Math 110, Spring 2019

Contents

1 Vector Space

1.1 Vector Space over a field and subspace

Recall that $(\mathbb{F}, +, \cdot)$ or $\langle \mathbb{F}, +, \cdot \rangle$, where $\mathbb{F}$ is a set, and $+, \cdot$ are binary operations.
We know that $(\mathbb{F}, +)$ and $(\mathbb{F} \setminus \{0\}, \cdot)$ and $+, \cdot$ satisfy distributivity.

Definition 1.1

V is a vector space over a field $\mathbb{F}$ is V is equipped with vector addition $+ : V \times V \to V$ and scalar multiplication $\cdot : \mathbb{F} \times V \to V$.

1.1.1 Lists (and vector spaces of lists)

Example 1.2

$\mathbb{R}^n, \mathbb{C}^n$, or generally $\mathbb{F}^n$.

$$\mathbb{R}^n = \{(x_1, x_2, \cdots x_n) : x_i \in \mathbb{R} \ \forall \ j = 1, 2 \cdots n\}$$

$$\mathbb{F}^n = \{(x_1, x_2, \cdots x_n) : x_i \in \mathbb{F} \ \forall \ j = 1, 2 \cdots n\}$$

We claim that $\mathbb{F}^n$ is a vector space over $\mathbb{F}$ provided $\mathbb{F}$ is a field. We can define addition and scalar multiplication as

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

$$\alpha \cdot (x_1, x_2, \ldots, x_n) = (\alpha \cdot x_1, \alpha \cdot x_2, \cdots, \alpha \cdot x_n) \quad \alpha \cdot x_i \in \mathbb{F}$$

What rules / axioms should we impose?

- Commutativity
  $\overline{v} + \overline{w} = \overline{w} + \overline{v} \ \forall \ \overline{v}, \overline{w} \in V$.
- Associativity
  $(\overline{v} + \overline{w}) + \overline{u} = \overline{v} + (\overline{w} + \overline{u}) \ \forall \ \overline{v}, \overline{w}, \overline{u} \in V$.
- Additive Identity
  $\exists \ \overline{0} \in V : \overline{v} + \overline{0} = \overline{v} + \overline{0} = \overline{v}$
- Additive Inverse $\forall v \in V \ \exists \ \overline{w} \in V : \overline{v} + \overline{w} = \overline{0}$.
- (Mixed) Scalar Multiplication Rules
  $1 \cdot v \in v \qquad \forall v \in V$
- Distributivity:
  $(\alpha + \beta) \cdot \overline{v} = \alpha \cdot \overline{v} + \beta \cdot \overline{v} \qquad \forall \ a, b, \in \mathbb{F} \qquad \forall \ \overline{v} \in V$
  $\alpha \cdot (\overline{v} + \overline{w}) = \alpha \cdot \overline{v} + \alpha \cdot \overline{w} \qquad \forall \ a \in \mathbb{F} \qquad \forall \ \overline{v}, \overline{w} \in V$

Now we can check that these rules hold in $\mathbb{F}^2$:

$$(0,0,\cdots,0)+(x_1,x_2,\ldots,x_n)=(x_1,x_2,\ldots,x_n)$$
$$(x_1,x_2,\ldots,x_n)+(-x_1,-x_2,\ldots,-x_n)=(0,0,\cdots,0)$$

Basic Observation $\overline{0}$ is unique

Proof. Suppose $\overline{0}_1$ and $\overline{0}_2$ are both identity element with respect to +:

$$\overline{0}_1=\overline{0}_1+\overline{0}_2+\overline{0}_2$$

A contradiction. ∎

Additive inverse ate unique, i.e., if $\overline{v}+\overline{w}=\overline{0}$ and $\overline{v}+\overline{w}=0$, then $\overline{u}=\overline{w}$.

Proof. Suppose $\overline{v}+\overline{w}=\overline{0}$ and $\overline{v}+\overline{w}=0$, then

$$\overline{w}=\overline{w}+\overline{0}=\overline{w}+(\overline{v}+\overline{u})=(\overline{w}+\overline{v})+\overline{u}=\overline{0}+\overline{u}=\overline{u}$$

A contradiction. ∎

Additive Inverse

$$\begin{aligned}(-1)\cdot\overline{v}+\overline{v}&=(-1)\cdot\overline{v}+1\cdot\overline{v}\\&=((-1)+1)\cdot\overline{v}\\&=\overline{0}\cdot v\\0\cdot\overline{v}+0\cdot\overline{v}&=(0+0)\cdot\overline{v}\\&=0\overline{v}\implies\boxed{0\cdot\overline{v}=\overline{0}}\end{aligned}$$

Additive inverse $\implies$ $0\cdot\overline{v}=\overline{0}$ on both sides.

