \exists b$ such that $a + b = 0$, namely $b = -a$ which is unique
8. Multiplicative Inverse: $\forall a \in K, \exists b$ such that $ab = 1$, name $b = 1/a$ which is unique

Example: R, Q are fields. Z is not a field since there is no multiplicative inverse of 2

Example: $C = \{a + bi \mid a, b \in R\}$, where $i = \sqrt{-1}$, is a field under

- $+ : (a + bi) + (c + di) = (a + c) + (b + d)i$
- $\times : (a + bi)(c + di) = (ac - bd) + (ad + bc)i$

Example: $F_2 = \{0, 1\}$ is a field under

- $+$: where

  $0 + 0 = 0$

  $0 + 1 = 1 + 0 = 1$

  $1 + 1 = 0$

- $\times$: where

  $0 \cdot 0 = 0$

  $0 \cdot 1 = 1 \cdot 0 = 0$

  $1 \cdot 1 = 1$

Example: For a prime p, let $F_p = \{0, \ldots, p - 1\}$. Then F_p is a field under

- $+ : a + b \pmod p$
- $\times : ab \pmod p$

Definition - Vector Space: For an arbitrary field K, a K-vector space is a set V, with a distinguished element O, such that any 2 elements in V can be added and scalar multiplied by $c \in K$

- $u, v \in V \implies u + v \in V$
- $c \in K, u \in V \implies cu \in V$

Satisfying the following properties

1. Commutative Addition: $u + v = v + u$
2. Associative Addition: $(u + v) + w = u + (v + w)$
3. Additive Identity: $u + O = u$
4. Additive Inverse: $\forall u \in V, \exists v \in V$ such that $u + v = O$, namely $v = -u$ which is unique
5. Distributive Laws: $\forall a, b \in K, a(u + v) = au + av$ and $(a + b)u = au + bu$
6. Commutative Scalar Multiplication: $(ab)u = a(bu)$
7. Multiplicative Identity: $1 \cdot u = u$

Example: R^3 is an R-vector space defined by the operations

$$R^3 = \{(x, y, z) \mid x, y, z \in R\}$$

- $+$: add componentwise so $(a, b, c) + (d, e, f) = (a + d, b + e, c + f)$
- $\times$: for $r \in R$, $r(a, b, c) = (ra, rb, rc)$
- Additive Identity is $O = (0, 0, 0)$

Example: For any field K, K^2 is a K-vector space defined by the operations

$$K^2 = \{(x, y) \mid x, y \in K\}$$

- $+$: add componentwise so $(a, b) + (c, d) = (a + c, b + d)$
- Scalar $\times$: for $k \in K$, $k(a, b) = (ka, kb)$
- Additive Identity is $O = (0, 0)$

Example: R is an R-vector space since clearly the necessary properties hold

Example R is a Q-vector space since clearly the necessary properties hold

- Notably, for $q \in Q$ and $r \in R$, we have $qr \in R$. Thus scalar multiplication is closed

Example: For any field K, the set $\{O\}$ is a K-vector space

Example: Let X be any non-empty set and let $\mathcal{F}(X)$ be the set of all functions $f : X \to R$. Then $\mathcal{F}$ is an R-vector space under the operations

- $+$: for $f, g \in \mathcal{F}(X)$, define $f + g := (f + g)(x)$
- $\times$: let $r \in R$, then define $rf := r(f(x))$
- Additive Identity is $O = f(x) = 0$, the function that takes any x to 0

Example: Take $X = N$ and let $F(X) = \{$ all functions $f : N \to R\}$ is a vector space

- **Note**: $f : N \to R$ is a sequence $(a_0, \dots, a_n)$ where $a_n = f(n)$

Lemma 1 - Cancellation: For $u, v, w \in V$ and if $u + v = w + v$, then $u = w$

Proof: $v \in V$ has an additive inverse, namely $-v$. Thus we have

$$u + v - v = w + v - v \implies u = w$$

Lemma 2 - Unique Additive Inverse: For all $v \in V$, there is a unique additive inverse, namely $-v$

Proof: Suppose u, w are both additive inverses of v. Then we have

$$v + u = v + w \implies u = w$$

Lemma 3 - 0 Times a Vector: For all $v \in V$, $0v = O$

Proof: $v = 1v = (0 + 1)v = 0v + 1v = 0v + v \implies 0v = O$

Lemma 4 - (-1)v is the Additive Inverse: For all $v \in v, (-1)v$ is the unique additive inverse of v

Proof: $(-1)v + v = (-1 + 1)v = 0v = O$. Thus $(-1)v$ is the additive inverse of v, which is unique by Lemma 2

Definition - Subspace: For a K-vector space V and a non-empty subset $W \subseteq V$, W is a **subspace** if it satisfies

- $w_1, w_2, \in W \implies w_1 + w_2 \in W$
- $\forall a \in K, w \in W \implies aw \in W$
- $O \in W$

Theorem 1: Every subspace of a K-vector space is a K-vector space

Proof: We need to show that $W \subseteq V$ satisfies all the necessary properties of a vector space

1. Verify $O \in W$

   Since W is non-empty and closed under scalar multiplication, take $0w = O \in W$ by Lemma 3
2. $u, v \in W \implies u + v \in W$ and $a \in K, v \in W \implies aw \in W$ by definition of subspace
3. Every $w \in W$ has an additive inverse, namely $-w$

   Since W is closed under scalar multiplication, $(-1)w = -w \in W$ by Lemma 4
4. Other conditions (associative addition, commutative addition, etc.) hold because $u, v, w \in W \implies u, v, w \in V$

   For example, choose $u, v \in W$, then $u + v = v + u$, since $u, v \in V$. Thus commutative addtion is satisfied

Example: Take $(5, 3, 2) \in R^3$. Then let $W = \{r(5, 3, 2) \mid r \in R\}$

Then W is an R-vector space. We prove this by showing that W is a subspace of R^3

- $+$: Choose 2 arbitrary elements of W, $r(5, 3, 2)$ and $s(5, 3, 2)$ for $r, s \in R$

  Then $r(5, 3, 2) + s(5, 3, 2) = (r + s)(5, 3, 2) \in W$
- $\times$: Choose $r(5, 3, 2) \in W$ and take $s \in R$

  Then $s(r(5, 3, 2)) = (sr)(5, 3, 2) \in W$

Example: Let $U = \{(x, y, z) \in R^3 \mid 2x + 3y = 0\}$. We show that U is a vector space by showing it's a subspace of R^3

- $+$: Take (x_1, y_1, z_1) and $(x_2, y_2, z_2) \in U \implies 2x_1 + 3y_1 = 0$ and $2x_2 + 3y_2 = 0$

  Then $2(x_1 + x_2) + 3(y_1 + y_2) = 0$

  Thus $(x_1 + x_2, y_1 + y_2, z_1 + z_2) \in U$

- $\times$: Let $(x, y, z) \in U$ and $r \in R$

  Then $2x + 3y = 0 \implies r(2x + 3y) = 2rx + 3ry = 0$

  Thus $r(x, y, z) \in U$

Example: Consider $\sin(x), \cos(x) \in \mathcal{F}(R)$ and let $W = \{a\sin(x) + b\cos(x) \mid a, b \in R\}$. Then W is a subspace of $\mathcal{F}(R)$

- $+$: Take $a_1 \sin(x) + b_1 \cos(x)$ and $a_2 \sin(x) + b_2 \cos(x) \in W$. Then $(a_1 + a_2)\sin(x) + (b_1 + b_2)\cos(x) \in W$
- $\times$: Take $r \in R$. Then $r(a\sin(x) + b\cos(x)) = (ra)\sin(x) + (rb)\cos(x) \in W$

1.2 Basis

Definition - Linear Combination: For vectors $\{v_1, \ldots, v_n\} \subseteq V$, a **linear combination** of $\{v_1, \ldots, v_n\}$ is a vector of the form

$$a_1 v_1 + \cdots + a_n v_n \qquad a_i \in K$$

Definition - Span: $\text{span}(\{v_1, \ldots, v_n\}) = \{$ all linear combinations of $\{v_1, \ldots, v_n\}\}$

Proposition 1: $W = \text{span}(\{v_1, \ldots, v_n\})$ is a subspace of V and thus is itself a K-Vector Space

Proof: We show that W satisfies the necessary criteria to be a subspace of V

- $+$: Let $a = a_1 v_1 + \cdots + a_n v_n \in W$ and $b = b_1 v_1 + \cdots + b_n v_n \in W$

  Then $a + b = (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n \in W$

  Thus W is closed under addition

- Scalar $\times$: Let $a = a_1 v_1 + \cdots + a_n v_n \in W$ and let $c \in K$

  Then $ca = (ca_1)v_1 + \cdots + (ca_n)v_n \in W$

  Thus W is closed under scalar multiplication

Example: Take $(5, 3, 1)$ and $(4, 0, -2) \in R^3$

$\text{span}(\{(5, 3, 1), (4, 0, -2)\})$ is a plane in R^3 passing through $(0, 0, 0)$

Example: Take $(5, 3, 1)$ and $(10, 6, 2) \in R^3$

$\text{span}(\{(5, 3, 1), (10, 6, 2)\})$ is a line in R^3 passing through $(0, 0, 0)$

- **Note**: $(10, 6, 2) = 2(5, 3, 1)$. Thus $\text{span}(\{(5, 3, 1), (10, 6, 2)\}) = a_1(5, 3, 1) + a_2(10, 6, 2) = (a_1 + 2a_2)(5, 3, 1)$

Definition - Linearly Independent: $\{v_1, \cdots, v_n\}$ is **linearly independent** if whenever $a_1 v_1 + \cdots + a_n v_n = 0$, then $a_1 = \cdots = a_n = 0$

- Otherwise $\{v_1, \ldots, v_n\}$ is **linearly dependent**

Proposition 2: $\{v_1, \ldots, v_n\}$ is linearly independent if and only if no v_i is a linearly combination of the other $n-1$ vectors

Proof: $\implies$ Assume $\{v_1, \ldots, v_n\}$ is linearly independent

BWOC, assume some $v_i = a_1v_1 + \cdots + a_nv_n$ for some $v_i \notin \{v_1, \ldots, v_n\}$

Then we have

$$O = a_1v_1 + \cdots + a_nv_n + (-1)v_i$$

Since v_i is a linear combination of $\{v_1, \ldots, v_n\}$, the above equation shows that $\{v_1, \ldots, v_n\}$ is linearly dependent. Contradiction

Thus v_i cannot be written as a linear combination of the other vectors

$\impliedby$ Assume by way of contraposition that $\{v_1, \ldots, v_n\}$ is not linearly independent

Thus choose $a_1, \ldots, a_n \in K$, not all 0 such that

$$a_1v_1 + \cdots + a_nv_n = O$$

WLOG, assume $a_1 \neq 0$. Then $v_2a_2 + \cdots + a_nv_n = a_1v_n$

Since $a_1 \neq 0$ and K is a field, we have

$$v_1 = \frac{a_2}{-a_1}v_2 + \cdots + \frac{a_n}{-a_1}v_n$$

Thus we have shown that v_1 is a linear combination of the other $n-1$ vectors

Corollary 3: $\{v_1, \ldots, v_n\}$ is linearly independent if and only if for each i, $v_i \notin \text{span}(\{v_1, \ldots, v_n\} \setminus \{v_i\})$

Proof: This follows from the previous proposition

Definition - Spans: Let W be a K-Vector Space and $\{v_1, \ldots, v_n\} \subseteq W$. If $\text{span}(\{v_1, \ldots, v_n\}) = W$, then $\{v_1, \ldots, v_n\}$ **spans** W, so every $w \in W$ is a linear combination of $\{v_1, \ldots, v_n\}$

Definition - Basis: $\{v_1, \ldots, v_n\}$ is a **basis** of W if it spans W and is linearly independent

Example: $\{(5, 3, 1), (4, 0, -2)\}$ is a basis for $\text{span}(\{(5, 3, 1), (4, 0, -2)\})$

Example: $\{(5, 3, 1), (10, 6, 2)\}$ is not a basis for $\text{span}(\{(5, 3, 1), (10, 6, 2)\})$ since it is not linearly independent

Proposition 4: Let $\{v_1, \ldots, v_n\}$ be a basis for W and let $w \in W$ be arbitrary. Then w can be written uniquely as

$$w = a_1 v_1 + \cdots + a_n v_n \qquad a_i \in K$$

Proof: Since $\{v_1, \ldots, v_n\}$ spans W, every $w \in W$ is a linear combination of $\{v_1, \ldots, v_n\}$

For uniqueness, suppose

$$w = a_1 v_1 + \cdots + a_n v_n = b_1 v_1 + \cdots + b_n v_n$$

Then we have

$$O = (b_1 - a_1) v_1 + \cdots (b_n - a_n)$$

Since $\{v_1, \ldots, v_n\}$ is linearly independent, we must have $b_i - a_i = 0$, and thus $b_i = a_i$ for each i

Thus each $w \in W$ can be written uniquely as a linear combination of $\{v_1, \ldots, v_n\}$

Example: Let $W = \text{span}(\{\sin(x), \cos(x)\} = \{a \sin(x) + b \cos(x) \mid a, b \in R\}$

We know that W is an R-Vector Space

$\{\sin(x), \cos(x)\}$ is linearly independent. Otherwise $\sin(x) = r \cos(x)$ for all $x \in X$ and some $r \in R$. However, this cannot hold for when $x = \pi/2$ since $\sin(\pi/2) = 1 \neq r \cos(\pi/2) = r0$

1.3 Dimension

Let $\{v_1, \ldots, v_n\} \subseteq V$ and let $W = \text{span}(\{v_1, \ldots, v_n\})$

Now let $X = \{w_1, \ldots, w_m\} \subseteq W$. Then there are 2 desirable properties of X

- **X is Big**: X spans W if $\text{span}(X) = W$, i.e. all $w \in W$ is a linear combination of elements from X
- **X is Small**: X is linearly independent, i.e. no element in X is a linear combination of the remaining elements

Note: the empty set $\emptyset$ is linearly independent since no element in $\emptyset$ is a linear combination of the others. Notably, $\emptyset$ is the basis for $\{O\}$

Shrinking Lemma: Let $X = \{w_1, \ldots, w_m\} \subseteq W$ and spans W but X is not linearly independent. Then $X \setminus \{w_i\}$ still spans W for some $w_i \in X$

Proof: Since X is not linearly independent, we know that some w_i is a linear combination of elements in $X \setminus \{w_i\}$. Suppose

$$w_i = a_1 w_1 + \cdots + a_m w_m \qquad \text{without } w_i \text{ occurring}$$

Then take arbitrary $u \in W$ where

$$u = b_1 w_1 + \cdots + b_m w_m$$

Replacing w_i above with the previous equation, we see that u is a linear combination of $X \setminus \{w_i\}$

Thus $X \setminus \{w_i\} = \text{span}(W)$

Shrinking Theorem: Let $X = \{w_1, \dots, w_m\}$ span W. Then for some subset $Y \subseteq X$ is a basis of W

Proof:

Case 0: If X is linearly independent, then X is a basis by definition

Otherwise, apply the shrinking lemma to get $X_1 = X \setminus \{w_i\}$, which spans W

Case 1: If X_1 is linearly independent, then X_1 is a basis

...

