Lecture 5:

Zorn's Lemma

Let S be a partially ordered set. If every chain of S has an upper bound in S, then S contains a maximal elements.

Definition 1: (Partially ordered) A partially ordered on a (non-empty) set S is a binary relation on S, denoted $\leq$, which satisfies:

- for $\forall s \in S$, $s \leq s$
- if $s \leq s'$ and $s' \leq s$, then $s = s'$
- if $s \leq s'$ and $s' \leq s''$, then $s \leq s''$.

Definition 2: If every elements in a partially ordered set S is comparable under $\leq$, then S is called a totally ordered set.

Definition 3: A chain is a collection of elements in S satisfying:
if $s_1 \in S$ and $s_2 \in S$, then either $s_1 \leq s_2$ or $s_2 \leq s_1$.

For simplicity, we may consider countable chain $s_1 \leq s_2 \leq \ldots$ in our discussion
(not necessarily countable) for the ease of explanation

Definition 4: A maximal element m of a partially ordered set S is defined as follows: for $\forall s \in S$ to which m is comparable, $s \leq m$.

Definition 5 An upper bound $\tilde{C}$ of a chain $\{C_i\}_{i \in I}$ is that for all i, ^I
$C_i \leq \tilde{C}$.

Zorn's lemma (adapted to 2048)

Let V be a vector space. Let S be the collection of linearly independent subsets of V. Then S is partially ordered under $\subseteq$.

Assume every chain $\{L_\alpha\}_{\alpha \in I}$ of S has a upper bound.

(Obviously, $\bigcup_\alpha L_\alpha$ is an upper bound of the chain. We need to show that $\bigcup L_\alpha$ is linearly independent)

Then: S has a maximal element.

Theorem: Every vector space has a basis.

Proof: Let $\mathcal{C}$ be the collection of all linearly independent subsets of V.

For any chain $\{S_i\}_{i \in I}$ (we may consider a countable chain $S_1 \subset S_2 \subset \cdots$ for easier interpretation.)

$\bigcup_{i \in I} S_i$ is also linearly independent. $\therefore \bigcup_i S_i \in \mathcal{C}$.

By Zorn's lemma, $\exists$ maximal linearly independent set M.

We claim that $\mathrm{span}(M) = V$.

If not, $\exists \vec{v} \in V \ni \vec{v} \notin \mathrm{Span}(M)$.

Then: $M \cup \{\vec{v}\}$ is linearly independent.

But $M \subset M \cup \{\vec{v}\}$. Contradiction to Zorn's lemma.

$\therefore$ ① $\mathrm{Span}(M) = V$

② M is L.I.

$\Rightarrow$ M is a basis.

Theorem: Every spanning set of a non-zero vector space V contains a basis of V.

Proof: Let $\mathcal{S}$ be a spanning set of V.

Let $\mathcal{C}$ be the collection of linearly independent subsets of $\mathcal{S}$.

Then $\mathcal{C} \neq \phi$ (as $\{\vec{v}\} \in \mathcal{C}$ for any $\vec{v} \neq \vec{0} \in \mathcal{S}$)

Then, $\mathcal{C}$ under $\subseteq$ is partially ordered.

Let $\{L_i\}_{i \in I}$ be a chain in $\mathcal{C}$.

Then $\bigcup_i L_i \in \mathcal{C}$ and $\bigcup_i L_i$ is an upper bound.

By Zorn's lemma, there is a maximal element $\mathcal{B}$ in $\mathcal{C}$.

(i.e. a linearly independent subset of $\mathcal{S}$ which is maximal)

We'll show that $\mathrm{Span}(\beta) = \mathrm{Span}(\mathcal{S}) = V$

It suffices to show that for any $\vec{v} \in \mathcal{S}$, $\vec{v} \in \mathrm{Span}(\beta)$.

If not, suppose $\vec{v} \notin \mathrm{Span}(\beta)$.

Then: $\beta \cup \{\vec{v}\}$ is linearly independent subset of $\mathcal{S}$.

Hence, $\beta \cup \{\vec{v}\} \in \mathcal{C}$. But $\beta \cup \{\vec{v}\} \supsetneq \beta$.

Contradicting to the fact that β is maximal.

$\therefore$ $\vec{v} \in \mathrm{Span}(\beta)$.

$\therefore$ $\mathrm{Span}(\beta) = \mathrm{Span}(\mathcal{S}) = V$.