1.2 Subspaces

Definition 1.3

V is a vector space over a field $\mathbb{F}$, Let $W\subseteq V$.
W is called a subspace of V if W equipped with the same operations $+,\cdot$ inherited from V is still a vector space.

Is it enough for W to be just a subset of V?
Suppose $V=\mathbb{R}^3$ is a vector space over $\mathbb{R}$. Let $W:=\{(1,1,1)\}$, the additive inverse doesn't exist. Note that W is not closed in addition and scalar multiplication.

$$W:=\{(x,0,0)\;:\;x_1\in\mathbb{R}\}$$

Why is $\overline{0}$ in every subspace?
We know that a vector space is a *non empty* set, and W is closed under multiplication, so since $0\in\mathbb{F}$, therefore $0\cdot\overline{v}=\overline{0}\in W$.

Remark 1.4

If $\overline{v}+\overline{w}=\overline{v}$, for some $\overline{v}\in V$, then $\overline{w}=\overline{0}$.

Proof. Suppose $\overline{v}+\overline{w}=\overline{v}$, then $-\overline{v}+\overline{v}+\overline{w}=-\overline{v}+\overline{v} \implies \overline{0}+\overline{w}=\overline{0} \implies \overline{w}=\overline{0}$ ∎

We recall that $\mathbb{F}^n$ is defined as

$$\mathbb{F}^n=\{(x_1,x_2,\ldots,x_n)\ :\ x_i\in\mathbb{F}\}$$

We can define $\mathbb{F}^S$ for S being a set as $\mathbb{F}^S=\left\{f:\underbrace{S}_{\text{no structure needed}}\to\underbrace{\mathbb{F}}_{\text{field}}\right\}$

We can define addition and multiplication as

$$(f+g)(s):=f(s)+g(s)\ \forall s\in S$$

$$(\lambda\cdot f)(s):=\lambda\cdot f(s)\in\mathbb{F}$$

Suppose $S=\{1,2,3\}$, what is $\mathbb{F}^S$ or $\mathbb{R}^S$? We can thought of $\mathbb{R}^S$ as $\mathbb{R}^3$..... why?

Remark 1.5

We can conclude $\mathbb{F}^S\cong\mathbb{F}^{|S|}$, where $|S|$ is the cardinality of S. If S is finite.

What is $\mathbb{R}^{\mathbb{N}}$? ← the set of all of all real sequences.

Remark 1.6

In the book we uses $\mathbb{R}^\infty$, we can conclude that

$$\mathbb{R}^\infty\cong\mathbb{R}^{\mathbb{N}}$$

We say that W is a subspace of $\mathbb{R}^\infty$ with $+,\cdot$

$$W:=\left\{s\ :\ \lim_{n\to\infty}s(n)=0\right\}$$

Proof. We can see that if $\lim_{n\to\infty}s(n)=0$ and $\lim_{n\to\infty}t(n)=0$, then $\lim_{n\to\infty}(s+t)(n)=0$.
If $\lambda\in\mathbb{R}$, then $\lim_{n\to\infty}s(n)=0\implies\lim_{n\to\infty}(\lambda\cdot s)(n)=0$.
The zero sequence is in W so $\overline{0}\in V$. Therefore W is a subspace of V. ∎

Theorem 1.7

W is a subspace of V iff W is closed under addition, multiplication by scalar multiplication by scalars, and $\overline{0}\in V$.

Since the operation in inherent from vector space V, we do not need to verify the other property since they all for all V and W is a subspace of V.
How do we form new subspaces from existing ones?

Theorem 1.8

Suppose W_1, W_2 are subspaces of V, then $W_1 \cap W_2$ is a subspace of V.