Since X is finite (it has m elements), we will stop eventually. Either

- Some X_i is linearly independent. Thus X_i is a basis for W
- Otherwise if we hit case m: $X_m = \emptyset$, which is linearly independent, and thus X_m spans $W = \{O\}$

Corollary: If $W = \text{span}(\{v_1, \dots, v_n\})$, then some subset of $\{v_1, \dots, v_n\}$ is a basis

- **Note**: In particular, W has to have a basis

Enlarging Lemma: Suppose $X = \{w_1, \dots, w_m\} \subseteq W$ and is linearly independent but doesn't span W. Then for any $w \in W \setminus \text{span}(X)$, $X \cup \{w\}$ is still linearly independent

Proof: Suppose $a_1 w_1 + \cdots + a_m w_m + bw = O$. We show that $a_1 = \cdots = a_m = b = 0$

Suppose BWOC, $b \neq 0$, then we can solve for w

$$w = \frac{-a_1}{b} w_1 + \cdots + \frac{-a_m}{b} w_m$$

Which means that $w \in \text{span}(X)$. Contradiction

Thus $b = 0$. This gives

$$a_1 w_1 + \cdots + a_m w_m + 0w = O$$

Since $X = \{w_1, \dots, w_m\}$ is linearly independent, we also have $a_1 = \cdots = a_m = 0$

Thus $X \cup \{w\}$ is linearly independent

Main Question: Does the enlarging process above terminate? After some steps, do we get a set $\{w_1, \dots, w_m\}$ that spans W?

Exchanging Lemma: Let $X = \{v_1, \dots, v_n\}$ be any basis for W. Choose any $w \in W$ but $w \notin \text{span}(\{v_k, \dots, v_n\})$. Then $\exists v_i, i < k$, such that $Y = (X \setminus \{v_i\}) \cup \{w\}$ is still a basis

- **Note**: If $k > n$, then $\{v_k, \dots, v_n\} = \emptyset$

Proof: First we show that $\text{span}(Y) = W$. Since X spans W, we can write

$$w = a_1 v_1 + \cdots + a_n v_n \implies v_1 = \frac{1}{a_1} w + \frac{-a_2}{a_1} v_2 + \cdots + \frac{-a_m}{a_1} v_m$$

Since $w \notin \text{span}(\{v_k, \dots, v_n\})$, we must have $a_i \neq 0$ for some $i < k$

WLOG, let $a_1 \neq 0$. We show that Y spans W

Since X spans W, for arbitrary $u \in W$, we have

$$u = d_1 v_1 + \cdots + d_n v_n$$

Replacing v_1 above with the previous equation, we see that u is a linear combination of elements of Y and thus $u \in \text{span}(Y)$

Thus $\text{span}(Y) = W$

Next we show that Y is linearly independent

Suppose we have

$$cw + b_2v_2 + b_nv_n = O$$

We show that $c = b_2 = \cdots = b_n = 0$

- If $c = 0 \implies b_2 = \cdots = b_n = 0$ since $\{b_2, \ldots, b_n\}$ is linearly independent
- Otherwise suppose $c \neq 0$, then we can solve for w

$$w = \frac{-b_2}{c}v_2 + \cdots + \frac{-b_n}{c}v_n \implies v_1 = \frac{1}{a_1}(\frac{-b_2}{c}v_2 + \cdots + \frac{-b_n}{c}v_n) + \frac{-a_1}{a_1}v_2 + \cdots + \frac{-a_m}{a_1v_m}$$

Thus v_1 is a linear combination of $\{v_2, \ldots, v_n\}$. Contradiction since we said X was linearly independent. Thus $c = 0$

Theorem: Let $X = \{v_1, \ldots, v_n\}$ be a basis for W, and let $\{w_1, \ldots, w_m\} \subseteq W$ be linearly independent. Then $m \leq n$

Proof: If $m < n$, we are done

Now assume $m \geq n$, we show that $m = n$

Since $\{w_1, \ldots, w_m\}$ is linearly independent, we have that $w_1 \neq O = \text{span}(\emptyset)$

Now apply the Exchanging Lemma to the basis X, with $k > n$ and w_1 Then $\exists v_i$ such that $X_1 = (X \setminus \{v_i\}) \cup \{w_1\}$ is a basis

After reindexing, we see that X_1 has $n-1$ vectors from X and 1 vector from w_1

Now take $k = n$. Since $\{w_1, \ldots, w_m\}$ is linearly independent, $w_2 \notin \text{span}(\{w_1\})$

Thus applying the Exchanging Lemma again, there exists $j < k = n$ such that $X_2 = (X_1 \setminus \{v_j\}) \cup \{w_2\}$ is a basis

Reindexing again, we get that $X_2 = \{v_1, \ldots, v_{n-2}, w_1, w_2\}$ is a basis

After n steps, X_n has no elements from X and $X_n = \{w_1, \ldots, w_n\}$ is a basis

Furthermore, we see that $w_m \in \text{span}(\{w_1, \ldots, w_n\})$, contradicting that $\{w_1, \ldots, w_m\}$ is linearly independent

Thus $m = n$

Corollary: If W is any K-vector space and some basis of W has n elements, then every basis of W has n elements

Definition - Finite Dimensional: Let W be a K-vector space. Then W is **finite dimensional** if some basis for W is finite

Definition - Dimension: Number of elements in any basis for a vector space W

Corollary: Suppose $\dim(W) = n$ and $X = \{w_1, \ldots, w_n\}$ are any n-vectors

1. If X spans W, then X is a basis for W
2. If X is linearly independent, then X is a basis for W

Proof:

1. By Shrinking Theorem, there exists a basis $Y \subseteq X$

   However, $|Y| < n$ contradicts that $\dim(W) = n$

   Thus $Y = X$, i.e. X is a basis

2. By Enlarging Lemma, we can expand X to a basis Y

   However, $|Y| > n$ contradicts that $\dim(W) = n$

   Thus $Y = X$, i.e. X is a basis

1.3.1 Toolbox Corollaries and Results

The following are useful corollaries that can be used to prove additional interesting results

Let V be a K-Vector Space with $\dim(V) = n$, i.e. V has some basis with n elements

1. Every basis for V has n elements
2. If $X \supseteq V$ and $\text{span}(X) = V$, then X has at least n elements and some subset $Y \subseteq X$ is a basis for V
3. If $Z \subseteq V$ is linearly independent, then Z has at most n elements and Z can be extended to a basis $Y \supseteq Z$ for V

Example: Let $V = R^3$. Since $\dim(V) = 3$, V has a basis with 3 elements

- Consider the **Standard Basis**: $B = \{(1,0,0), (0,1,0), (0,0,1)\}$

Suppose $X = \{v_1, v_2, v_3\} \subseteq V$ for arbitrary vectors

- If $\text{span}(X) = V$ then X is a basis
- If X is linearly independent, since $|X| = 3$, X is a basis for V

Example: Describe all subspaces $W \subseteq R^3$

Note: Since $\dim(V) = 3$, we must have $\dim(W) \leq \dim(V) = 3$

- Case 0: $\dim(W) = 0$

  Clearly $W = \{O\}$

- Case 1: $\dim(W) = 1$

  W is a line going through $(0,0,0)$

  Thus a basis for W will be $\{w\}$ for any nonzero $w \in W$

- Case 2: $\dim(W) = 2$

  W is a plane containing $(0,0,0)$

  Thus a basis for W will be any 2 element set $\{w_1, w_2\} \subseteq W$ such that

  - Neither element is O
  - w_2 is not a scalar multiple of w_1

- Case 3: $\dim(W) = 3$

  Only possibility is $W = V = R^3$

Examples: Consider subspaces of $\mathcal{F}(R)$ and look at small subspaces

- $W = \text{span}(\{e^x\}) = \{re^x \mid r \in R\}$

  This can be thought of as a 1-dimensional subpsace of $\mathcal{F}(R)$

- $V = \text{span}(\{\sin(x), \cos(x)\}) = \{a\sin(x) + b\cos(x) \mid a, b \in R\}$

  Clearly $\dim(V) = 2$

  Consider $f(x) = \sin(x) \quad g(x) = \cos(x) \quad h(x) = 3\sin(x) - 2\cos(x)$

  Since $h = 3f + (-2)g$, $\{f, g, h\}$ is not linearly independent

  Thus $\text{span}(\{f, g, h\}) = \text{span}(\{f, g\})$

1.4 Direct Sums

Let V be a K-Vector Space with $\dim(V) = n$. Let $W \subseteq V$ be a subspace of V. Then $\dim(W) \leq n$

Now choose another subspace $U \subseteq V$

Note: $W \cap U \neq \emptyset$ since both must contain O

Thus the smallest we can make $W \cap U$ is $\{O\}$

Furthermore, it can be shown that both $U \cap W$ and $U + W$ are both subspaces of V

Definition - Direct Sum: $U \oplus W$ is called a **direct sum** if

- $U \oplus W = U + W$
- $U \cap W = \{O\}$

We often look at cases where $V = U \oplus W$

Example: Consider R^3 and let W be any plane containing $(0, 0, 0)$

If U is any line through $(0, 0, 0)$ such that $U \notin W$, then $R^3 = W \oplus U$

Theorem: Let V be a K-Vector Space with $\dim(V) = n$. Let $W \subseteq V$ be any subspace of V. Then there exists a subspace $U \subseteq V$ such that

$$V = U \oplus W$$

Proof: Choose any basis $Z = \{w_1, \dots, w_m\}$ of W (we know that $m \leq n$)

Now extend Z to $Y = Z \cup \{u_1, \dots, u_r\}$, which is a basis for V

Let $U = \text{span}(\{u_1, \dots, u_r\})$. Then U is a subspace of V and $\{u_1, \dots, u_r\}$ is a basis for U

- Show that $U \cap W = \{O\}$

  Choose $v \in U \cap W$

  Then we have $v = a_1 u_1 + \cdots + a_r u_r = b_1 w_1 + \cdots + b_m w_m$

  Since Y is a basis for V, then $\{u_1, \dots, u_r, b_1, \dots, b_m\}$ is linearly independent

  Thus $v - v = a_1 u_1 + \cdots + a_r u_r - b_1 w_1 - \cdots - b_m w_m = O \implies a_1 = \cdots = a_r = b_1 = \cdots = b_m = 0$

  Thus $v = O$

- Show that $V = U + W$

  Choose any $v \in V$

  Since Y is a basis for V

  $v = \underbrace{a_1 u_1 + \cdots + a_r u_r}_{u \in U} + \underbrace{b_1 w_1 + \cdots + b_m w_m}_{w \in W}$

  Thus $v = u + w \implies V = U + W$

2 Matrices

Definition - $\mathbf{m \times n}$ Matrix: Entries $\in K$ of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{bmatrix}$$

Example: $A = \begin{bmatrix} 4 & 0 & 2 \\ -1 & 3 & 6 \end{bmatrix}$ is a 2×3 matrix with entries $\in Q$

Note: Any 2×3 matrices can be added together componentwise or multiplied by a scalar, resulting in a 2×3 matrix

- Here the additive identity is $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
- Here the additive inverse of A (from previous example) is $-A = \begin{bmatrix} -4 & 0 & -2 \\ 1 & -3 & -6 \end{bmatrix}$

Thus $\mathrm{Mat}_{2\times 3}(K)$, the set of all 2×3 matrices with entries in K is a K-Vector Space

Here the basis is $B = \{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \}$

- Clearly spans since any 2×3 matrix $\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix}$ can be written as a linear combination of elements in B
- Clearly B is linearly independent since the only way to write O is to take each scalar $a_i = 0$

Thus $\dim(\mathrm{Mat}_{2\times 3}(K)) = 6$

Upshot: We can generalize the discussion above to show that $\mathrm{Mat}_{m\times n}(K)$ is a K-Vector Space of $\dim = m \times n$

Example: $\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} \}$, **Symmetric 2×2 matrices**, is a subspace of $\mathrm{Mat}_{2\times 2}(K)$, has dimension 3

Non-Example: $\mathrm{Mat}(K)$ is NOT a Vector Space since addition between 2×2 and 3×3 matrices is not defined

Notation: $A_i = (a_{i1}, \ldots, a_{in})$, the ith row vector, is a $1 \times n$ matrix

Notation: $A^j = (a_{1j}, \ldots, a_{mj})$, the jth column vector, is an $m \times 1$ matrix

Definition - Transpose: Given an $m \times n$ matrix A, the **transpose** A^t is an $n \times m$ matrix that swaps the rows and columns, and vice versa

- **Note**: If A is a square $n \times n$ matrix, then A^t is also a square $n \times n$ matrix

Example: $\begin{bmatrix} 4 & 0 & 3 \\ -1 & 3 & 0 \end{bmatrix}^t = \begin{bmatrix} 4 & -1 \\ 0 & 3 \\ 2 & 6 \end{bmatrix}$

Definition - Matrix Multiplication: An $m \times n$ matrix A can multiply with an $n \times k$ matrix B where

$$C_{il} = \sum_{d=1}^{n} a_{ij} b_{d,l}$$

- **Note**: If A, B are both $n \times n$ matrices, then AB is an $n \times n$ matrix

Upshot: Square matrices are closed under transposition and matrix multiplication

Example: $\begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 22 \\ 4 \end{bmatrix}$

2.1 Linear Equations

Consider the following system

$$5x_1 + 3x_2 - 6x_3 = 8$$
$$x_1 - 2x_2 + x_3 = 4$$

We can represent this using

$$A = \begin{bmatrix} 5 & 3 & -6 \\ 1 & -2 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad B = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \implies AX = B$$

3 Mappings

Definition - Function: Mapping between 2 sets D, R such that for each $x \in D$, there exists a unique $y \in R$ such that $f(x) = y$

$$F : D \to R$$

- **Note**: D here is the **domain** of F and R is the **range** of F

Definition - Image: $F(D) = \{F(x) \mid x \in D\} \subseteq R$

Example: $F : R \to R \qquad F(x) = x^2$

- $\text{Domain}(F) = \text{Range}(F) = R$
- Image of $F = \{y \in R \mid y \geq 0\} = [0, \infty)$

Example: $G[0, \infty) \to R \qquad G(x) = \sqrt{x}$

- Image of $G = [0, \infty)$

Example: $\mathcal{F} =$ all functions $F :\to R$

Let S be all “infinitely” differentiable functions

Let $\frac{d}{dx} : S \to S$ where $\frac{d}{dx}(f) = f'$

Thus $\frac{d}{dx}$ is a function

Example: $t : \text{Mat}_{2\times 3}(K) \to \text{Mat}_{3\times 2}(K)$

Then $t(A) = A^t$ is a function

Definition - Onto: A function $F : D \to R$ is **onto** if Image of $F = R$

Definition - 1-1: A function $F : D \to R$ is **1-1** if different elements from D get mapped to different elements of R

$$F(d) = F(e) \implies d = e$$

Definition - Bijection: A function that is both onto and 1-1

Definition - Inverse Function: If $F : D \to R$ is a bijection, there exists an inverse function $F^{-1} : R \to D$ such that

$$\forall r, \in R, F(F^{-1}(r)) = r$$
$$\forall d, \in D, F^{-1}(F(d)) = d$$

Definition - Linear Transformation: For fixed K-Vector Spaces V, W, a **linear transformation** $T : V \to W$ is a function satisfying

1. $\forall v_1, v_2 \in W$, $T(v_1 + v_2) = T(v_1) + T(v_2)$
2. $\forall c \in K, v \in W$, $T(cv) = cT(v)$