$\therefore$ β is a basis.

Remark: To find a basis inside a spanning set $\mathcal{S}$, it's natural to find a minimal spanning set of V inside $\mathcal{S}$.

If $\mathcal{M} \subset \mathcal{S}$ is minimal, then $\mathcal{M}$ is linearly independent.

If not, we can find $v \in \mathcal{M} \ni \operatorname{Span}(\mathcal{M} \setminus \{\vec{v}\}) = \operatorname{Span}(\mathcal{M})$, contradicting the minimality of $\mathcal{M}$.

One might consider Zorn's lemma as follows:

Let $\mathcal{C}$ be the set of all spanning subsets of $\mathcal{S}$, partially order $\mathcal{C}$ by reverse inclusion. That is: $S_1 \in \mathcal{C}$ and $S_2 \in \mathcal{C}$, $S_1 \le S_2$ iff $S_1 \supseteq S_2$.

For any chain $\{S_i\}_{i \in I}$ in $\mathcal{C}$, $\bigcap_{i \in I} S_i$ is the upper bound.

If $\bigcap_{i \in I} S_i \in \mathcal{C}$, then Zorn's lemma tells us $\mathcal{C}$ has a maximal element (i.e. minimal spanning set)

BUT: $\bigcap_{i \in I} \mathcal{S}_i$ MAY NOT always in $\mathcal{C}$!!

Linear Transformation

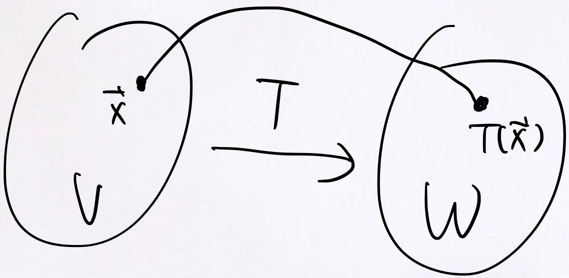
Definition: Let V and W be vector spaces over F.

A *linear transformation* from V to W is a map $T: V \to W$

such that: (a) $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$

(b) $T(a\vec{x}) = aT(\vec{x})$

for all $\vec{x}, \vec{y} \in V$, $a \in F$.



<u>Proposition:</u> Let $T: V \to W$ be a linear transformation. Then:

(i) $T(\vec{0}_V) = \vec{0}_W$

(ii) $T\left(\sum_{i=1}^{n} a_i \vec{x}_i\right) = \sum_{i=1}^{n} a_i T(\vec{x}_i)$ $\forall\ \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n \in V$

$a_1, a_2, \ldots, a_n \in F$.

(T preserves linear combination)

(i) $T(\vec{0}_V) = T(\vec{0}_V + \vec{0}_V) = T(\vec{0}_V) + T(\vec{0}_V)$

$\Rightarrow T(\vec{0}_V) = \vec{0}_W$. (Cancellation law)

(ii) Use math. induction (exercise)

Examples: • For any vector spaces V and W, we have:

(a) The zero transformation $T_0 : V \to W$ defined by $T_0(\vec{x}) := \vec{0}_W$ for $\forall \vec{x} \in V$

(b) The identity transformation $I_V : V \to V$ defined by $I_V(\vec{x}) = \vec{x}$ for $\forall \vec{x} \in V$.

• Let $A \in M_{m\times n}(F)$ be a $m\times n$ matrice . F.

Define: $L_A : F^n \to F^m$ as: (F^n = space of col vectors of size n)

$$L_A(\vec{x}) \overset{\text{def}}{=} A\vec{x}$$

L_A is called the left multiplication by A.

• $T : M_{m\times n}(F) \to M_{n\times m}(F)$ defined by $T(A) \overset{\text{def}}{=} A^t$ (transpose of A)

- $T : P_n(\mathbb{R}) \to P_{n-1}(\mathbb{R})$ defined by $T(f(x)) = f'(x)$ (derivative of f)
  is a lin. transf.