Proof. We know that W_1, W_2 are subspaces of V, therefore $\overline{0} \in W_1$ and $\overline{0} \in W_2$, then $0 \in W_1 \cap W_2$. Suppose $\overline{v}, \overline{u} \in W_1 \cap W_2$, we know that $\overline{v}, u \in W_1$ and $\overline{v}, u \in W_2$. Since W_1, W_2 is a subspace, therefore $\overline{u} + \overline{v} \in W_1 \wedge \overline{u} + \overline{v} \in W_2 \implies \overline{u} + \overline{v} \in W_1 \cap W_2$, therefore $W_1 \cap W_2$ is closed under vector addition.
Suppose $\alpha \in \mathbb{F}$ and $\overline{v} \in W_1 \cap W_2$. We know that $\alpha \cdot \overline{v} \in W_1$ and $\alpha \cdot \overline{v} \in W_2$ since they are both subspaces of V. Therefore we conclude $\alpha \cdot \overline{v} \in W_1 \cap W_2$, therefore $W_1 \cap W_2$ is closed under multiplication.
Therefore $W_1 \cap W_2$ is a subspace of V. ∎

Proposition 1.9

The union of two subspace of V are generally not a subspace of V

Proof. We can see that span $\{e_1\}$ and span $\{e_2\}$ is not a subspace if $\mathbb{R}^2$ as $(1,1) \notin W_1 \cup W_2$ ∎

Theorem 1.10

Union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

Proof. The proof is left as an exercise. ∎

Theorem 1.11

$W_1 + W_2$ is a subspace of V.

Proof. (identity) $\overline{0} \in W_1 \wedge \overline{0} \in W_2 \implies \overline{0} + \overline{0} = \overline{0} \in W_1 + W_2$.
(closure under addition) Suppose $\overline{w}_1 + \overline{w}_2 \in W_1 + W_2$ and $\tilde{w}_1 + \tilde{w}_2 \in W_1 + W_2$. We compute $(\overline{w}_1 + \overline{w}_2) + (\tilde{w}_1 + \tilde{w}_2) = \underbrace{(\overline{w}_1 + \tilde{w}_1)}_{\in W_1} + \underbrace{(\overline{w}_2 + \tilde{w}_2)}_{\in W_2} \implies (\overline{w}_1 + \overline{w}_2) + (\tilde{w}_1 + \tilde{w}_2) \in W_1 + W_2$.
(closure under scalar multiplication) Suppose $\overline{w}_1 + \overline{w}_2 \in W_1 + W_2$, and $\lambda \in \mathbb{F}$, we compute $\lambda \cdot (\overline{w}_1 + \overline{w}_2) = \underbrace{(\lambda \cdot \overline{w}_1)}_{\in W_1} + \underbrace{(\lambda \cdot \overline{w}_2)}_{\in W_2} \implies \lambda \cdot (\overline{w}_1 + \overline{w}_2) W_1 + W_2$ ∎

Remark 1.12

$W_1 + W_2 + \cdots + W_n$ if the smallest subspace containing $W_1, W_2, \ldots, W_n$.
If $\tilde{v}$ is a subspace of $V \supseteq W_j \ \forall j$, since $\tilde{v}$ is closed under $+$, $w_1 + w_2 + \cdots + w_n \in W_n$

Example 1.13

Suppose $V = \mathbb{R}^3$. Let $W_1 = \text{span}\{e_1, e_2\}$, $W_2 = \text{span}\{(0,1,1)\}$, $W_3 = \text{span}\{(x,y,z) : x+y+z=0\}$. What is $W_1 + W_2 + W_3$?

Note that $(0,0,1) = \underbrace{\left(0, \frac{1}{2}, \frac{1}{2}\right)}_{\in W_2} + \underbrace{\left(0, -\frac{1}{2}, \frac{1}{2}\right)}_{\in W_3}$. We also know that $(1,0,0) \in W_1$ and $(0,1,0) \in W_2$, therefore $W_1 + W_2 + W_3 = \mathbb{R}^3$

Discussion

Definition 1.14

A vector space, is often denoted as $(\underbrace{\mathbb{F}}_{\text{scalars}}, \underbrace{V}_{\text{vectors}}, \underbrace{\cdot : \mathbb{F} \to B}_{scaling})$

Example 1.15

$(\mathbb{R}, \mathbb{R}^n, \cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n)$ is a vector space.

Example 1.16

$(\mathbb{R}, \mathbb{R}, \cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R})$ is also a vector space.

Notion of a field

Suppose $F = \{0,1,2,3\}$. Can F be a field?

Definition 1.17

A subset W of the vector space V is a subspace of V if it satisfy the following:

1) $\overline{0} \in W$
2) $+ : W \times W \to W \subseteq V$ (closure under addition)
3) $\cdot : \mathbb{F} \times W \to W \subseteq V$ (closure under scalar multiplication)

Example 1.18

Can we find a subset W of V such that W satisfy property 1), 2) but not 3)?
Suppose $W = \{(x,0) : x \in \mathbb{Z}\}$ the proof is trivial and is left as an exercise.