Examples

1. $F : R \to R, F(x) = x^2$
   - Not onto since x^2 cannot be negative
   - Not 1-1 since $1^2 = (-1)^2 = 1$
   - Not a linear transformation since $(1+2)^2 = 9 \neq 1^2 + 2^2$
2. $F : [0, \infty) \to R, F(x) = \sqrt{x}$
   - Not onto since x^2 cannot be negative
   - 1-1 since $\sqrt{x} = \sqrt{y} \implies x = y$
   - Not a linear transformation since $[0, \infty)$ isn't a Vector Space
3. Let S be the set of all infinite differentiable functions. Consider $\frac{d}{dx} : S \to S$ where $\frac{d}{dx}(f) = f'$
   - Onto by the Fundamental Theorem of Calculus
   - Not 1-1 since f and $f + 5$ share the same derivative
   - Is a linear transformation by addition and scalar multiplication properties of derivatives
4. Let C be the set of continuous functions on $[0, 1]$. Consider $I : C \to R, I(f) = \int_0^1 f(t)\, dt$
   - Onto since we can generate any value of R by taking the integral of the constant function
   - Not 1-1 since the definite integral of 2 functions could yield the same result
   - Is a linear transformation by additional and scalar multiplication properties of integrals
5. $I^* : G \to C, I^*(f) = \int_0^x f(t)\, dt$
   - Not onto since not all functions of $f(0) = 0$
   - 1-1 since indefinite integral yields a unique function
   - Is a linear transformation by additional and scalar multiplication properties of integrals
6. Fix $(4, 0, 2)$ and consider $T_{(4,0,2)} : R^3 \to R^3, T_{(4,0,2)}((x, y, z)) = (x + 4, y, z + 2)$
   - Clearly onto
   - Clearly 1-1
   - Not a linear transformation since $T_{(4,0,2)}((0, 0, 0) + (1, 1, 1)) = (5, 0, 3) \neq T_{(4,0,2)}((0, 0, 0)) + T_{(4,0,2)}((1, 1, 1))$
7. $E_\pi : R^3 \to R^3, E_\pi((x, y, z)) = (\pi x, \pi y, \pi z)$
   - Clearly onto
   - Clearly 1-1
   - Is a linear transformation since $E_\pi((a, b, c) + (d, e, f)) = (\pi(a + d), \pi(b + e), \pi(c + f)) = E_\pi((a, b, c)) + E_\pi((d, e, f))$

3.1 Consequences of Properties of Linear Transformations

Proposition: For any linear transformation $T : V \to W$, we have that

$$T(O_V) = O_W$$

Proof: Let $w = T(O_V)$

Since $O_V = 0 * O_V$, we have that

$$T(O_V) = T(0 * O_V) = 0 * T(O_V) = 0 * w = O_W$$

Proposition: $T(a_1v_1 + \cdots + a_nv_n) = a_1T(v_1) + \cdots + a_nT(v_n)$

Proof: Follows from linearly properties of linear transformations

- **Note**: If $x = \{v_1, \ldots, v_n\}$ is a basis for V and if $w_1, \ldots, w_n$ are arbitrary vectors in W, then there is a unique linear transformation $T : V \to W$ such that
$$T(v_1) = w_1, \ldots, T(v_n) = w_n$$

Lemma: $\mathrm{Im}(T)$ is a subspace of W

Proof: We show the necessary conditions for a subspace

- $+ : w_1, w_2 \in \mathrm{Im}(T) \implies \exists v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$

  Then $w_1 + w_2 = T(v_1) + T(v_2) = T(\underbrace{v_1 + v_2}_{\in V}) \in \mathrm{Im}(T)$
- $\times : w \in \mathrm{Im}(T) \implies \exists v \in V$ such that $T(v) = w$

  Then for $c \in K$, we have $cw = c(Tv) = T(\underbrace{cv}_{\in V}) \in \mathrm{Im}(T)$

Definition - Pull Back: Suppose $Y = \{w_1, \ldots, w_m\} \subseteq \mathrm{Im}(T)$. Then a **pull-back** is any set $\{v_1, \ldots, v_m\} \subseteq V$ such that

$$T(v_1) = w_1, \ldots, T(v_m) = w_m$$

Lemma: If $\{w_1, \ldots, w_m\}$ is linearly independent in $\mathrm{Im}(T)$ (or in W), then any pull back $\{v_1, \ldots, v_m\} \subseteq V$ is linearly independent in V

Proof: Let $a_1v_1 + \cdots + a_mv_m = O_V$

Thus $T(a_1, v_1 + \cdots + a_mv_m = O_V) = a_1w_1 + \cdots + a_mw_m = O_W$

Since $\{w_1, \ldots, w_m\}$ is linearly independent, we have $a_1 = \cdots = a_m = 0$ as desired

Pull Back Property: Suppose $\{w_1, \ldots, w_m\}$ is a basis for $\mathrm{Im}(T)$, and let $\{v_1, \ldots, v_m\} \subseteq V$ be any pull back. Furthermore, let $S = \mathrm{span}(\{v_1, \ldots, v_m\}) \subseteq V$ be a subspace. Then $\{v_1, \ldots, v_m\}$ is a basis for S

Proof: By the previous lemma, $\{v_1, \ldots, v_m\}$ is linearly independent

Furthermore, $\{v_1, \ldots, v_m\}$ spans S by definition

Corollary: If $T : V \to W$ is any linear transformation and if $\dim(V) = n$, then $\dim(\mathrm{Im}(T)) \leq n$

Proof: BWOC, suppose $\dim(\mathrm{Im}(T)) > n$, thus we can create a set of $n + 1$ linearly independent elements in $\mathrm{Im}(T)$.

By the Pull Back Property, this pulls back to $n + 1$ linearly independent elements in V. Contradiction since $n + 1 > n = \dim(V)$

Note: $T : V \to W$, where $T(v) = \{O_W\}$, is a linear transformation with $\dim(\text{Im}(T)) = 0$, regardless of the value of $\dim(V)$

3.2 Kernel

Definition - Kernel: For $T : V \to W$, the **kernel** $\text{Ker}(T) = \{v \in V \mid T(v) = O_W\}$

Proposition: $\text{Ker}(T)$ is a subspace of V

Proof: Clearly $O_V \in \text{Ker}(T)$

- $+$: For $v_1, v_2 \in \text{Ker}(T)$, we see that $T(v_1 + v_2) = T(v_1) + T(v_2) = O_W + O_W = O_W$. Thus $v_1 + v_2 \in \text{Ker}(T)$
- $\times$: For $c \in K$ and $v \in \text{Ker}(T)$, we see that $T(cv) = cT(v) = O_W$. Thus $cv \in \text{Ker}(V)$

Proposition: Let $T : V \to W$ be any linear transformation. For any basis $B = \{w_1, \dots, w_m\} \subseteq \text{Im}(T)$ and for any pullback $\{v_1, \dots, v_m\} \subseteq V$, we have

$$V = \text{Ker}(T) \oplus S \qquad S = \text{span}(\{v_1, \dots, v_m\})$$

Proof: We need to show $V = \text{Ker}(T) + S$ and $\text{Ker}(T) \cap S = \{O_V\}$

- Take arbitrary $v \in V \implies T(v) \in \text{Im}(T) = a_1 w_1 + \cdots + a_m w_m$

  Let $s = a_1 v_1 + \cdots a_m v_m \in S$.

  Then $T(s) = T(v) \implies T(v - s) = T(v) - T(s) = O_W \implies v - s \in \text{Ker}(T)$

  Let $u = v - s \in \text{Ker}(T)$

  Thus clearly $v = u + s$ for $u \in \text{Ker}(T)$ and $s \in S$
- Clearly $O_V \in \text{Ker}(T) \cap S$ since both are subspaces of V

  Take any arbitrary $v \in \text{Ker}(T) \cap S$

  $v \in S \implies v = b_1 v_1 + \cdots b_m v_m \implies T(v) = b_1 w_1 + \cdots b_m w_m$

  Since $v \in \text{Ker}(T)$, we have that $T(v) = O_W \implies b_1 = \cdots = b_m = 0$ since $\{w_1, \dots, w_m\}$ is linearly independent

  Thus we have $v = 0v_1 + \cdots + 0v_m = O_V \implies \text{Ker}(T) \cap S = \{O_V\}$

Thus we have shown the necessary properties for $V = \text{Ker}(T) \oplus S$

Theorem: $\dim(V) = \dim(\text{Ker}(T)) + \dim(\text{Im}(T))$

Proof: Choose a basis $B = \{w_1, \dots, w_m\}$ for $\text{Im}(T)$ and a pullback $\{v_1, \dots, v_m\}$

Let $S = \text{span}(\{v_1, \dots, v_m\})$

Since $V = \text{Ker}(T) \oplus S$, we have $\dim(\text{Ker}(T)) + \dim(S) = \dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$

3.2.1 Consequences of Kernel

Corollary 1: For linear $T : R^3 \to R^4$, T is NOT onto

Proof: $\dim(\text{Im}(T)) \leq \dim(R^3) = 3 < 4 \implies \text{Im}(T) \neq R^4 \implies T$ is NOT onto

Corollary 2: For linear $T : R^4 \to R^3$, T is NOT 1-1

Proof: $\dim(\text{Ker}(T)) + \underbrace{\dim(\text{Im}(T))}_{\leq 3} = \dim(R^4) = 4 \implies \dim(\text{Ker}(T)) \geq 1$

Thus $\text{Ker}(T)$ has something non-zero mapped to $O_W \implies T$ is NOT 1-1

Definition - Isomorphism: $T : V \to W$ such that T is linear transformation and a bijection

Corollary 3: $\dim(V) = \dim(W)$ and $T : V \to W$ is a linear transformation and 1-1 $\implies T$ is an isomorphism (i.e. T is onto)

Proof: $\dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$

But we know that $\dim(\text{Ker}(T)) = 0 \implies \dim(\text{Im}(T)) = \dim(V) = \dim(W)$

Furthermore $\text{Im}(T)$ is a subspace of W and $\dim(\text{Im}(T)) = \dim(W) \implies T$ is onto

Corollary 4: $\dim(V) = \dim(W)$ and $T : V \to W$ is a linear transformation and onto $\implies T$ is an isomorphism (i.e. T is 1-1)

Proof: $\dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$

But we know that $\dim(\text{Im}(T)) = \dim(V) \implies \dim(\text{Ker}(T)) = 0$

3.3 Compositions and Inverse Linear Mappings

Consider Vector Spaces U, V, W and linear transformations $T : U \to V$ and $S : V \to W$

Proposition: $S \circ T : V \to W$ is a linear transformation

Proof:

- $+$: For $u_1, u_2 \in U$ we have that

$$\begin{aligned} S \circ T(u_1 + u_2) &= S(T(u_1 + u_2)) \\ &= S(T(u_1) + T(u_2)) \\ &= S(T(u_1)) + S(T(u_2)) \\ &= S \circ T(u_1) + S \circ T(u_2) \end{aligned}$$

- $\times$: For $u \in U$ and $c \in K$

$$\begin{aligned} S \circ T(cu) &= S(T(cu)) \\ &= S(cT(u)) \\ &= cS(T(u)) \\ &= cS \circ T(u) \end{aligned}$$

Thus $S \circ T : V \to W$ is a linear transformation

Definition - Inverse Mapping: $T^{-1} : W \to V$ where $T^{-1}(w) =$ the unique $v \in V$ such that $T(v) = w$

Proposition: $T^{-1} : W \to V$ is a linear transformation (and thus an isomorphism)

Proof:

- $+$: Take $w_1, w_2 \in W$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$ for $v_1, v_2 \in V$. Then we see that

  $$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2$$

  However, by definition of inverse mapping, $v_1 + v_2$ is the unique element such that $T(v_1 + v_2) = w_1 + w_2$

  Thus by definition of T^{-1}, we have that $T^{-1}(w_1 + w_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2)$

- $\times$: Similar

4 Linear Maps and Matrices

Definition - $\mathbf{L_A}$: For a $m \times n$ matrix A, L_A determines a linear transformation from $R^n \to R^m$

Example: Consider $L_A : R^3 \to R^2$ where $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix}$

Then we see that $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + 3y + z \\ 4y + 2z \end{bmatrix}$

It can be clearly shown that L_A is a linear transformation (follows from logic of dot products)

4.1 Bases, Matrices, and Linear Maps

For a given transformation $T : V \to W$, the matrix of T with respect to the standard basis is given by

$$A = (T(E_1), \dots, T(E_n))$$

Example: $T : R^2 \to R^3 \qquad T(x, y) = (5x + y, x - y, x)$

$T(E_1) = (5, 1, 1) \qquad T(E_2) = (1, -1, 0)$

Thus we see that $A = \begin{bmatrix} 5 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$

- $T({}^t(3, 2)) = \begin{bmatrix} 5 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} =^t (17, 1, 3)$

Example: $T : R^2 \to R^2$ where we stretch the x-coordinate by 2

$T({}^t(1, 0)) =^t (2, 0) \qquad T({}^t(0, 1)) =^t (0, 1)$

Thus we see that $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

Example: $S \circ T : R^2 \to R^2$ where we first stretch by x by 3 then stretch y by 3

$T({}^t(1, 0)) =^t (2, 0) \qquad T({}^t(0, 1)) =^t (0, 3)$

Thus we see that $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Upshot: Applying functions just corresponds to matrix multiplication $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Example: Fix $\theta \in R$, then rotate by θ

$R_\theta({}^t(1,0)) = {}^t(\cos(\theta), \sin(\theta)) \qquad R_\theta({}^t(0,1)) = {}^t(-\sin(\theta), \cos(\theta))$

Thus $A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$

Thus given any ${}^t(x,y) \in R^2$, we see that $T_\theta({}^t(x,y)) = \begin{bmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{bmatrix}$

Example: Stretch x by 2, rotate by $\pi/4$, and stretch y by 3

$A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

Note: Given $T : K^n \to K^m$, the matrix A for T depends on our choosing of bases for K^n and K^m

Example: $T : R^2 \to R^3 \qquad T(x,y) = (5x+y, x-y, x)$

Let $B = \{\underbrace{(1,4)}_{v_1}, \underbrace{(3,0)}_{v_2}\}$ be a basis for R^2 and $B' = \{\underbrace{(3,0,0)}_{w_1}, \underbrace{(0,5,0)}_{w_2}, \underbrace{(0,0,1)}_{w_3}\}$ be a basis for R^3

We can define a matrix of T with respect to B and B'

$$M^B_{B'}(T) = (\underbrace{T(v_1) \qquad T(v_2)}_{\text{in terms of } w_1, w_2, w_3})$$

$T(v_1) = T(1,4) = (9,-3,1) = 3w_1 - \frac{3}{5}w_2 + w_3$

$T(v_2) = T(3,0) = (15,3,3) = 5w_1 + \frac{3}{5}w_2 + 3w_3$

Thus we see that $M^B_{B'}(T) = \begin{bmatrix} 3 & 5 \\ -3/5 & 3/5 \\ 1 & 3 \end{bmatrix}$

Upshot: Any vector, written in B coordinates, when multiplied by this matrix, yields an answer in B' coordinates. Thus for $v = av_1 + bv_2$, we have

$$T(v) = (3a+5b)w_1 + (-3/5a + 3/5b)w_2 + (a+3b)w_3$$

- As a sanity check, for $v = (5,8) \in R^2$
  - Normal Transformation: $T(v) = (33,-3,5)$
  - Linear Map: writing v in terms of v_1, v_2, we get $(5,8) = a(1,4) + b(3,0) \implies (5,8) = 2(1,4) + 2(3,0)$
    Thus we have

$$M^B_{B'}(T) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ -3/5 & 3/5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ -3/5 \\ 5 \end{bmatrix} \implies 11(3,0,0) - 3/5(0,5,0) + 5(0,0,1) = (33,-3,5)$$

Example: Consider $P_n = \{a_0 + a_1x + \cdots + a_nx^n \mid a_i \in R\}$

It's easily verifiable that P_n is a subspace of $\mathcal{F}(R)$. Furthermore, the basis for P_n is $\{1, x, \ldots, x^n\} \implies \dim(P_n) = n+1$

Let $D : P_2 \to P_2$ be the derivative

$$D(a_0 + a_1x + a_2x^2 = a_1 + 2a_2x)$$

Easily verifiable that D is a linear transformation. Consider what is the matrix of D with respect to $B = \{1, x, x^2\}$?