- Let a and $b \in \mathbb{R}$, $a < b$. Then,

  $T : C(\mathbb{R}) \to \mathbb{R}$ defined by:
  (Space of continuous functions)

$$T(f) \stackrel{def}{=} \int_a^b f(t)\,dt$$

Null space or Range

Definition: Let V and W be vector spaces and $T: V \to W$ be a linear transformation

Then, the null space (or kernel) of T is defined as:

$$N(T) :\stackrel{\text{def}}{=} \{ \vec{x} \in V : T(\vec{x}) = \vec{0}_W \} \subset V$$

the range (or image) of T is defined as:

$$R(T) := \{ T(\vec{x}) : \vec{x} \in V \} \subset W$$

e.g. For $I_V : V \to V$, $N(I_V) = \{\vec{0}_V\}$, $R(I_V) = V$
(identity)

For $T_0 : V \to W$, $N(T_0) = V$, $R(T_0) = \{\vec{0}_W\}$
(zero transf)

- $L_A : F^n \to F^m \quad (A \in M_{m\times n}(F))$

$N(L_A) = N(A) =$ null space of A

$R(L_A) = \mathcal{C}(A) =$ col space of A (space of linear combination of col vectors of A)

$$A = \begin{pmatrix} | & | & \cdots & | \end{pmatrix}$$

- For $T : P_n(\mathbb{R}) \to P_{n-1}(\mathbb{R})$ defined by $T(f(x)) = f'(x)$, then:

$N(T) = \{ a_0 \in P_n(\mathbb{R}) : a_0 \in \mathbb{R} \}$

$R(T) = P_{n-1}(\mathbb{R})$

Proposition: Let $T: V \to W$ be a linear transformation.

Then: $N(T)$ and $R(T)$ are subspaces of V and W respectively.

Proof: $\because$ $T(\vec{0}_v) = \vec{0}_w$

$\therefore$ $\vec{0}_v \in N(T)$ and $\vec{0}_w \in R(T)$

Let $\vec{x}$ and $\vec{y} \in N(T)$ and $a \in F$. Then:

$$T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y}) = \vec{0}_w + \vec{0}_w = \vec{0}_w \quad \text{and}$$

$$T(a\vec{x}) = aT(\vec{x}) = a\vec{0}_w = \vec{0}_w$$

$\therefore$ $\vec{x}+\vec{y} \in N(T)$ and $a\vec{x} \in N(T)$

$\therefore$ $N(T)$ is a subspace of V.

Now, let $\vec{u}, \vec{v} \in R(T)$ and $a \in F$.

Then: $\exists\, \vec{x}, \vec{y} \in V$ such that $\vec{u} = T(\vec{x})$ and $\vec{v} = T(\vec{y})$

So, $T(\underbrace{\vec{x}+\vec{y}}^{\in V}) = T(\vec{x}) + T(\vec{y}) = \vec{u} + \vec{v} \Rightarrow \vec{u}+\vec{v} \in R(T)$

$T(\overbrace{a\vec{x}}^{\in V}) = a\,T(\vec{x}) = a\vec{u} \Rightarrow a\vec{u} \in R(T)$

$\therefore$ $R(T)$ is a subspace of W.

Remark: $T: V \to W$ is onto iff $R(T) = W$

(follows from the def)

Proposition: A linear transformation $T: V \to W$ is one-to-one iff $N(T) = \{\vec{0}\}$.

Pf: (Recap: One-to-one $\Leftrightarrow$ " $T(\vec{x}) = T(\vec{y}) \Rightarrow \vec{x} = \vec{y}$ "

($\Rightarrow$) If T is one-to-one, then: for any $\vec{x} \in N(T)$, we have

$$T(\vec{x}) = \vec{0}_W = T(\vec{0}_V)$$

$$\Rightarrow \quad \vec{x} = \vec{0}_V$$

This implies $N(T) = \{\vec{0}_V\}$.

($\Leftarrow$) Suppose $N(T) = \{\vec{0}_V\}$

Let $\vec{x}, \vec{y} \in V$ such that $T(\vec{x}) = T(\vec{y})$.

Then: $T(\vec{x}) - T(\vec{y}) = T(\vec{x} - \vec{y}) = \vec{0}$

This implies $\vec{x} - \vec{y} \in N(T) = \{\vec{0}_V\}$

$\therefore \vec{x} - \vec{y} = \vec{0}_V$ or $\vec{x} = \vec{y}$.

$\therefore$ T is 1-1.

Definition: Let $T: V \to W$ be a linear transformation.

If $N(T)$ and $R(T)$ are finite-dimensional, we define:

- Nullity is denoted as $\text{Nullity}(T)$ is the dimension of $N(T)$.
- Rank is denoted as $\text{Rank}(T)$ is the dimension of $R(T)$.