Example 1.19

The set of functions $\{f : (0,\infty) \to \mathbb{R}\} = \mathbb{R}^{(0,\infty)}$ is a vector space. We claim that W is a subspace of V.

$$W = \{f : (0,\infty) \to \mathbb{R} : f'(1) = 0\}$$

Proof. We begin by verifying the three properties

1) The zero function is in W
2) Suppose $f, g \in W$, then $(f+g)'(1) = f'(1) + g'(1) = 0 + 0 = 0 \implies f(x) + g(x) \in V$
3) Suppose $f \in W$ and $\lambda \in \mathbb{F}$, then $\lambda \cdot f'(1) = \lambda \cdot 0 = 0 \implies \lambda \cdot f(x) \in V$

Therefore W is a subspace of V. ∎

1.3 Direct Sum

Definition 1.20

Let $(\mathbb{F}, V, \cdot : \mathbb{F} \times V \to V)$ be a vector space. Given that $U_1, U_2, \dots, U_n \subseteq V$ are subspaces of V, we can define the sum of the subspaces as

$$U_1 + U_2 + \cdots + U_n = \{u_1 + u_2 + \cdots + u_n : u_i \in U_i\}$$

Proof. 1) We can see that $\overline{0} \in U_i\ \forall i$, and $\overline{0} + \overline{0} + \cdots + \overline{0} = \overline{0}$

2) Suppose $\overline{x} = \sum_{i=1}^{n} \overline{x}_i \in U_i$ and $\overline{y} = \sum_{j=1}^{n} \overline{y}_j \in U_j$, we can see that $\overline{x} + \overline{y} = \sum_{k=0}^{n} \overline{x}_k + \overline{y}_k \in U_k$, therefore it's closed under addition.
3) Suppose $\lambda \in \mathbb{F}$ and $\overline{x} = \sum_{i=1}^{n} \overline{x}_i \in U_i$, we compute $\lambda \cdot \overline{x} = \lambda \cdot \sum_{i=1}^{n} \overline{x}_i \in U_i$, therefore it's closed under scalar multiplication.

∎

Definition 1.21

We say that $U_1 + U_2 + \cdots + U_n$ is a direct sum, denoted as $U_1 \oplus U_2 \oplus \cdots \oplus U_n$ if for every $\overline{v} \in U_1 + U_2 + \cdots + U_n$, $\overline{v} = \overline{u}_1 + \overline{u}_2 + \cdots + \overline{u}_n$ has a unique representation.

Remark 1.22

How best to check $U_1 + U_2 + \cdots + U_n$ is a direct sum?
Check that $U_i \cap U_j = \{\overline{0}\}$. We will go over in depth later.

What about $W_1 + W_2 + \cdots W_n$ being a direct sum?

Theorem 1.23

The sum of subspaces $W_1, W_2, \ldots, W_n$:

$$W_1 + W_2 + \cdots + W_n$$

is a direct sum iff $\overline{0}$ can be written in only **one way** as a sum

$$\overline{w}_1 + \overline{w}_2 + \cdots + \overline{w}_n = 0$$

namely $\overline{0} + \overline{0} + \cdots + \overline{0} = 0$.

Remark 1.24

If $W_1 \cap W_2 \neq \{0\}, W_1 \cap W_3 \neq \{0\}, W_2 \cap W_3 \neq \{0\}$, it is not possible for W_1, W_2, W_3 to be a direct sum, However, the opposite of the proposition is not sufficient for being a direct sum as demonstrated in Remark 1.25.

Remark 1.25

$W_1 \cap W_2 = \{0\}$, $W_1 \cap W_2 = \{0\}$, $W_2 \cap W_3 = \{0\}$ and $W_1 + W_2 + W_3$ being not a direct sum is possible. For example, consider $\mathbb{R}_2$, for line $x = y, y = 0$ and $x = 0$, we can see that they only have the trivial intersection but they are not a direct sum. (credit: Catherine)

2 Finite Dimensional Vector Spaces

2.1 Linear Dependence and Independence

Definition 2.1

We will works with lists of vectors $v_1, v_2, \ldots, v_k$, then the span of $v_1, v_2, \ldots, v_k$ can be defined as

$$\operatorname{span}(\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_k) := \{\alpha_1\overline{v}_1 + \alpha_2\overline{v}_2 + \cdots + \alpha_k\overline{v}_k\}$$
$$\forall \alpha_i \in \mathbb{F}$$

If the the list happens to cover the entire vector space V, we call the list a spanning list of V.