$$A = \begin{bmatrix} D(1) & D(x) & D(x^2) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus we see that for $p(x) = 5 + 3x + 4x^2$,

$$D(p(x)) = 3 + 8x = 5(0,0,0) + 3(0,1,0) + 4(0,2,0)$$

Upshot: For a linear transformation $T : V \to W$, with $\dim(V) = n$ and $\dim(W) = m$, if $B = \{v_1, \ldots, v_n\}$ and $B' = \{w_1, \ldots, w_m\}$ are bases for V, W, then

$$M_{B'}^{B}(T) = \begin{bmatrix} T(v_1) & T(v_2) & \cdots & T(V_n) \end{bmatrix}$$

is a $m \times n$ matrix with column vectors containing coefficients of $T(v_1)$ WRT B'

Furthermore, for any $v \in V, v = x_1v_1 + \cdots + x_nv_n$, we have

$$M_{B'}^{B}(T) \begin{bmatrix} x_1 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \cdots \\ y_m \end{bmatrix}$$

Thus $T(v) = y_1w_1 + \cdots + y_mw_m$ (**Note** coordinate is WRT to B')

Definition - Change of Basis: Let $B = \{v_1, \ldots, v_n\}$ and $B' = \{w_1, \ldots, w_n\}$ be basis for the same vector space V, and let $T : V \to V$ be the identity mapping. Then

$$M_{B'}^{B}(\mathrm{id}) = \underbrace{\begin{bmatrix} \mathrm{id}(v_1) & \mathrm{id}(v_2) & \cdots & \mathrm{id}(v_n) \end{bmatrix}}_{\text{WRT } B'}$$

is the **Change of Basis** matrix for V

Example: Let $V = P_1 = \{a_0 + a_1 x \mid a_i \in R\}$ and let $B = \{1, x\}$ and $B' = \{3 + x, 5 + 2x\}$, which are both bases for V

$1 = a(3 + x) + b(5 + 2x) \implies a = 2, b = -1 \implies 1 = 2(3 + x) - (5 + 2x)$

$x = c(3 + x) + d(5 + 2x) \implies c = -5, d = 3 \implies x = -5(3 + x) + 3(5 + 2x)$

$$M_{B'}^{B}(\text{id}) = \underbrace{\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}}_{\text{WRT } B'}$$

Furthermore, consider

$$M_{B}^{B'}(\text{id}) = \underbrace{\begin{bmatrix} \text{id}(w_1) & \text{id}(w_2) \end{bmatrix}}_{\text{WRT } B} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$$

Finally, we see that $M_{B}^{B'}(M_{B'}^{B}(\text{id})) = \text{id}$

Thus the inverse of $M_{B'}^{B}$ is $M_{B}^{B'}$

5 Scalar Products and Orthogonality

5.1 Scalar Products

Definition - Scalar Product: For a Vector Space V, we define $\langle , \rangle : V \times V \to K$

- **Example**: Think of dot products in $R^n \times R^n \to R$

Properties of Scalar Products

1. $\langle v, w \rangle = \langle w, v \rangle$
2. $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
3. $\langle v, cw \rangle = c\langle v, w \rangle \qquad \langle cv, w \rangle = c\langle v, w \rangle$

Consequences of Properties

- $\forall v_1, v_2, w \in V, \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$

  Proof: Follows from applying properties 1 and 2
- $\forall v \in V, \langle v, O_V \rangle = 0 = \langle O_V, v \rangle$

  Proof: For any $w \in V$, we have $\langle v, O_V \rangle = \langle v, 0w \rangle = 0\langle v, w \rangle$

Definition - Non-Degenerate: Scalar product that satisfies $\forall v \neq 0, \exists w \in V$ such that $\langle v, w \rangle \neq 0$

Example: $\mathcal{F}([0, 1])$, all functions $f : [0, 1] \to R$

Let $C([0, 1])$ be the set of all continuous functions $f : [0, 1] \to R$, which is clearly an R subspace

Now define $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$. We claim that this is a scalar product

Proof:

- $\int_0^1 f(x)g(x)\,dx = \int_0^1 g(x)f(x)\,dx$ so property 1 holds
- $\int_0^1 f(x)(g_1(x) + g_2(x))\,dx = \int_0^1 f(x)g_1(x)\,dx + \int_0^1 f(x)g_2(x)\,dx$ so property 2 holds
- $\int_0^1 f(x)cg(x)\,dx = c\int_0^1 f(x)g(x)$ so property 3 holds

We also claim that $\langle f, g \rangle$ is non-degenerate since for $f \neq 0$, we have $\langle f, f \rangle = \int_0^1 f(x)^2$, which is always > 0 and is continuous

Example: $f(x) = 2x + 3 \qquad g(x) = x^2$

$\langle 2x+3, x^2 \rangle = \int_0^1 (2x+3)x^2 \, dx = 3/2$

Defintion - Orthogonal: Elements $v, w \in V$ are **orthogonal**, denoted $v \perp w$, if $\langle v, w \rangle = 0$

Definition - Orthogonal Complement: Suppose $W \subseteq V$ is a subspace, then the **orthogonal complement** of W is

$$W^\perp = \{v \in V \mid \forall w \in W, v \perp w\}$$

- **Note**: $W^\perp \subseteq V$ is a subspace

Definition - Positive Definite: Scalar product that satisfies $\forall v \neq O, \langle v, v \rangle > 0$. Otherwise $\langle v, v \rangle = 0 \implies v = O$

Definition - Length: $\|v\| = \sqrt{\langle v, v \rangle}$

- Length between v and w: $\|v - w\|$
- $\|cv\| = |c|\|v\|$
- $\|v+w\|^2 = \langle v+w, v+w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2$
- $v \perp w \implies \langle v, w \rangle = 0 \implies \|v+w\|^2 = \|v-w\|^2 = \|v\|^2 + \|w\|^2$

Pythagoras Theorem: For $v \perp w$,

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

Proof:

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \|w\|^2 \end{aligned}$$

Parallelogram Law: For any $v, w \in V$, we have

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

Proof: Follows from the definition/properties of length

Definition - Unit Vector: $v \in V$ such that $\|v\| = 1$

- If $v \neq O$, then $(\frac{1}{\|v\|})v$ is a unit vector

Let $w \in V$. For any $v \in V$ there exists $c \in K$ such that $v - cw \perp w$. That is $\langle v - cw, w \rangle = 0$

Definition - Projection: $\text{proj}_w v$ represents v as a scalar multiple of w where $\text{proj}_w v = (\frac{\langle v, w \rangle}{\langle w, w \rangle})w$

- Definition comes from creating a right triangle where $v - cw \perp w \implies \langle v - cw, w \rangle = 0$

  Thus we have $\langle v, w \rangle - \langle cw, w \rangle = \langle v, w \rangle - c\langle w, w \rangle \implies c = \frac{\langle v, w \rangle}{\langle w, w \rangle}$
- Special case where $\langle w, w \rangle = 1 \implies \text{proj}_w v = \langle v, w \rangle w$

Schwartz Inequality: For any $v, w \in V$ we have

$$|\langle v, w\rangle| \leq \|v\|\|w\|$$

Proof: If $w = O$, then $|\langle v, w\rangle| = 0 \leq \|v\|\|w\| = 0$

Otherwise, assume that w is a unit vector. Using the definition of projection, we have $cw \perp v - cw$. Thus we see

$$\begin{aligned}
\|v\|^2 &= \|v - cw\|^2 + \|cw\|^2 \\
&= \|v - cw\|^2 + c^2 \\
&\geq c^2 \\
\implies \|v\| &\geq c = \frac{\langle v, w\rangle}{\langle w, w\rangle} \\
\implies \langle v, w\rangle &\leq \|v\|\|w\|
\end{aligned}$$

Triangle Inequality: For $v, w \in V$, we have

$$\|v + w\| \leq \|v\| + \|w\|$$

Proof:

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w\rangle \\
&= \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \\
&\leq \|v\|^2 + \underbrace{2\|v\|\|w\|}_{\text{by Schwartz}} + \|w\|^2 \\
&\leq (\|v\| + \|w\|)^2 \\
\implies \|v + w\| &\leq \|v\| + \|w\|
\end{aligned}$$

Proposition: Suppose $\{w_1, \ldots, w_r\} \subseteq V$ is pairwise orthogonal and assume that each $w_i \neq O$. Then $\{w_1, \ldots, w_r\}$ is linearly independent

Proof: Let $a_1w_1 + \cdots + a_rw_r = O_V$. Then we have

$$\langle w_i, a_1w_1 + \cdots + a_rw_r\rangle = \langle w_i, a_1w_1\rangle + \cdots + \langle w_i, a_nw_n\rangle = 0 \qquad \text{since each w is pairwise orthogonal}$$

Thus $\langle w_i, a_iw_i\rangle = 0 \implies a\langle w_i, w_i\rangle = 0 \implies a_i = 0$ since $\langle w_i, w_i\rangle > 0$ since positive definite

Let $W = \operatorname{span}(\{w_1, \ldots, w_r\}) \subseteq V$. Then clearly $\dim(W) = r$

Now take $v \in V$ and define $\operatorname*{proj}_W v = \sum_{i=1}^{r} c_iw_i$ where $c_iw_i = \operatorname{proj}_{w_i} v$

Clearly $\operatorname{proj}_W v \in W$

Proposition: $\left(v - \sum_{j=1}^{r} c_j w_j\right) \perp$ each w_i

Proof: Fix i, then

$$\sum_{j=1}^{r} c_j w_j = c_i w_i + \sum_{j \neq i} c_j w_j$$

Now take

$$v - \sum_{j=1}^{r} c_j w_j = (v - c_i w_i) - \sum_{j \neq i} c_j w_j$$

and take the inner product with w_i

$$\underbrace{\langle w_i, v - c_i w_i \rangle}_{\text{0 b/c of projection}} - \langle w_i, \underbrace{\sum_{j \neq i} c_j w_j \rangle}_{\text{0 b/c orthogonal}}$$

Thus we have $w_i \perp v - \sum_{j=1}^{r} c_j w_j$

Corollary: $(v - \sum_{j=1}^{r} c_j w_j) \perp$ every $w \in W$

Proof: Since each w_i in the basis is orthogonal to $v - \sum_{j=1}^{r} c_j w_j$, we must have

$$\langle w, v - \sum_{j=1}^{r} c_j w_j \rangle = 0$$

Corollary: $(v - \sum_{j=1}^{r} c_j w_j) \in W^{\perp}$

Proof: Follows from the previous corollary

Geometric Interpretation: For any $v \in V$, $\operatorname{proj}_W v$ is the closest point to v in W

$$\|v - \operatorname*{proj}_W v\| \leq \|v - w\| \qquad \text{for any arbitrary } w \in W$$

Proof: Choose any $w \in W = \operatorname{span}(\{w_1, \dots, w_r\})$, then $w = \displaystyle\sum_{i=1}^{r} a_i w_i$. Then we have

$$\begin{aligned}
\|v - w\|^2 &= \|v - \sum_{i=1}^{r} a_i w_i\|^2 \\
&= \| \underbrace{v - \sum_{i=1}^{r} c_i w_i}_{\in W^\perp} + \underbrace{\sum_{i=1}^{r} (c_i a_i) w_i}_{\in W} \|^2 \\
&= \|v - \sum_{i=1}^{r} c_i w_i\|^2 + \|\sum_{i=1}^{r} (c_i - a_i) w_i\|^2 \qquad \text{by Pythagoras}
\end{aligned}$$

Thus $\|v - w\|^2 \geq \|v - \displaystyle\sum_{i=1}^{r} c_i w_i\|^2 \implies \|v - w\| \geq \|v - \displaystyle\sum_{i=1}^{r} c_i w_i\|$

Corollary: Suppose $w \in W$, then $\operatorname{proj}_W w$ is the element of W closest to w

But we have $w = \displaystyle\sum_{i=1}^{r} c_i w_i \implies c_i = \dfrac{\langle w, w_i \rangle}{\|w_i\|^2}$

5.2 Orthonormal Basis

Definition - Orthonormal Basis: $\{w_1, \dots, w_r\} \subseteq W$ is an **orthonormal basis** if

1. $\{w_1, \dots, w_r\}$ are pairwise orthogonal and none are zero
2. $\|w_i\| = 1$ for $i \in \{1, \dots, r\}$

Corollary: If $\{w_1, \dots, w_r\}$ is orthonormal, then $\forall w \in W, w = \displaystyle\sum_{i=1}^{r} \langle w, w_i \rangle w_i$

Gram-Schmidt Process: Turn any basis $B = \{v_1, \dots, v_n\}$ into an orthonormal basis $B' = \{u_1, \dots, u_n\}$

1. Given v_1, let $u_1 = \frac{1}{\|v_1\|} v_1$. Then we have $\operatorname{span}(\{u_1\}) = \operatorname{span}(\{v_1\})$
2. Let $p_2 = v_2 - \operatorname{proj}_{u_1} v_2 = v_2 - \langle v_2, u_1 \rangle u_1$

   Now let $u_2 = \frac{1}{\|p_2\|} p_2$. Then $\operatorname{span}(\{u_1, u_2\}) = \operatorname{span}(\{v_1, v_2\})$
3. Let $p_3 = v_3 - \operatorname{proj}_{\operatorname{span}(\{u_1, u_2\})} v_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$

   Now let $u_3 = \frac{1}{\|p_3\|} p_3$. Then $\operatorname{span}(\{u_1, u_2, u_3\}) = \operatorname{span}(\{v_1, v_2, v_3\})$
4. Repeat

Upshot: Any finite R-Vector Space V with a positive definite inner product has an orthonormal basis

Theorem: Let V be a finite dimension R Vector Space with a positive definite scalar product. Then for any subspace $W \subseteq V$

$$V = W \oplus W^{\perp}$$

Proof:

- Show that $V = W + W^{\perp}$

  Choose $v \in V$ and let $w^* = \text{proj}_W v \in W$. Then $v - w^* \in W^{\perp}$

  Thus $v = \underbrace{w^*}_{\in W} + \underbrace{(v - w^*)}_{\in W^{\perp}}$

- Show that $W \cap W^{\perp} = \{O\}$

  Choose $w \in W \cap W^{\perp}$

  Since $w \in W^{\perp}$, w is orthogonal to all vectors in W

  In particular, $w \perp w \implies \langle w, w \rangle = 0 \implies w = O$ since the scalar product is positive definite

Corollary: If $W \subseteq V$ is a subspace, then

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

5.3 Application to Linear Equations: Rank

Let A be an $m \times n$ matrix with entries in R

- Let $C_A \subseteq R^m$ be the span of column vectors of A
- Let $R_A \subseteq R^n$ be the span of row vectors of A
- Let $\text{Null}(A) = \{v \in R^n \mid Av = O\}$

Recall that any $m \times n$ matrix A describes a linear transformation $L_A : R^n \to R^m$ where $L_A(v) = Av \in R^m$

Thus $\text{Im}(L_A) = C_A$

Furthermore, $\text{Ker}(L_A) = \{v \in R^n \mid Av = O\} = \text{Null}(A)$

Thus we have

$$\begin{aligned} \dim(R^n) &= \dim(\text{Im}(L_A)) + \dim(\text{Ker}(L_A)) \\ &= \dim(C_A) + \dim(\text{Null}(A)) \end{aligned}$$

Now consider using scalar products

Take $v \in \text{Null}(A)$. Thus $Av = O$

Thus $A_i \cdot v = 0 \iff A_i \perp v \iff v \perp \text{all } u \in R_A \implies v \in (R_A)^{\perp}$

Thus $\text{Null}(A) = \text{Ker}(A) = (R_A)^{\perp}$

Thus $R_A \subseteq R^n$ is a subspace of R^n.