Lemma: Let $T: V \to W$ be a linear transformation. If $\beta = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is a basis for V, then:

$$R(T) = \text{Span}(T(\beta)) \overset{\text{def}}{=} \text{span}\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_n)\}$$

Proof: $\because\ T(\vec{v}_j) \in R(T)$ for $j = 1, 2, \dots, n$

and $R(T)$ is subspace.

$$\therefore\ \mathrm{Span}\{\underset{\substack{\uparrow \\ R(T)}}{T(\vec{v}_1)}, \underset{\substack{\uparrow \\ R(T)}}{T(\vec{v}_2)}, \dots, \underset{\substack{\uparrow \\ R(T)}}{T(\vec{v}_n)}\} \subset R(T)$$

Conversely, let $\underset{\substack{\parallel \\ T(\vec{x})}}{\vec{w}} \in R(T)$ where $\vec{x} \in V$.

Then: $\exists\ a_1, a_2, \dots, a_n \in F$ s.t. $\vec{x} = \sum_{j=1}^{n} a_j \vec{v}_j$.

$$\text{So, } \vec{w} = T(\vec{x}) = T\left(\sum_{j=1}^{n} a_j \vec{v}_j\right) = \sum_{j=1}^{n} a_j T(\vec{v}_j) \in \mathrm{Span}\{T(\vec{v}_1), \dots, T(\vec{v}_n)\}$$

$\therefore\ R(T) \subset \mathrm{Span}(T(\beta))$ $\qquad \therefore\ R(T) = \mathrm{Span}(T(\beta))$

Example: $T: P_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ defined by:

$$T(f) = \begin{pmatrix} f(0) & f(1) \\ 0 & f'(0) \end{pmatrix}$$

Take $\beta = \{1, x, x^2\}$ as basis of $P_2(\mathbb{R})$.

We have: $R(T) = \text{span}\{T(1), T(x), T(x^2)\} = \text{Span}(T(\beta))$

$$= \text{span}\underbrace{\left\{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\}}_{\text{Lin. indep.}}$$

$\Rightarrow \text{Rank}(T) = 3$

Theorem: (Rank - Nullity Theorem)

Let V and W be vector spaces s.t. V is finite-dimensional.

Then for any linear transformation $T: V \to W$, we have:

$$\text{nullity}(T) + \text{Rank}(T) = \dim(V)$$

Proof: Let $n = \dim(V)$ and $k = \dim(N(T)) \leq n$

Choose a basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ for $N(T)$ and extend it to a basis $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ for V.

Claim: $S = \{T(\vec{v}_{k+1}), T(\vec{v}_{k+2}), \ldots, T(\vec{v}_n)\}$ is a basis for $R(T)$.

① $$R(T) = \operatorname{span}\{\underset{\overset{\parallel}{0}}{T(\vec{v}_1)}, T(\vec{v}_2), \underset{\overset{\parallel}{0}}{\overset{T(\vec{v}_k)}{\ldots}}, T(\vec{v}_n)\}$$

$$= \operatorname{span}\{\underbrace{T(\vec{v}_{k+1}), \ldots, T(\vec{v}_n)}_{S}\} = \operatorname{span}(S)$$

② Now suppose $\exists\, b_{k+1}, b_{k+2}, \ldots, b_n \in F$ s.t.

$$\sum_{i=k+1}^{n} b_i T(\vec{v}_i) = \vec{0}.$$

Then, by linearity, we have: $$T\left(\sum_{i=k+1}^{n} b_i \vec{v}_i\right) = \vec{0}$$

$$\Rightarrow \sum_{i=k+1}^{n} b_i \vec{v}_i \in N(T)$$

$\therefore \sum_{i=k+1}^{n} b_i \vec{v}_i = \sum_{i=1}^{k} c_i \vec{v}_i$ for some $c_1, \ldots, c_k \in F$.

But then: $\sum_{i=1}^{k} (-c_i) \vec{v}_i + \sum_{i=k+1}^{n} b_i \vec{v}_i = \vec{0}$

$\therefore$ $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for V and so it is lin. ind.

$\therefore$ $(-c_i) = 0$ for $i = 1, 2, \ldots, k$

$b_i = 0$ for $i = k+1, k+2, \ldots, n$

$\therefore$ S is lin. ind.

$\therefore$ S is basis for $R(T)$.

$$\therefore \quad \mathrm{Nullity}(T) + \mathrm{Rank}(T)$$

$$= k + (n-k)$$

$$= n = \dim(V)$$