Definition 2.2

V is finite dimensional if V is a span of finitely many vectors.

Remark 2.3

V is not finite dimensional is logically equivalent to V is infinite dimensional.

Example 2.4

Consider the vector spaces: $\mathcal{P}(x) := \{\alpha_0 + \alpha_1 x + \cdots + a_k x^k : a_j \in \mathbb{F} \text{ for some } k\}$. We can see that $\mathcal{P}(x) \subseteq \mathbb{F}^{\mathbb{F}}$, and $\mathcal{P}(x)$ is infinite dimensional.

Definition 2.5

We can define the degree of a polynomial, denoted as $\deg(f(x))$, is the highest power of x where hose coefficient (α_k) is nonzero. The zero function $f(x) = 0$ has $-\infty$ degree.

Example 2.6

$\mathcal{P}(x)$ is infinite dimensional.

Proof. Suppose $\mathcal{P}(x) = \operatorname{span}(f_1, f_2, \ldots, f_k)$, where f_j is polynomials, for all j. Let

$$D := \max\{\deg(f_1), \deg(f_2), \cdots, \deg(f_k)\}$$

Suppose $f(x) = x^{D+1} \in \mathcal{P}(x)$ however, $x \notin \operatorname{span}(f_1, f_2, \ldots, f_k)$. Since $f(x)$ is not a linear combination of $f_1, f_2, \ldots, f_k$. A contradiction, therefore $\mathcal{P}(x)$ is a infinite dimensional vector space. ∎

Definition 2.7

V has dimension k over $\mathbb{F}$ if you can find vectors $v_1, v_2, \ldots, v_k$ such that

$$\forall v \in V : v = \sum f_i v_i \text{ uniquely}$$

Definition 2.8

$\mathcal{P}_d(x) :=$ all polynomials in $g(x)$ of degree $\leq d$.
Note that $\{1, x, x^2, \cdots, x^d\}$ is a spanning list for $\mathcal{P}_d(x)$

Definition 2.9

A list $v_1, v_2, \ldots, v_k \in V$ is called **linearly independent** if the equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0 \implies \alpha_1, \alpha_2, \ldots, \alpha_k = 0$$

Definition 2.10

A list $v_1, v_2, \ldots, v_k \in V$ is called **linearly dependent** if it is not independent.

Digression on Logic

Logic: $A \implies B$ is equivalent to $\neg A \vee B$. Then we know that

$$\neg(A \implies B) \iff (\neg(\neg A \vee B)) \iff A \wedge \neg B$$

Definition 2.11 (The better definition)

A list $v_1, v_2, \ldots, v_k \in V$ is called **linearly dependent** if for equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

has a nontrivial solution such that $v_1, v_2, \ldots, v_k \neq 0$

Example 2.12

Is $\{\}$ linearly independent?
By definition, it is linearly independent, because it is not linearly dependent. A set S is linearly dependent if there exists a finite set of vectors $v_1, v_2, \dots, v_n$ and corresponding scalars $\alpha_1, \alpha_2, \dots, \alpha_n$ such that there exists at least one $\alpha_i \neq 0$ so that

$$\sum_{i=0}^{n} \alpha_i v_i = 0$$

since α_i doesn't exist, we know that $\{\}$ is linearly independent.

Example 2.13

Is $\{(1,0,0),(0,1,0),(0,0,1)\}$ linearly independent on $\mathbb{R}^3$?

$$\begin{aligned} \alpha_1(1,0,0) + \alpha_2(0,1,0) + \alpha_3(0,0,1) &= (0,0,0) \\ \implies (\alpha_1, \alpha_2, \alpha_3) &= (0,0,0) \\ \implies \alpha_1 = \alpha_2 = \alpha_3 = 0 \end{aligned}$$

Remark 2.14

We can remove vectors from a linearly independent list can still remain independent, however, we cannot guarantee the result if we are still adding vectors; In mathematical terms, any sublist of the list is linearly independent, since $\{\}$ is a sublist of any list, therefore its linearly independent.