Thus we have

$$\begin{aligned} \dim(R^n) &= \dim(R_A) + \dim((R_A)^{\perp}) \\ n &= \dim(R_A) + \dim(\text{Null}(A)) \end{aligned}$$

Thus we have $\dim(R_A) = \dim(C_A)$

Definition - Rank: The **rank** of a matrix A is $\dim(R_A) = \dim(C_A)$

5.4 Scalar Products Under Complex Numbers

We want a positive definite scalar product for C

Take the **complex conjugate**

$$(a+bi)(a-bi) = a^2 + b^2$$

Then we see that

$$\|z\| = \sqrt{\langle z, \bar{z} \rangle} \in R$$

Definition - Hermitian Inner Product: For $(y_1, \dots, y_n)$ and $(z_1, \dots, z_n) \in C^n$, define

$$\langle y, z \rangle = y_1 \overline{z_1} + \cdots + y_n \overline{z_n}$$

- **Note**: This is NOT a scalar product since $\langle y, z \rangle \neq \langle z, y \rangle$

Now we list the properties of the Hermitian Inner Product

- $\langle w, v \rangle = \overline{\langle v, w \rangle}$
- $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
- $\langle cv, w \rangle = c\langle v, w \rangle$ AND $\langle v, cw \rangle = \bar{c}\langle v, w \rangle$

Proposition: The Hermitian Inner Product is positive definite

Proof: We look at

$$\langle v, v \rangle = x_1 \overline{x_1} + \cdots + x_n \overline{x_n} = \|x_1\|^2 + \cdots + \|x_n\|^n \in R$$

We see that $\langle v, v \rangle \geq 0$. If it happens that $\langle v, v \rangle = 0 \implies x_1 = \cdots = x_n = 0$

5.5 General Orthogonal Bases

5.5.1 Properties and Types of Scalar Products

Repeating a lot of what was stated above for clarity

A **scalar product** satisfies

1. Symmetry: $\langle v, w \rangle = \langle w, v \rangle$
2. Linear: $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
3. Scalar $\langle cv, w \rangle = c\langle v, w \rangle = \langle v, cw \rangle$

Types of scalar products (each progressively weaker):

- **Positive Definite**: $\forall v \in V, \langle v, v \rangle \geq 0$ AND $\langle v, v \rangle = 0 \implies v = O$
- **Non-Degenerate**: For $v \neq O, \exists w \in V$ such that $\langle v, w \rangle \neq 0$
- **Non-Trivial**: $\exists v, w \in V$ such that $\langle v, w \rangle \neq 0$

Upshot: positive definite $\implies$ non-degenerate $\implies$ non-trivial

We also consider **Trivial Scalar Products** where $\forall v, w \in V$, we have $\langle v, w \rangle = 0$

For a positive definite $\langle , \rangle$, we proved that

1. Every finite dimentional Vector Space V has an orthonormal basis (**Gram Schmidt Process**)
2. For any subspace $W \subseteq V$, we have $V = W \oplus W^{\perp}$ (**Projection**)

Observation: If $\langle , \rangle$ is trivial, then any basis of V is orthogonal

Lemma: Suppose $\langle v, v \rangle = 0$ for all $v \in V$, then $\langle , \rangle$ is trivial

Proof: Choose any $v, w \in V$. Then we see

$$\langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle$$

Thus we have

$$\langle v, w \rangle = \frac{1}{2}(\langle v + w, v + w \rangle - \langle v, v \rangle - \langle w, w \rangle) = 0$$

Corollary: If $\langle v, v \rangle = 0$ for all $v \in V$, then any basis of V is orthogonal

Proof: Since $\langle , \rangle$ is trivial (shown from the Lemma), by the observation above, any basis of V is orthogonal

Theorem 1: If $\langle , \rangle$ is any scalar product on V, then V has an orthogonal basis

Proof: By Induction on $n = \dim(V)$

Claim: If $\langle , \rangle$ is any scalar product on any finite dimensional Vector Space V with $\dim(V) \leq n$, then V has an orthogonal basis

Base Case: $n = 0 : \dim(V) \implies B = \{\}$ is a basis and is an orthogonal basis

Base Case: $n = 1 : \dim(V) = 1 \implies \{v_1\}$ is an orthogonal basis for $v_1 \in V, v_1 \neq 0$

IH: Assume the claim holds for $\dim(V) = n - 1$

IS: Suppose $\dim(V) = n$

- Case 1: $\forall v \in V, \langle v, v \rangle = 0$. Then by the preceding Lemma, $\langle , \rangle$ is trivial and any basis for V is an orthogonal basis
- Case 2: $\exists v_1 \in V$ such that $\langle v_1, v_1 \rangle \neq 0$

  Let $V_1 = \operatorname{span}(\{v_1\}) \subseteq V$ be a subspace. We show that $V = V_1 \oplus V_1^{\perp}$

  - Show that $V = V_1 + V_1^{\perp}$

    Choose $v \in V$. Since $\langle v_1, v_1 \rangle \neq 0$ we can use projection: $\operatorname{proj}_{v_1} v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \in V_1$

    Thus $(v - \operatorname{proj}_{v_1} v) \perp v_1 \implies (v - \operatorname{proj}_{v_1} v) \in V_1^{\perp}$

    Thus $v = \underbrace{(\operatorname{proj}_{v_1} v)}_{\in V_1} + \underbrace{(v - \operatorname{proj}_{v_1} v)}_{\in V_1^{\perp}}$

  - Show $V_1 \cap V_1^{\perp} = \{O\}$

    Choose $v \in V_1 \cap V_1^{\perp}$

    $v \in V_1^{\perp}$ and $v \in V_1 \implies v \perp v \implies \langle v, v \rangle = 0$

    However, $v \in V_1 \implies v = dv_1 \implies 0 = \langle v, v \rangle = \langle dv_1, dv_1 \rangle = d^2 \underbrace{\langle v_1, v_1 \rangle}_{\neq 0}$

    Thus we see that $d = 0 \implies v = O$

  Now we have $\dim(V) = \dim(V_1) + \dim(V_1^{\perp}) \implies \dim(V_1^{\perp}) = n - 1$ which by IH has an orthogonal basis $\{v_2, \ldots, v_n\}$

  Finally, since $v_1 \perp v_i$ for $2 \leq i \leq n$, we see that $\{v_1, v_2, \ldots, v_n\}$ is a orthogonal basis for V

Definition - Dual Space: K-Vector Space $V^* = \mathcal{L}(V, K)$ where each element of V^* is a linear transformation $\phi : V \to K$

- **Note**: For any $w_1, \ldots, w_n \in W$, there is exactly one Linear Transformation $T : V \to W$ such that $T(v_i) = w_i$ for $1 \leq i \leq n$

Example: Let $B = \{v_1, \dots, v_n\}$ be a basis for V and take

$$\begin{aligned} \phi_1 : V \to K \qquad & \phi_1(v) = \phi_1(a_1 v_1 + \cdots + a_n v_n) = a_1 \\ \phi_2 : V \to K \qquad & \phi_2(v) = \phi_2(a_1 v_1 + \cdots + a_n v_n) = a_2 \\ & \dots \end{aligned}$$

Thus we see that $\phi_i(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

Let $B' = \{\phi_1, \dots, \phi_n\}$. Then we see that B' is a basis for V^*

- Show linear independence: Take $a_i \in K$ such that $\underbrace{O}_{O \text{ mapping}} = \underbrace{a_1\phi_1 + \cdots + a_n\phi_n}_{\text{mapping}}$

  This equality means that $\forall w \in V$, we have $(a_1\phi_1 + \cdots + a_n\phi_n)(w) = O(w)$

  Now applying the transformation to v_1, we see that $a_1\phi_1(v_1) = O(v_1) = 0 \implies a_1 = 0$

  Similar logic shows that $a_i = 0$ for $1 \leq i \leq n$

- Show B' spans $\mathcal{L}(V, K)$

  Choose any $T \in \mathcal{L}(V, K)$. Then we see

  $T(v_1) = b_1 \in K, \dots, T(v_n) = b_n \in K$

  Now let $\phi^* = b_1\phi_1 + \cdots + b_n\phi_n$. Clearly $\phi* \in \text{span}(B')$

  We show that $\phi^* = T$ (they need to agree on all input)

  It suffices so show that $\phi^*(v_j) = T(v_j)$ for $v_j \in B$ since B is a basis of V

  Simple calculations show that $\phi^*(v_j) = (b_1\phi_1 + \cdots + b_n\phi_n)(v_j) = b_j = T(v_j)$

  Thus $T' \in \text{span}(B)$

Corollary: $\dim(V^*) = \dim(V) = n$ (so same size as basis)

Corollary: V is isomorphic to V^*. Namely, there exists a 1-1, onto linear transformation $F : V \to V^*$ where

$$F(v_1) = \phi_1, \dots, F(v_n) = \phi_n$$

These ϕ_i uniquely describe F

Consider a subspace $W \subseteq V$

Definition - Annihilator: $\text{Ann}(W) = \{\phi \in V^* \mid \forall w \in W, \phi(w) = 0\}$, so the set of linear transformations ϕ in V^* such that $W \subseteq \text{Ker}(\phi)$

Annihilator Theorem: For any $W \subseteq V$

$$\dim(W) + \dim(\text{Ann}(W)) = \dim(V) = n$$

Proof: Choose a basis for W, $\{w_1, \ldots, w_r\}$

Now extend it to a basis for V, $B = \{w_1, \ldots, w_r, w_{r+1}, \ldots, w_n\}$

Let $B' = \{\phi_1, \ldots, \phi_n\}$ be the dual basis of V^* corresponding to B

We claim that $\{\phi_{r+1}, \ldots, \phi_n\}$ is a basis for $\text{Ann}(W)$

- For any $w \in W$, $w = a_1 w_1 + \cdots + a_r w_r$, and $j \geq r+1$, we have that $w_j = O \implies \phi_j(w) = 0 \implies \{\phi_{r+1}, \ldots, \phi_n\} \subseteq \text{Ann}(W)$
- $\{\phi_{r+1}, \ldots, \phi_n\}$ is linearly independent since B' is linearly independent
- To show that $\text{span}(\{\phi_{r+1}, \ldots, \phi_n\}) = \text{Ann}(W)$

  Take $T \in \text{Ann}(W) \implies T : V \to K$ is a linear transformation

  Furthermore, we have $T(w_1) = 0, \ldots, T(w_r) = 0$

  Since $T \in B'$ (since B' is a basis for V^*), we have that $T = a_1\phi_1 + \cdots + a_r\phi_r + \cdots + a_n\phi_n$

  Now we see $T(w_1) = (a_1\phi_1 + \cdots + a_n\phi_n)(w_1) = a_1 = 0$

  Similarly, we see $a_i = 0$ for $1 \leq i \leq r$

  Thus $T = a_{r+1}\phi_{r+1} + \cdots + a_n\phi_n \in \text{span}(\{\phi_{r+1}, \ldots, \phi_n\})$

Theorem 2: If $\langle , \rangle$ is non-degenerate, then for every subspace $W \subseteq V$, we have

$$V = W \oplus W^{\perp}$$

Now consider a $\langle , \rangle$ non-degenerate

Claim: $\forall v \in V$, given a linear transformation $L_v : V \to K$, let $L_v(w) = \langle v, w \rangle \in K$, then $F : V \to V^*$ where $F(v) = L_v$ is an isomorphism

5.6 Quadratic Forms

Definition - Symmetric Bilinear Form: Another way of calling scalar products on a vector space V

- **Symmetric** comes from $\langle v, w \rangle = \langle w, v \rangle$
- **Bilinear** comes from $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$ and $\langle v, cw \rangle = c\langle v, w \rangle = \langle cv, w \rangle$
- **Form** comes from the mapping $(v, w) \to \langle v, w \rangle$, often denoted as a function

$$g : V \times V \to K \qquad g(v, w) = \langle v, w \rangle$$

Definition - Quadratic Form: Given a scalar product $g = \langle , \rangle$, the **quadratic form** determined by g is a function

$$f : V \to K \qquad f(v) = g(v, v) = \langle v, v \rangle$$

Example: If $V = K^n$ then $f(X) = X \cdot X = x_1^2 + \cdots + x_n^2$ is the quadratic form determined by regular dot product

In general, if $V = K^n$ and C is a symmetric matrix, then the quadratic form is given by

$$F(X) =^t XCX = \sum_{i,j=1}^{n} c_{ij}x_ix_j$$

For a diagonal matrix C, this simplifies to

$$F(X) = c_1x_1^2 + \cdots + c_nx_n^2$$

5.7 Sylvester's Theorem

Let $V = R^2$ and let the form be represented by the symmetric matrix

$$C = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then the vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

form an orthogonal basis using $f(X) = \langle X, X \rangle =^t XCX$. Indeed

$$\langle v_1, v_1 \rangle = -1 \qquad \langle v_2, v_2 \rangle = 0$$

Now we generalize the situation above to arbitrary dimensions

Let $\{v_1, \ldots, v_n\}$ be an orthogonal basis of V and let

$$c_i = \langle v_i, v_i \rangle$$

After some renumbering of elements in our basis, we can assume that

$$\begin{aligned} c_1, \ldots, c_r &> 0 \\ c_{r+1}, \ldots, c_s &< 0 \\ c_{s+1}, \ldots, c_n &= 0 \end{aligned}$$

We are interested in looking at the number of positive, negative, and zero terms among $c_i = \langle v_i, v_i \rangle$ i.e. the numbers r and s

Let X be the coordinate vector of an element of V with respect to our basis and let f be the quadratic form associated with our scalar product. Then

$$F(X) = c_1x_1^2 + \cdots + c_rx_r^2 + \cdots + c_sx_s^2$$

Here we see r positive terms, $s - r$ negative terms, and that $n - s$ of the terms have disappeared

We can see this more clearly by normalizing the basis

Definition - Orthonormal: A basis $\{v_1, \ldots, v_n\}$ is **orthonormal** if for each i we have

$$\langle v_i, v_i \rangle = 1 \qquad \text{or} \qquad \langle v_i, v_i \rangle = -1 \qquad \text{or} \qquad \langle v_i, v_i \rangle = 0$$

If $\{v_1, \ldots, v_n\}$ is a orthogonal basis, we can always obtain an orthonormal basis by taking

- $c_i = 0 \implies v_i' = v_i$
- $c_i > 0 \implies v_i' = \dfrac{v_i}{\sqrt{c_i}}$
- $c_i < 0 \implies v_i' = \dfrac{v_i}{\sqrt{-c_i}}$

Then $\{v_1', \ldots, v_n'\}$ is an orthonormal basis

Now suppose that X is the coordinate vector of an element of V. In terms of the orthonormal basis, we have

$$f(X) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_s^2$$

Thus we can clearly see the number of positive and negative terms

We now show that number of positive, negative, and zero terms don't depend on the orthonormal basis

Theorem 8.1: Let V be a finite dimensional vector space over R with a scalar product. Take the subspace $V_0 \subseteq V, V_0 = \{v \in V \mid \forall w \in V, \langle v, w\rangle = 0\}$. Then the number of integers i such that $\langle v_i, v_i\rangle = 0$ is equal to the dimension of V_0

Proof: Suppose $\{v_1, \ldots, v_n\}$ is ordered such that

$$\langle v_1, v_1\rangle \neq 0, \ldots, \langle v_s, v_s\rangle \neq 0 \qquad \text{but} \langle v_i, v_i\rangle = 0 \qquad \text{for } i > s$$

Since $\{v_1, \ldots, v_n\}$ is orthogonal, clearly $v_{s+1}, \ldots, v_n \in V_0$

Now we take $v \in V_0$

$$v = x_1 v_1 + \cdots + x_s v_s + \cdots + x_n v_n$$

Taking the scalar product with any v_j for $j \leq s$, we get

$$0 \langle v, v_j\rangle = x_j \langle v_j, v_j\rangle \implies x_j = 0 \implies v \in \operatorname{span}(\{v_{s+1}, \ldots, v_n\})$$

Furthermore, since $\{v_{s+1}, \ldots, v_n\}$ is linearly independent, we have that $\{v_{s+1}, \ldots, v_n\}$ is a basis for V_0