Lemma 2.15 (Linear Dependence Lemma)

Suppose $\overline{v}_1, \overline{v}_2, \dots, \overline{v}_h$ is linearly independent. Then there exists j between 1 and k such that

- $v_j \in \operatorname{span}\{\overline{v}_1, \overline{v}_2, \overline{v}_{j-1}\}$
- $\operatorname{span}\{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_h\} = \operatorname{span}\{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_{j-1}, \overline{v}_{j+1}, \cdots, \overline{v}_k\}$

Proof. If $\overline{v}_1, \overline{v}_2, \dots, \overline{v}_k$ is a linearly dependent list, there are coefficients $\alpha_1, \alpha_2, \dots, \alpha_k$ not all 0, such that

$$\alpha_1 \overline{v}_1 + \alpha \overline{v}_2 + \cdots \alpha_k v_k = 0$$

Take j such that α_j is the largest index with $\alpha_j \neq 0$. Then $\alpha_{j+1} = \alpha_{j+2} = \cdots = \alpha_k = 0$ and

$$\overline{v}_j = \frac{-1}{\alpha_j}(\alpha_1 \overline{v}_1 + \alpha \overline{v}_2 + \cdots + \alpha_{j-1} v_{j-1})$$

hence $\overline{v}_j \in \operatorname{span}\{\overline{v}_1, \overline{v}_2, \dots, \overline{v}_{j-1}\}$. ■

Remark 2.16 (Very Important, a.k.a. Magic Lemma)

The length of the independent list $\leq$ length of any spanning list.

Proof. Say $\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_n$ is linearly independent say $\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_m$ is spanning. Then we want to establish that $m \leq n$.

Step 1. Take the list $\overline{u}_1, \overline{v}_1, \overline{v}_2, \ldots, \overline{v}_m$ It is linearly independent since $\overline{u}_1 \in \operatorname{span}\{\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_m\}$. By the linear dependence lemma, there is a j such that $\overline{v}_j$ can be removed (noted that $\overline{u}_1$ cannot be subject to removal since $\overline{u}_1$ comes from a linearly independent list). Consider the new list $\{\overline{u}_1, \overline{v}_1, \overline{v}_2, \ldots, \overline{v}_{j-1}, \cdots, \overline{v}_m\}$

Step 2. We can continue this process by bringing $\overline{u}_2, \overline{u}_3, \cdots \overline{u}_n$, we know that $\overline{u}_i$ since they are linearly independent.

Note that this process preserves linear span of the whole list.
We know that this list contains all the $\overline{u}_i$ (plus possibly some remaining $\overline{v}_j$) and the length of the list is always n. So $\boxed{m \leq n}$. ■

2.2 Bases and Dimension

Definition 2.17

A basis a linearly independent spanning list.

Theorem 2.18

Any two basis in a finite dimensional space have the same number of vectors.

Remark 2.19

The span of $\{\}$ is the zero vector.

Theorem 2.20

Suppose V is a finite dimensional vector space. Let W be a subspace of V, then W is finite dimensional.

Proof. V is finite dimensional means that V is spanned by some k vectors. Consider W. If $W = \{\overline{0}\}$, then w is spanned by the empty list $\overline{0}$. If $W \neq \overline{0}$, there exists $\overline{w}_1 \in W$ such that $W = \operatorname{span}\{\overline{u}_1\}$, done. Otherwise take $\overline{w}_2 \in W \setminus \operatorname{span}\{\overline{w}_1\}$. Repeat this algorithm until it terminates. Now we want to show that this algorithm will terminate at $\overline{w}_k$, we know that $\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_j$ is linearly independent by construction and the linearly dependence lemma. By remark 2.16, we know that the length of any such list will not exceed length k, therefore we know the algorithm will terminate in finite steps. This implies that W is finitely spanned, or W is finite dimension. ■

2.2.1 Dimension

Definition 2.21

Dimension of a vector space V is the cardinality of any basis in a finite dimensional space.

Proposition 2.22 (Criterion for a Basis)

$\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_k$ is a basis for V if and only if any $v \in V$ can be uniquely written as a lineat combination

$$\lambda_1 \overline{v}_1 + \lambda_2 \overline{v}_2 + \cdots + \lambda \overline{v}_n$$

Proof. We know that "can be written as linear combination" is logically equivalent a $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_k$ is a spanning list for V. "uniqueness" is logically equivalent as linear independence. Suppose

$$\overline{v} = \alpha_1 \overline{v}_1 + \alpha_2 \overline{v}_2 + \cdots + \alpha_k \overline{v}_k = \beta_1 \overline{v}_1 + \beta_2 \overline{v}_2 + \cdots + \beta_k \overline{v}_k$$

Not all $\alpha_j = \beta_j$. Then $(\alpha_1 - \beta_1)\overline{v}_1 + (\alpha_2 - \beta_2)\overline{v}_2 + \cdots (\alpha_k - \beta k)\overline{v}_k = \overline{0}$ is a nontrivial linear combination of $\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_k$ and vice versa. ■

Theorem 2.23

Any spanning set for a finite dimensional space can be shrink down to a basis.

Proof. Trivial by the linear dependence lemma. ■

Example 2.24

Consider $\mathcal{P}_2(x)$ is spanned by $\{x^2, (x-1)^2, (x-3)^2, (x-3)^2\}$, we can see that this can be thinned down to $\{x^2, (x-1)^2, (x-2)^2\}$.

Corollary 2.25

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

Proof. Add a spanning list at the back of our given list, then do removal for the linearly independent lemma. ■

Theorem 2.26

Suppose V is finite dimensional and W is a subspace, then there is a subspace U such that $V = W \oplus U$.

Proof. We already know by proceeding stuff W is finite dimensional an its dimensional does not exceed that of V. Take any basis of $\overline{w}_1, \overline{w}_2, \dots, \overline{w}_k$ of W. It's linearly independent so can be enlarged to a basis for V. Suppose the resulting basis is $\overline{w}_1, \overline{w}_2, \dots, \overline{w}_k, \overline{u}_1, \overline{u}_2, \dots, \overline{u}_l$. Take $U = \operatorname{span}(\overline{v}_1, \overline{v}_2, \dots, \overline{v}_l)$. Then $W + U = V$ and $W \cap U = \{0\}$. ■

Remark 2.27

$\operatorname{span}(\overline{n}_1, \overline{n}_2, \dots, \overline{n}_n)$ is a subspace of V by construction.

Example 2.28

Consider $\mathcal{P}(x)T$. We define W as

$$W := \{f \in \mathcal{P}_3(x) : f'(5) = 0\}$$

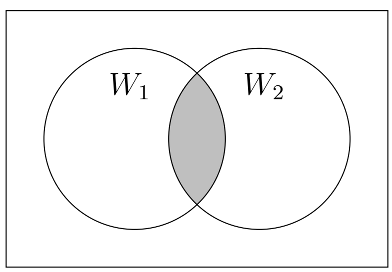
A basis for W can be taken as $\{1, (x-5)^2, (x-5)^3\}$.
Now consider

$$\tilde{W} := \{f \in \mathcal{P}_3(x) : f''(5)\}$$

. A basis for W can be taken as $\{1, (x-5)^2, (x-5)^3\}$.

2.2.2 Dimension of a Sum

Principal of Inclusion for subspaces



$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

Suppose $W_1 \cap W_2$ forms a basis $\overline{w}_1, \overline{w}_2, \dots, \overline{w}_k$. We can extend the basis to

$$\overline{w}_1^{(1)}, w_2^{(1)}, \dots, w_l^{(1)}, \overline{w}_1, \overline{w}_2, \dots, \overline{w}_k$$

is a basis for W_1. Similarly, we can extend the basis to

$$\overline{w}_1^{(2)}, w_2^{(2)}, \dots, w_m^{(2)}, \overline{w}_1, \overline{w}_2, \dots, \overline{w}_k$$

is a basis for W_2.
We want to establish that

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) = l + m + k$$

We want to prove that

$$\overline{w}_1^{(1)}, \overline{w}_2^{(1)}, \dots, \overline{w}_l^{(1)}, \overline{w}_1^{(2)}, w_2^{(2)}, \dots, \overline{w}_m^{(2)}, \overline{w}_1, \overline{w}_2, \dots, \overline{w}_k$$

is a basis for $W_1 + W_2$. We can see that

$$\operatorname{span}(\overline{w}^1{}_1, \overline{w}^1{}_2, \dots, \overline{w}^1{}_l, \overline{w}_1^{(2)}, \overline{w}_2^{(2)}, \dots, \overline{w}_m^{(2)}, \overline{w}_1, \overline{w}_2, \dots, \overline{w}_k) \supseteq U_1, U_2$$