Definition - Index of Nullity: From the proof above, we call V_0 the **index of nullity of the form**

- **Note**: Here form is non-degenerate if and only if the index of nullity $= 0$

Sylvester's Theorem: Let V be a finite dimensional vector space of R. Then there exists $r \geq 0$ such that if $\{v_1, \ldots, v_n\}$ is a basis, then there are precisely r integers such that

$$\langle v_i, v_i\rangle > 0$$

Proof Let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ be orthogonal bases for V. Arrange them such that

$$\begin{aligned}
\langle v_i, v_i\rangle &> 0 \qquad 1 \leq i \leq r \\
\langle v_i, v_i\rangle &< 0 \qquad r+1 \leq i \leq s \\
\langle v_i, v_i\rangle &= 0 \qquad s+1 \leq i \leq n \\
\langle w_i, w_i\rangle &> 0 \qquad 1 \leq i \leq r' \\
\langle w_i, w_i\rangle &< 0 \qquad r'+1 \leq i \leq s' \\
\langle w_i, w_i\rangle &= 0 \qquad s'+1 \leq i \leq n
\end{aligned}$$

We show that $v_1, \dots v_r, w_{r'+1}, \dots, w_n$ is linearly independent

Suppose that we have

$$x_1 v_1 + \cdots + x_r v_r + y_{r'+1} w_{r'+1} + \cdots + y_n w_n = 0 \implies x_1 v_1 + \cdots + x_r v_r = -(y_{r'+1} w_{r'+1} + \cdots + y_n w_n)$$

Let $c_i = \langle v_i, v_i \rangle$ and $d_i = \langle w_i, w_i \rangle$

Taking the scalar product of both sides with itself, we see that

$$c_1 x_1^2 + \cdots c_r x_r^2 = d_{r'+1} y_{r'+1}^2 + \cdots + d_{s'} y_{s'}^2$$

Clearly the LHS ≥ 0 and the RHS $\leq 0 \implies$ both sides are 0

Thus $x_1 = \cdots = x_r = 0 \implies y_{r'+1} = \cdots = y_n = 0$ by linear independence

Finally, since $\dim(V) = n$, we see that $r + n - r' \leq n \implies r \leq r'$

However, by symmetric we also get that $r' \leq r$

Thus we must have that $r = r'$

Definition - Index of Positivity: From Sylvester's Theorem, the integer r is called the **index of positivity**

5.8 Riesz Representation

Recall that $P_2(R) = \{a_0 + a_1 x + a_2 x^2 \mid a_i \in R\}$

Also recall that if $\langle , \rangle$ is non-degenerate, then $L^* : V \to V^*$ is an isomorphism where

$$L^*(v) = L_v : V \to K \qquad L_v(w) = \langle v, w \rangle$$

Riesz Representation Theorem: For any finite dimensional vector space V with a non-degenerate $\langle , \rangle$, for any linear function $\phi : V \to K \in V^*$, there exists a unique $u \in V$ such that $\phi = L_u$

Proof: Since $L^* : V \to V^*$ is an isomorphism, we let $u = (L^*)^{-1}(\phi)$

Proposition: There is a polynomial $u(x) \in P_2(R)$ such that for all $p(x) \in P_2(R)$

$$\int_0^1 p(x)u(x)\, dx = \int_\pi^\pi p(x)\cos(x)\, dx$$

Proof: Clearly $V = P_2(R)$ is finite dimensional and $\langle f, g \rangle = \displaystyle\int_0^1 fg$ is non-degenerate

Let

$$\phi : P_2(R) \to R \qquad \phi(p) = \int_\pi^\pi p(x)\cos(x)\, dx$$

Now we use Riesz Representation Theorem to get u such that

$$\int_0^1 p(x)u(x)\, dx = \langle u, p \rangle = \int_\pi^\pi p(x)\cos(x)\, dx$$

Proposition: There is a $u(x) \in P_2(R)$ such that for all $p(x) \in P_2(R)$ we have

$$\int_0^1 p(x)u(x)\,dx = P(0) = a_0$$

Proof: Let

$$\psi : P_2(R) \to R \qquad \psi(a_0 + a_1 x + a_2 x^2) = a_0$$

Then apply Riesz Representation Theorem

6 Operators

Definition - Operators: Linear transformations $T : V \to V$

Definition $\mathcal{L}(\mathbf{V}, \mathbf{V})$: Set of all linear transformations $T : V \to V$

- **Note**: $\mathcal{L}(\mathcal{V}, \mathcal{V})$ is a Vector Space

For the remainder of the course, we look at **operators** of V

For every linear transformation $T : V \to V$, we have an $n \times n$ matrix A

However, there are many different $n \times n$ matrices associated to the same transformation T

In fact, for any basis $B = \{v_1, \ldots, v_n\}$, we get a matrix $M_{n\times n}(T)_B^B$

In particular, we study properties of $n \times n$ matrices A that don't depend on the change of basis

6.1 Multilinear k-form

Defininition - Multilinear k-form: A function $\omega : \underbrace{V \times \cdots \times V}_{k \text{ factors}} \to K$ such that for all $1 \le i \le n$, for all $v_1, \ldots, v_i, w_i, v_{i+1}, \ldots, v_k$, and $a, b \in K$ we have

$$\omega(v_1, \ldots, v_{i-1}, (av_i + bw_i), v_{i+1}, \ldots, v_k) = a\omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_k) + b\omega(v_1, \ldots, v_{i-1}, w_i, v_{i+1}, \ldots, v_k)$$

Upshot: It's linear on each coordinate, provided that the other coordinates stay fixed

Let $\mathrm{ML}_k(V)$ be the set of all multilinear k-forms $\omega : V^k \to K$

- **Note**: $\mathrm{ML}_k(V)$ is a K-Vector Space

Consider: What is a multilinear 1-form

$\omega : V \to K$ is a linear transformation. Thus $\{\omega : V \to K\} = V^* =$ dual space

Consider: What is a multilinear 2-form (**bilinear form**)

$\omega : V \times V \to K$ is linear in each coordinate

$\mathrm{ML}_2(V)$ is the set of all bilinear forms on V

- **Note**: Scalar Products $\subseteq ML_2$

Definition - Alternating: A multilinear k-form $\omega : V^k \to K$ is **alternating** if some $v_i = v_j$ for $i \neq j$ then

$$\omega(v_1, \ldots, v_k) = 0$$

Example: $\begin{vmatrix} 5 & 0 & 0 \\ 4 & 3 & 3 \\ 2 & 6 & 6 \end{vmatrix} = 0$

Definition - $\mathbf{\Lambda(V)}$: All alternating multilinear k-forms

- **Note**: $\Lambda(V)$ is a subspace of $\mathrm{ML}_k(V)$
  - In particular $0 \in \Lambda(V) \subseteq \mathrm{ML}_k(V)$. This is the 0 mapping

Consider: For a fixed V with dimension n, what is $\Lambda(V)$?

Definition - Permutation: 1-1, onto mapping $\sigma : [n] \to [n]$

Example: For $n = 4$, $(1, 2, 3, 4) \to (2, 4, 1, 3)$ corresponds to $\sigma(1) = 2, \sigma(2) = 4, \sigma(3) = 1, \sigma(4) = 3$

We can also compose permutations

Let $\tau(1) = 1, \tau(2) = 3, \tau(3) = 3, \tau(4) = 4$. Then

- $\tau \circ \sigma(1) = 3$
- $\tau \circ \sigma(2) = 4$
- $\tau \circ \sigma(3) = 1$
- $\tau \circ \sigma(4) = 2$

Furthermore, every permutation $\sigma : [n] \to [n]$ has an inverse function σ^{-1}, satisfying $\sigma^{-1}\sigma = \mathrm{id}$

- $\sigma^{-1}(1) = 3$
- $\sigma^{-1}(2) = 1$
- $\sigma^{-1}(3) = 4$
- $\sigma^{-1}(4) = 2$

Definition - Transposition: A permutation τ that swaps two entries and fixes everything else

- **Note** For a transposition τ, we have that $\tau^{-1} = \tau \implies \tau^2 = \mathrm{id}$

Let S_n be the set of all permutations of $[n]$

Claim: S_n has $n!$ elements

Proof: on the homework

Claim: For all $n \geq 1$, every $\sigma \in S_n$ can be written as a (possibly empty) product of tranpositions

$$\sigma = \tau_r \circ \cdots \circ \tau_1$$

Proof by Induction:

Base Case: For $n = 1$, we have $S_1 \implies S_1 = \{\text{id}\}$ where id is the product of no transpositions

Base Case: For $n = 2$, we have $S_2 \implies S_2 = \{\text{id}, \tau_{1,2}\}$ where $\tau_{1,2}$ swaps $1, 2$

IH: Suppose for an arbitrary n, every $\sigma \in S_n$ can be written as a (possibly empty) product of transpositions

IS: Choose an arbitrary $\sigma \in S_{n+1}$

- Case 1: Suppose $\sigma(n+1) = n+1$. Then we can look at the remaining elements $[n]$, which by IH, any $\sigma \in S_n$ can be written as a product of transpositions
- Case 2: Suppose $\sigma(n+1) = j$ for some $j \leq n$. Then let τ be the transposition swapping $j, n+1$. Then $\tau \in S_{n+1}$ and $\tau\sigma(n+1) = n+1$

  By using Case 1, we can write

$$\tau\sigma = \tau_r \circ \cdots \circ \tau_1 \implies \tau\tau\sigma = \sigma = \tau(\tau_r \circ \cdots \circ \tau_1)$$

Definition - ϵ: Is a function $\epsilon : S_k \to \{-1, +1\}$

$$\epsilon(\sigma) = \begin{cases} +1 & \sigma \text{ is even} \\ -1 & \sigma \text{ is odd} \end{cases}$$

- **Note**: Any $\sigma \in S_k$ permutes $\{x_1, \ldots, x_k\} \to \{x_{\sigma(1)}, \ldots, x_{\sigma(k)}\}$

Notation: For each $\omega \in \Lambda_k(V)$ and each $\sigma \in S_k$, we let

$$(\sigma_\omega)(x_1, \ldots, x_k) = \omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$$

Example: For $k = 3$, suppose $\sigma(x_1, x_2, x_3) = (x_3, x_1, x_2)$

Then for any $(v_1, v_2, v_3) \in V^3$, we have

$$(\sigma_\omega)(v_1, v_2, v_3) = \omega(v_3, v_1, v_2)$$

Theorem: If $\omega \in \Lambda(V)$ and $\sigma \in S_k$, then

$$(\sigma_\omega) = \epsilon(\sigma)\omega$$

Meaning that for all $(v_1, \ldots, v_k) \in V^k$, we have

$$(\sigma_\omega)(v_1, \ldots, v_k) = \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \epsilon(\sigma)\omega(v_1, \ldots, v_k)$$

Proof: Since σ is a product of transpositions, it suffices to prove that when σ is a transposition τ swapping i, j

- **Note**: $\epsilon(\tau) = -1$

We need to show that for all $(v_1, \ldots, v_k) \in V^k$, we have

$$\omega(v_{\tau(1)}, \ldots, v_{\tau(k)}) = -\omega(v_1, \ldots, v_k)$$

Notationwise, let $\overline{\omega}(x, y)$ denote

$$\omega(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_{j-1}, y, v_{j+1}, v_k)$$

Note that $\overline{\omega}(x+y, x+y) = 0$ since ω is alternating

Thus we see that

$$\overline{\omega}(x+y, x+y) = \overline{\omega}(x, x) + \overline{\omega}(x, y) + \overline{\omega}(y, x) + \overline{\omega}(y, y) = 0$$

This shows that

$$\overline{\omega}(x, y) = \overline{\omega}(y, x) \implies \overline{\omega}(v_j, v_i) = -\overline{\omega}(v_i, v_j)$$

Theorem: Suppose $\{v_1, \ldots, v_k\} \subseteq V$ is linearly dependent. Then for all $\omega \in \Lambda_k(V)$, we have

$$\omega(v_1, \ldots, v_k) = 0$$

Proof: Suppose that v_i is a linear combination of the other vectors in the basis

$$v_i = \sum_{j \neq i} a_j v_j$$

Then we see that

$$\omega(v_1, \ldots, v_{i-1}, (\sum_{j \neq i} a_j v_j), v_{i+1}, \ldots, v_k) = \underbrace{\sum_{j \neq i} a_j \omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_k)}_{\text{by multilinearlity}} = 0$$

This last part follows since there are 2 v_j and ω is alternating

Upshot: Alternating multilinear k-forms preserve linear dependence

The Big Count: Suppose $\dim(V) = n$ and V has a basis $B = \{b_1, \ldots, b_n\}$. Take any $\omega \in \Lambda_k(V)$. Then for any $(v_1, \ldots, v_n) \in V^n$

$$\omega(v_1, \ldots, v_n) = (\sum_j a_{1j} b_j, \ldots, \sum_j a_{1n} b_j) = \underbrace{\sum_{1 \leq j_1, \ldots, j_n \leq n} a_{1j_1}, \ldots, a_{nj_n} \omega(b_{j1}, \ldots, b_{jn})}_{n^n \text{ terms}}$$

- **Note**: This follows from $v_i = \sum_j a_{1j} b_j$

However, the terms in the summation above are non-zero only when $j_1, \ldots, j_n$ are distinct

Thus the terms in the summation can be viewed as permutations $\sigma : \{1, \ldots, n\} \to \{j_1, \ldots, j_n\}$

Thus the summation actually only involves $n!$ terms

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \underbrace{\omega(b_{\sigma(1)}, \ldots, b_{\sigma(n)})}_{(\sigma_\omega)(b_1, \ldots b_n)}$$

Finally, we see that this is equal to

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \epsilon(\sigma) \omega(b_1, \ldots, b_n) = \omega(b_1, \ldots, b_n) \underbrace{\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \epsilon(\sigma)}_{\in K}$$

6.1.1 Consequences of the Big Count

Let $\{b_1, \dots, b_n\}$ be a basis of V

1. If $\omega(b_1, \dots, b_n) = 0 \implies \omega = 0$, the 0 mapping
2. If $\omega(b_1, \dots, b_n) \neq 0$ for some basis, then $\omega(c_1, \dots, c_n) \neq 0$ for any other basis of V, $\{c_1, \dots, c_n\}$
3. If $w, w' \neq 0$ are 2 different elements in $\Lambda_n(V)$,then they are linearly dependent
   - This means that $w' = cw$ for some $c \in K \implies \dim(\Lambda_n(V)) \leq 1$

Theorem: If $\dim(V) = n \geq 1$, then $\dim(\Lambda_n(V)) = 1$

- This means that there is some non-zero $\omega \in \Lambda_n(V)$

Proof by Induction on $k \leq n$

We will show that there is some non-zero $\omega \in \Lambda_n(V)$

Base case $k = 1$. Recall that $\mathrm{ML}_1(V) = V^*$, which has dimension ≥ 1

IH: Assume there exists a non-zero $\omega \in \Lambda_k(V)$ with $k < n$

IS: Show that there is a $\hat{\omega} \in \Lambda_{k+1}(V)$ where $\hat{\omega} \neq 0$

TODO FINISH THIS PROOF

Now take a linear transformation $T : V \to V$ that induces another linear transformation $T^* : \Lambda_n(V) \to \Lambda_n(V)$ defined by

$$T^*(\omega) : V^n \to K \qquad T^*(\omega)(u_1, \dots, u_n) = \omega(T(u_1), \dots, T(u_n))$$

Clearly, $T^* : \Lambda_n(V) \to \Lambda_n(V)$ is just scalar multiplication, meaning that there is some $d \in K$ such that

$$\forall \omega \in \Lambda_n(V), T^*(\omega) = d\omega$$

Definition - Determinant: The **determinate** of T is exactly the d above. That is $\det(T) = d \in K$

Properties of $\det(\mathbf{T})$:

1. Suppose $T : V \to V$ is multiplication by a. That is $T(v) = av$

   Then $T^* : \Lambda_n(V) \to \Lambda_n(V) \qquad T^*(\omega)(u_1, \dots, u_n) = \omega(T(u_1), \dots, T(u_n)) = \omega(au_1, \dots, au_n) = a^n\omega(u_1, \dots, u_n)$

   Thus $T^*(\omega) = a^n\omega$ for $\omega \in \Lambda_n(V)$

   Here $\det(T) = a^n$

   Special Cases:

   - $\mathrm{id} : V \to V \qquad \forall v \in V, \mathrm{id}(v) = v \implies \det(\mathrm{id}) = 1$
   - zero: $V \to V \qquad \forall v \in V, \mathrm{zero}(v) = 0 \implies \det(\mathrm{zero}) = 0$
2. Take two linear transformations $S, T : V \to V$

   Then the composition $S \circ T : V \to V$ is also a linear transformation

   We claim that $\det(S \circ T) = \det(S)\det(T)$

   For any $\omega \in \Lambda_n(V)$, we have that

$$
\begin{aligned}
(S\circ T)^*(\omega)(u_1,\dots,u_n) &= \omega(S\circ T(u_1),\dots,S\circ T(u_n)) \\
&= \det(S)\omega(T(u_1),\dots,T(u_n)) \\
&= \det(S)\det(T)\omega(u_1,\dots,u_n) \\
\implies (S\circ T)^*(\omega) &= \det(T)\det(S)\omega
\end{aligned}
$$

Special Cases:

- Suppose that $T : V \to v$ is invertible, then clearly $T^{-1}\circ T = \mathrm{id} \implies \det(T^{-1}\circ T) = \det(\mathrm{id}) = 1$

  Thus $\det(T^{-1})\det(T) = 1 \implies \det(T^{-1}) = \dfrac{1}{\det(T)}$

  Thus T is invertible if and only if $\det(T) \neq 0$

TFAE Theorem:

1. T is an isomorphism
2. T is invertible
3. $\mathrm{rank}(T) = n$
4. $\det(T) \neq 0$

Proof: $1 \iff 2 \iff 3$ is shown by the previous proof

To show that 4 must hold, by the special case before, if any of $1, 2, 3$ hold, then $\det(T) \neq 0$

Now we show that if $1, 2, 3$ fail, then $\det(T) = 0$

Since 3 fails, we must have that $\mathrm{rank}(T) = \dim(\mathrm{Im}(T)) < \dim(V) = n$

Now choose any $\omega \in \Lambda_n(V)$ and choose any $(u_1,\dots,u_n) \in V^n$

We see that

$$(T^*)(\omega)(u_1,\dots,u_n) = \omega(T(u_1),\dots,T(u_n))$$

Since $\{T(u_1,\dots,T(u_n)\}$ are n vectors in $\Im(T)$, they must be linearly dependent

Thus since ω respect linearly dependency, we see that

$$\omega(T(u_1),\dots,T(u_n)) = 0$$

Thus for any $\omega \in \Lambda_n(V)$, we must have

$$T^*(\omega) = 0 \implies \det(T) = 0$$

6.1.2 Matrix Representation

Now take $A \in M_{n\times n}(K)$

We know that A encodes a linear transformation, namely $T_A : K^n \to K^n$

Thus $\det(A) = \det(T_A)$

Consequences:

1. $\det(I_n) = 1$ since $T_{I_n} = \mathrm{id} : K^n \to K^n$ and $\det(\mathrm{id}) = 1$
2. $\det(O) = 0$ since $T_{\text{zero}} = \text{zero} : K^n \to K^n$ and $\det(\text{zero}) = 0$

3. For $A, B \in M_{n\times n}(K), \det(AB) = \det(A)\det(B)$

   The linear transformation T_{AB} is described by the composition $T_A \circ T_B \implies \det(T_{AB}) = \det(T_A)\det(T_B) = \det(A)\det(B)$

TFAE Theorem: For $A \in M_{n\times n}(K)$, the following are equivalent

1. T_A is an isomorphism
2. A is invertible
3. $\text{rank}(A) = n$
4. $\det(A) \neq 0$

Now we can compute $\det(A)$ for $A \in M_{n\times n}(K)$ by applying The Count Theorem

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

Example: For $n = 2$ we have $S_2 = \{\text{id}, \tau\}$ where $\epsilon(\text{id}) = 1$ and $\epsilon(\tau) = -1$

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \epsilon(\text{id}) a_{11}a_{22} - \epsilon(\tau) a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}$$

Note: Since any linear transformation can be represented as a matrix A, we have that

$$\det(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \det({}^tA)$$

7 Determinants

Determinants only make sense for square $n \times n$ matrices. We define the **determinate** as

- $1 \times 1 \implies \det(a) = a$
- $2 \times 2 \implies \det : M_{2\times 2}(K) \to K$ where $\det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ad - bc$
- $3 \times 3 \implies \det : M_{3\times 3}(K) \to K$ where $\det(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}) = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$

Example: $\begin{vmatrix} 2 & 1t \\ 3 & 5t \end{vmatrix} = 2(5t) - 3(t) = 10t - 3t = t\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix}$

Example: $\begin{vmatrix} a + a' & b \\ c + c' & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b \\ c' & d \end{vmatrix}$

- **Upshot**: Freezing a column gives us linearity with the other column

Example: $\begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -1\begin{vmatrix} a & b \\ c & d \end{vmatrix}$

- **Upshot**: Switching columns changes the sign of the determinant

Example: $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$

Example: $\begin{vmatrix} 5 & 1 & 2 \\ 3 & 2 & 0 \\ 4 & 1 & 3 \end{vmatrix} + \begin{vmatrix} 5 & 2 & 2 \\ 3 & -1 & 0 \\ 4 & 0 & 3 \end{vmatrix} = 11 - 25 = -14 = \begin{vmatrix} 5 & 3 & 2 \\ 3 & 1 & 0 \\ 4 & 1 & 3 \end{vmatrix}$

7.1 Row Determinants

We look at what exactly are row reductions and their impact on determinants

Suppose $A^1, \dots, A^n$ are columns of A. Let B have the same columns, except two swapped columns

From the rules of determinants, we have that

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

Now consider replacing a column by itself plus some scalar multiple of another column

That is $B = [A^1 + cA^2, A^2, \dots]$. Then we see that

$$\det(B) = \det(A^1, A^2, \dots) + c\det(A^2, A^2, A^3, \dots) = \det(A)$$

Finally, since $\det({}^tA) = \det(A)$, these equalities work under row operations as well

8 Symmetric, Hermitian, Unitary Operators

Definition - Operator: A linear transformation $T : V \to V$

Consider when V is a K-Vector Space and $\langle , \rangle$ is a positive definite scalar product

Recall that

- $\|v\| = \sqrt{\langle v, v \rangle}$
- Gram-Schmidt process takes a basis B and produces an orthonormal basis
- V^* is the set of linear transformation $\phi : V \to R$ and that $V \approx V^*$ under

$$L^* : V \to V^* \qquad L^*(w) : V \to R \qquad L^*(w)(v) = \langle v, w \rangle \forall v \in V$$

Fundamental Fact

For any operator $A : V \to V$, there exists a unique operator $B : V \to V$ such that for all $v, w \in V$

$$\langle Av, w \rangle = \langle v, Bw \rangle$$

Here B is called the **transpose** of A, namely $B =^t A$

Now given A, how do we find B?

Take $w \in V$ and let $L_w^A : V \to R$ be defined by $L_w^A(v) = \langle Av, w \rangle$

- It can be shown that L_w^A is a linear transformation. Thus $L_w^A \in L^*$

Furthermore, since $L^* : V \to V^*$ is an isomorphism, there exists a unique $w' \in V$ such that

$$L^*(w') = L_w^A$$

- Importantly, $L^*(w')$ is the same function as L_w^A

But then we have that

$$\forall v \in V \qquad L_w^A(v) = L^*(w')(v) \implies \langle Av, w\rangle = \langle v, w'\rangle$$

Now we define $B : V \to V$ such that $B(w) = w'$. Thus we have

- **Note**: It can be shown that B is a linear transformation

$$\langle Av, w\rangle = \langle v, Bw\rangle$$

Furthermore we have that

$$\langle Av, w\rangle = \langle v,^t Aw\rangle$$

Definition - Symmetric: An operator $A : V \to V$ is **symmetric** if and only if any $n \times n$ matrix representing A is a symmetric matrix

$$^tA = A$$

Definition - Unitary: An operator $A : V \to V$ is **unitary** if

$$\forall v, w \in V \qquad \langle Av, Aw\rangle = \langle v, w\rangle$$

- **Note**: We say A is **norm-preserving** if for all $v \in V$, $\langle Av\rangle = \langle v\rangle$

Proposition: A Is unitary if and only if A is norm-preserving

Proof: $\implies$ Assume that A is unitary and choose $v \in V$. Clearly

$$\|Av\|^2 = \langle Av, Av\rangle = \langle v, v\rangle = \|v\|^2$$

$\impliedby$ Assume A is norm-preserving and chooose $v, w \in V$. Then

$$\langle v + w, v + w\rangle - \langle v - w, v - w\rangle = 4\langle v, w\rangle$$

Similarly, we have that

$$\langle A(v + w), A(v + w)\rangle - \langle A(v - w), A(v - w)\rangle = 4\langle Av, Aw\rangle$$

Thus we have that $\|v + w\|^2 - \|v - w\|^2 = \|A(v + w)\|^2 - \|A(v - w)\|^2 \implies \langle v, w\rangle = \langle Av, Aw\rangle$

8.1 Symmetric Operators

Here we look at any vector space K^n and non-degenerate scalar products

Lemma: For $A : V \to V$, there exists a unique $B : V \to V$ such that for all $u, w \in V$

$$\langle Av, w\rangle = \langle v, Bw\rangle$$

Proof: For $w \in V$, we define $L : V \to K$ as $L(v) = \langle Av, w\rangle$

- Clearly L is linear and thus $L \in V^*$

- Furthermore, we know that there exists a unique $w' \in V$ such that for all $v \in V$,

$$L(v) = \langle Av, w \rangle = \langle v, w' \rangle = \langle v, Bw \rangle$$

Now define $B : V \to V$ as $B(w) = Bw = w'$

Now we show that B is linear

- $$\begin{aligned} \langle v, B(w_1 + w_2) \rangle &= \langle Av, w_1 + w_2 \rangle \\ &= \langle Av, w_1 \rangle + \langle Av, w_2 \rangle \\ &= \langle v, Bw_1 \rangle + \langle v, Bw_2 \rangle \\ &= \langle v, Bw_1 + Bw_2 \rangle \end{aligned}$$

- $$\begin{aligned} \langle v, B(cw) \rangle &= \langle Av, cw \rangle \\ &= c\langle Av, w \rangle \\ &= c\langle v, Bw \rangle \\ &= \langle v, cBw \rangle \end{aligned}$$

Here $B = A^t$ is called the **transpose** of A

Definition - Symmetric: An operator A is **symmetric** if $A^t = A$

- Thus for any operator of V

$$\langle Av, w \rangle = \langle v, A^t w \rangle$$

Upshot: A is symmetric if and only if $\langle Av, w \rangle = \langle v, Aw \rangle$ for all $v, w \in V$

Example: Consider the typical dot product in $V = K^n$ and consider a matrix A and column vectors X, Y

$$\langle X, Y \rangle = X^t Y \implies \langle AX, Y \rangle = (AX)^t Y = X^t A^t Y = \langle X, A^t Y \rangle$$

Theorem: For any operators, we have that

1. $(A + B)^t = A^t + B^t$
2. $(AB)^t = B^t A^t$
3. $(cA)^t = cA^t$
4. $A^{tt} = A$

Proof of 2:

$$\langle ABv, w \rangle = \langle Bv, A^t w \rangle = \langle v, B^t A^t w \rangle \implies (AB)^t = B^t A^t$$

8.2 Hermitian Operators

Here we consider a vector space over C and a Hermitian inner product

Recall that a Hermitian inner product is one such that

$$\langle X, Y\rangle = X^t\overline{Y}$$

Theorem: For any vector space V over C with a Hermitian inner product, given a function L, there exists a unique $w' \in V$ such that

$$L(v) = \langle v, w'\rangle \qquad \forall v \in V$$

Now we denote the mapping $w \to w'$ by A^*

Lemma: For $A : V \to V$ there exists a unique operator $A^* : V \to V$ such that

$$\langle Av, w\rangle = \langle v, A^*w\rangle \qquad \forall v, w \in V$$

Definition - Adjoint: Here A^* is called the **adjoint** of A

Example: Consider

$$\begin{aligned} \langle X, Y\rangle = X^t\overline{Y} \implies \langle AX, Y\rangle = (AX)^t\overline{Y} = X^tA^t\overline{Y} = X^t\overline{(\overline{A^t}Y)} \\ \langle X, A^*Y\rangle = X^t\overline{(A^*Y)} \end{aligned}$$

Thus we see that $A^* = \overline{A^t}$

Definition - Hermitian: An operator is **Hermitian** if $A^* = \overline{A^t} = A$, or $A^t = \overline{A}$

Theorem:

1. $(A + B)^* = A^* + B^*$
2. $(AB)^* = B^*A^*$
3. $(\alpha A)^* = \overline{\alpha}A^*$
4. $A^{**} = A$

Proof 3:

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

Thus $(\alpha A)^* = \overline{\alpha}A^*$

8.3 7.3 Unitary Operators

Consider a vector space over R and a positive definite scalar product

Definition - Real Unitary: An operator A is **real unitary** if

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

Theorem: For a vector space V over R and an operator $A : V \to V$, the following are equivalent

1. A is unitary
2. A preserves the norm $\|Av\| = \|v\|$
3. For every unit vector $v \in V$, Av is also a unit vector

Proof $1 \to 2$: Since $\langle Av, Av \rangle = \langle v, v \rangle$, which means that their norms are also equivalent

Proof $2 \to 1$: We see that

$$\langle A(v+w), A(v+w) \rangle - \langle A(v-w), A(v-w) \rangle = 4\langle Av, Aw \rangle = 4\langle v, w \rangle = \langle v+w, v+w \rangle - \langle v-w, v-w \rangle$$

Proof $2 \leftrightarrow 3$: Follows since $\|Av\| = \|v\|$

Upshot: Unitary mappings map unit vectors to unit vectors

- Also preserves perpendicularity

$$\langle Uv, Uw \rangle = \langle v, w \rangle = 0$$

Theorem: For a vector space over R with a positive definite sclar product, a linear map $A : V \to V$ is unitary if and only if

$$A^t A = I$$

Proof: An operator A is unitary if and only if

$$\langle Av, Aw \rangle = \langle v, w \rangle \implies \langle A^t Av, w \rangle = \langle v, w \rangle$$

Thus this equivalency holds when $A^t A = I$

Note: A unitary matrix is invertible since $Av = O \implies v = O$ since A preserve the norm

Under $V = R^n$ and the usual dot product, a real matrix is **real unitary** if

$$A^t A = I_n \iff A^t = A^{-1}$$

Example: The only unitary maps on R^2 to itself are rotations

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

Under $V = C^n$ and the usual Hermitian product, A is **complex unitary** if

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

This holds if and only if

$$A^* A = I \iff \overline{A^t} A = I \iff \overline{A^t} = A^{-1}$$

Theorem: Let V be a vector space either over R with a positive definite scalar product, or over C with a positive definite hermitian product. For a linear map $A : V \to V$ and an orthonormal basis $\{v_1, \ldots, v_n\}$, we have that

- A is unitary $\implies \{Av_1, \ldots, Av_n\}$ is an orthonormal basis
- If $\{w_1, \ldots, w_n\}$ is another orthonormal basis such that $Av_i = w_i$, then A is unitary

9 Eigenvectors and Eigenvalues

For the rest of the course, we consider a Finite Dimensional Vector Space V over R or C with a positive definite $\langle , \rangle$ that is

- Scalar if we are dealing with R
- Hermitian if we are dealing with C

We study operators $A : V \to V$

Some operators are easy to understand. For example consider $A : R^2 \to R^2$ $\quad A = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$

Then we have $A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ and $A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

Thus A stretches x by 5 and streches y by 3 $\implies A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5x \\ 3y \end{bmatrix}$

However, some operators are more difficult to understand. For example, consider $B : R^2 \to R^2$ $\quad B = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

Then we have $B \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $B \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Clearly using the standard basis, it's difficult to describe the operator

However, if we instead use the basis $B = \{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\}$, we get

$$B \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad B \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Which is much easier to describe the operator

Definition - Eigenvector: Nonzero $v \in V$ such that $Av = \lambda v$ for some $\lambda \in K$

Definition - Eigenvalue: Some $\lambda \in K$ such that $Av = \lambda v$ for some nonzero $v \in V$

So for the examples above

- Eigenvalues for A are $\lambda \in \{5, 3\}$ corresponding to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- Eigenvalues for B are $\lambda \in \{4, 2\}$ corresponding to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Definition - Eigenspace: For any λ, the **eigenspace** is

$$W_\lambda = \{v \in V \mid Av = \lambda v\}$$

- **Note**: W_λ is a subspace of V

Theorem: Every non-zero element of W_λ is an eigenvector of A having λ as its eigenvalue

Proof: Let $v_1, v_2 \in V$ such that $Av_1 = \lambda v_1$ and $Av_2 = \lambda v_2$. Then

$$A(v_1 + v_2) = Av_1 + Av_2 = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$$

If $c \in K$ then

$$A(cv_1) = cAv_1 = c\lambda v_1 = \lambda c v_1$$

Thus any linear combination of elements of W_λ has the same eigenvalue

Example: Rotations. Fix an angle θ

Define the operator $R_\theta = R^2 \to R^2$ (rotation counter-clockwise) by

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

- **Note**: There are no eigenvalues in R and thus no eigenvectors
- **Note**: However, for $v \in C^2, R_\theta(v) = e^{i\theta} v$

Theorem: Let $v_1, \ldots, v_m$ be eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_m$. Also assume that these eigenvalues are distinct, then $v_1, \ldots, v_m$ are linearly independent

Proof by Induction:

Base Case: $m = 1$ then clearly $v_1 \in V$ is linearly independent

IH: Assume statement holds for $m - 1$

IS: Show for m. Suppose we have the relation

$$c_1 v_1 + \cdots + c_m v_m = O$$

Then we multiply the equation above by λ_1 and by A to get

$$\begin{aligned} c_1\lambda_1 v_1 + \cdots + c_m\lambda_1 v_m &= O \\ c_1\lambda_1 v_1 + \cdots + c_m\lambda_m v_m &= O \end{aligned}$$

Combining the equations, we get

$$c_2(\lambda_2 - \lambda_1)v_2 + \cdots + c_m(\lambda_m - \lambda_1)v_m = O$$

Since λ_i, λ_j are distinct, we conclude (by IH) that

$$c_2 = \cdots = c_m = 0$$

Finally looking back at the original equation, we get $c_1 v_1 = O \implies c_1 = 0$

Definition - Diagonalization: Suppose we have a linear map $L : V \to V$ and a basis of V $\{v_1, \ldots, v_n\}$. We say this basis **diagonalizes** L if each v_i is an eigenvector of L such that $Lv_i = c_i v_i$

Then the matrix representing L with respect to this basis is a diagonal matrix

$$A = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}$$

We say that a linear map L can be **diagonalized** if there exists a basis of V consisting of eigenvectors

9.1 Characteristic Polynomial

Now show how to determine if λ is an eigenvalue

By definition, λ is an eigenvalue if and only if there is some nonzero $v \in V$ such that $Av = \lambda v$

But this means that $(\lambda I - A)v = O$

Which means that $v \in \text{Ker}(\lambda I - A)$

Which means that $\text{Ker}(\lambda I - A) \neq \{O\}$

Which means that $\det(\lambda I - A) = 0$

Definition - Characteristic Polynomial: For an operator A, we define the **characteristic polynomial**

$$P_A(t) = \det(tI - A)$$

- **Note**: λ is an eigenvalue of A if and only if $P_A(\lambda) = 0$

Example: Consider $A : R^2 \to R^2$ $\quad A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. Determine the eigenvalues of A

We look at

$$\det(\lambda I - A) = \lambda^2 - 4\lambda + 3 - 8 = 0 \implies t \in \{5, -1\}$$

Thus eigenvalues of A are $\{5, -1\}$

Note: By the Fundamental Theorem of Algebra, every non-constant monic polynomial has a root in C. Thus every $A : V \to V$ has an eigenvalue (hence eigenvector) in C

9.2 Diagonalization of Symmetric Linear Maps

In this section, we let V be a vector space over R with a positive definite scalar product

Recall that A is symmetric if we have the relation

$$\langle Av, w \rangle = \langle v, Aw \rangle$$

Theorem: Let $A : V \to V$ be a symmetric matrix, then A has a nonzero eigenvector

Proof: If V is a vector space under C, we know that it has a nonzero eigenvector

Now consider V as a vector space under R with eigenvalue $\lambda \in C$ and eigenvector $Z = X + iY \in C$. We show that X, Y are real eigenvectors of A with eigenvalue λ

Let $Z = (z_1, \dots, z_n)^t$ have complex coordinates z_i. Then

$$Z \cdot \overline{Z} = \overline{Z} \cdot Z = \overline{Z}^t Z = |z_1|^2 + \cdots + |z_n|^2 > 0$$

By hypothesis, we have that $AZ = \lambda Z$. Thus

$$\overline{Z}^t AZ = \overline{Z}^t \lambda Z = \lambda \overline{Z}^t Z$$

Since the transpose of a 1×1 matrix equals itself, we also have

$$Z^t A^t \overline{Z} = \overline{Z}^t AZ = \lambda \overline{Z}^t Z$$

However since $\overline{AZ} = A\overline{Z}$ (since A is real) and $\overline{AZ} = \overline{\lambda Z}$. Thus we have

$$\lambda \overline{Z}^t Z = \overline{\lambda} Z^t \overline{Z}$$

Since $Z^t \overline{Z} \neq 0$, it follows that $\lambda = \overline{\lambda}$ and thus λ is real

Now from $AZ = \lambda Z$, we see that

$$Ax + iAY = \lambda X + i\lambda Y$$

Since A, X, Y are real and $AX = \lambda X$ and $AY = \lambda Y$, this proves the theorem

Example: Consider

$$A = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$$

Then for $X^t = (x, y)$, we see that

$$X^t AX = 3x^2 - 2xy + 2y^2$$

Theorem: If V is a finite dimensional vector space with a positive definite scalar product, with $A : V \to V$ a symmetric linear map, then A has a nonzero eigenvector

Proof: Follows from the previous theorem

Definition - Invariant: For a subspace $W \subseteq V$ and symmetric linear map $A : V \to V$, we say that W is **invariant** under A if $A(W) \subseteq W$, that is $\forall u \in W, Au \in W$

Theorem: Let $A : V \to V$ be a symmetric linear map and let v be a non-zero eigenvector of A

- If $w \in V$ and $w \perp v$, then $Aw \perp v$
- If W is a subspace of V and is invariant under A, then $W^\perp$ is also invariant under A

Proof: Suppose that v is an eigenvector of A, then

$$\langle Aw, v\rangle = \langle w, Av\rangle = \langle w, \lambda v\rangle = \lambda\langle w, v\rangle = 0$$

Thus $Aw \perp v$

Now suppose that W is invariant under A and take $u \in W^{\perp}$. Then for all $w \in W$ we have that

$$\langle Au, w\rangle = \langle u, Aw\rangle = 0$$

Thus $Au \in W^{\perp} \implies W^{\perp}$ is invariant under A

Spectral Theorem: For a symmetric linear map $A : V \to V$, V has an orthonormal basis consisting of eigenvectors

Proof: Since A is a symmetric, there exists a non-zero eigenvector v for A

Let W be the one-dimensional space generated by v. Then W is invariant under A

Thus by the theorem above, $W^{\perp}$ is also invariant under A and is a vector space of dimension $n-1$

By applying induction, we can find a basis $\{v_2, \dots, v_n\}$ of $W \perp$ consisting of eigenvectors

Finally we have an orthogonal basis of V consisting of eigenvectors $\{v_1, \dots, v_n\}$

Taking the norm of each vector gives us an orthonormal basis

Finally, let $\{e_1, \dots, e_n\}$ be an orthonormal basis of V such that each e_i is an eigenvector, then the matrix of A with respect to this basis is diagonal and consists of the eigenvalues

$$A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Definition - Spectral Basis: A basis $\{v_1, \dots, v_n\}$ such that each v_i is an eigenvector for A is called a **spectral basis** of A

- This basis **diagonalizes** A since the matrix of A with respect to this basis is a diagonal basis

10 Polynomials and Matrices

10.1 Polynomials

Definition - Polynomial: A **polynomial** over a field K is an expression

$$f(t) = a_n t^n + \cdots + a_0$$

Where t is a variable

We also define operations on polynomials

- Addition

$$(f+g)(t) = (a_n + b_n)t^n + \cdots + (a_0 + b_0)$$

- Scalar Multiplication

$$(cf)(t) = ca_n t^n + \cdots + ca_0$$

- Multiplication

$$(fg)(t) = c_{n+m}t^{n+m} + \cdots + c_0 \qquad c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$$

Theorem: For polynomials f, g

$$\deg(fg) = \deg(f) + \deg(g)$$

- **Note**: We let $f = 0$ have degree $-\infty$

Proof:

$(fg)(t) = a_n b_m t^{n+m} + \cdots$ and $a_n b_m \neq 0$. Thus $\deg(fg) = n + m = \deg(f) + \deg(g)$

If either f, g is 0, then our convention of having the 0 function have degree $-\infty$ works out

Definition - Root: A **root** α of f is a number such that $f(\alpha) = 0$

Theorem: Let f be a polynomial with complex coefficients, leading coefficient 1, and $\deg(f) = n \geq 1$, then there exist complex numbers $\alpha_1, \ldots, \alpha_n$ such that

$$f(t) = (t - \alpha_1) \cdots (t - \alpha_n)$$

10.2 Polynomials of Matrices and Linear Maps

We denote the set of polynomials with coefficients in K as $K[t]$

Now for a polynomial $f \in K[t]$ and a square matrix A with coefficients in K, we can write

$$f(A) = a_n A^n + \cdots + a_0 I$$

Theorem: For $f, g \in K[t]$ and a square matrix, we have

$$\begin{aligned} (f+g)(A) &= f(A) + g(A) \\ (fg)(A) &= f(A)g(A) \\ (cf)(A) &= cf(A) \qquad c \in K \end{aligned}$$

Theorem: For a $n \times n$ matrix A, there exists a non-zero polynomial $f \in K[t]$ such that $f(A) = O$

Proof: Since the vector space of $n \times n$ matrices has dimension n^2, successive powers of A are dependent independent for some $N > n^2$

Thus there exists numbers $a_0, \ldots, a_N \in K$, not all 0, such that

$$a_N A^N + \cdots + a_0 I = O$$

Setting, $f(t) = a_N t^N + \cdots + a_0$ gets us the desired polynomial

Note: The theorem above also applies for any linear map A of a finite dimensional vector space

11 Triangulation of Matrices and Linear Maps

11.1 Existence of Triangulation

Definition - Fan: A **Fan** of an operator A is a sequence of subspaces $\{V_1, \dots, V_n\}$ such that each $V_i \subset V_{i+1}$, where each $\dim(V_i) = i$ and each V_i is A-invariant

Definition - Fan Basis: For a fan of A, $\{V_1, \dots, V_n\}$, a **fan basis** is a basis $\{v_1, \dots, v_n\}$ such that $\{v_1, \dots, v_i\}$ is a basis for V_i

- **Note**: A fan basis automatically exists since we can inductively construct a basis for V_i; namely, let v_1 be a basis for V_1. We can extend this to $\{v_1, v_2\}$, a basis of V_2, and so on until we have a basis $\{v_1, \dots, v_n\}$ of V_n

Theorem: Let $\{v_1, \dots, v_n\}$ be a fan basis for A, then a matrix associated with A relative to this basis is an upper triangular matrix

Proof: Since AV_i is contained in V_i, there exists numbers a_{ij} such that

$$\begin{aligned} Av_1 &= a_{11}v_1 \\ Av_2 &= a_{12}v_1 + a_{22}v_2 \\ &\vdots \\ Av_n &= a_{1n}v_1 + \cdots + a_{nn}v_n \end{aligned}$$

Thus the matrix associated with A with respect to our basis is a triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Note: Let A be an upper triangular matrix as above, we can view A as a linear map $K^n \to K^n$. Thus the column unit vectors $e^1, \dots, e^n$ form a fan basis for A

Indeed, if we let V_i be the space generated by $e^1, \dots, e^i$, then $\{V_1, \dots, V_n\}$ is the corresponding fan

Thus the converse of the above theorem is true

Definition - Triangulable: For an operator $A : V \to V$, if there exists a basis for V for which the associated matrix of A is triangular, the A is **triangulable**

Theorem: For a finite dimensional vector space V over the complex numbers, let $A : V \to V$ be a linear map. Then there exists a fan of A in V

Proof by Induction: If $\dim(V) = 1$, then there is nothing to prove

IH: Assume the statement holds for $\dim(V) = n - 1$

IS: Show the statement holds for $\dim(V) = n$

Since V is over the complex numbers, it has a non-zero eigenvector v_1 for A. Let V_1 be the subspace generated by v_1

Thus for some subspace W, we can write

$$V = V_1 \oplus W$$

The only issue is that W isn't A-invariant

Define P_1 to be the projection of V onto V_1 and P_2 to be the project of V onto W

Then clearly P_2A is a linear map $V \to V$ that maps W into W (since P_2 maps any element of V into W)

Thus by IH, P_2A has a fan, call it $\{W_1, \ldots, W_{n-1}\}$

Now we let

$$V_i = V_1 + W_{i-1} \qquad i \geq 2$$

Then $V_i \subseteq V_{i+1}$ and $\dim(V_i) = i$

Now to show that $\{V_1, \ldots, V_n\}$ is a fan for A, we show that AV_i is contained in V_i

$$A = IA = (P_1 + P_2)A = P_1A + P_2A$$

Now take $v \in V_i$, we can write $v = cv_1 + w_{i-1}$ and we see that $P_1(Av) \in V_1$ and $P_2Av = P_2A(cv_1) + P_2Aw_{i-1} = cP_2(c\lambda_1 v_1) + P_2AW_{i-1} = P_2AW_{i-1} \in W_{i-1} \subset V_i$

Corollary: For a vector space V over the complex numbers and an operator $A : V \to V$, there exists a basis of V such that the matrix of A with respect to this basis is triangular

Corollary: For a matrix M of complex numbers, there exists a non-singular matrix B such that $B^{-1}MB$ is triangular

11.2 Theorem of Hamilton-Cayley

Suppose we have eigenvectors $\{v_1, \ldots, v_n\}$ and eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. Consider the characteristic polynomial of A

$$P(A) = (A - \lambda_1 I) \cdots (A - \lambda_n I)$$

If we apply $P(A)$ to any eigenvector v_i, then the factor $A - \lambda_i I$ will result in 0 and thus

$$P(A)v_i = O \implies P(A) = O$$

Theorem: For a vector space over the complex numbers and an operator $A : V \to V$, then for the characteristic polynomial P, we have

$$P(A) = O$$