Hence $\operatorname{span}(\dots) \supseteq U_1 + U_2$. Suppose the equation

$$\alpha_1 \overline{w}_1^{(1)} + \alpha_2 \overline{w}_2^{(1)} + \cdots + \alpha_l \overline{w}_l^{(1)} + \beta_1 \overline{w}_1^{(2)} + \beta_1 w_2^{(2)} + \cdots + \beta_m \overline{w}_m^{(2)} + \gamma_1 \overline{w}_1 + \gamma_2 \overline{w}_2 + \cdots + \gamma_k \overline{w} k = \overline{0}$$

Manipulate the equation and we can see

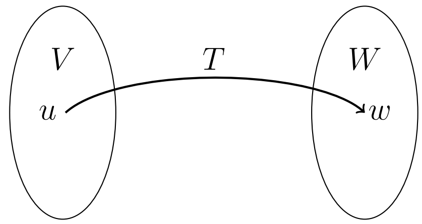
$$\underbrace{\left(\alpha_1 \overline{w}_1^{(1)} + \alpha_2 \overline{w}_2^{(1)} + \cdots + \alpha_l \overline{w}_l^{(1)} + \beta_1 \overline{w}_1^{(2)} + \beta_1 w_2^{(2)} + \cdots + \beta_m \overline{w}_m^{(2)}\right)}_{\in W_1} = -\underbrace{\left(\gamma_1 \overline{w}_1 + \gamma_2 \overline{w}_2 + \cdots + \gamma_k \overline{w} k\right)}_{\in W_2 \setminus W_1}$$

Since they belongs to different sets, clearly they cannot span each other. Therefore

$$\alpha_1 = \alpha_2 = \cdots = \alpha_l = \beta_1 = \beta_2 = \cdots = \beta_m = \gamma_1 = \gamma_2 = \cdots = \gamma_n = 0$$

Hence the list of vectors is also linearly independent.

3 Linear Maps



3.1 Linear Maps as Vector Space

Suppose V and W are two linear spaces over $\mathbb{F}$. T is a function with domain V and codomain W. T is called linear iff

1. $T(u_1 + u_2) = T(u_1) + T(u_2)$
2. $T(\lambda u) = \lambda \cdot T(u)$

$\forall \overline{v}_1, \overline{v}_2 \in V$ and $\forall \overline{v} \in V, \forall \lambda v \in \mathbb{F}$.

Example 3.1

Let $V = \mathbb{R}^3, W = \mathbb{R}^4$. Define T as $(x_1, x_2, x_3) \mapsto (x_1, 0, 0, 0)$

Example 3.2

$T : \mathcal{P}(x) \to \mathcal{P}(x)$, where $f(x) \mapsto \int_{10}^{x} f(x)\ dx$ is a linear map.

Definition 3.3

$\mathcal{L}\{V, W\}$ denotes the set of all linear maps from V to W. Note that $\mathcal{L}\{V, W\}$ with $+$ and $\cdot$ becomes a vector space over $\mathbb{F}$. This requires the additions of functions and multiplications of linear maps by scalars (from $\mathbb{F}$). Given $T_1, T_2 \in \mathcal{L}(V, W)$ we define addition as $(T_1 + T_2)(u) \coloneqq T_1(u) + T_2(u)$, multiplication as $(\lambda T)(u) \coloneqq \lambda \cdot T(u)$.

Theorem 3.4

In finite vector space V, W, let $u_1, u_2, \ldots, u_n$ be a basis for V, let $w_1, w_2, \ldots, w_m$ be any vectors in W. Then there exist a unique linear map $T \in \mathcal{L}\left\{V, W\right\}$ such that $T(u_j) = w_j \forall j$.

Proof. Any vector in V has a unique representation $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = u$.
Define $T(u) \coloneqq \underbrace{\alpha_1 T(u)_1 + \alpha_2 T(u)_2 + \cdots + \alpha_n T(u)_n}_{\in W}$ This makes T a linear map from V to W.
Indeed if $\lambda \in \mathbb{F}$, then $T(\lambda u) = T(\sum_{j=1}^{n} \lambda \alpha_j u_j) = \lambda \sum_{j=1}^{n} \alpha_j wj$. Suppose $\tilde{T}(u_j) = w_j$ for all j, then $T = \tilde{T}$ as a map function by linearity and basis. ∎

3.2 Null Space and Range

Theorem 3.5

Let Null $(T) := \{u \in v : T(u) = 0\}$. Null (T) is a subspace of V.

Theorem 3.6

Let Range $(T) := \{w \in W : T(u) = w\}$. Range (T) is a subspace of W.

Proof. The proof is trivial and is left as an exercise for the reader . ■

Example 3.7

Let $T : f \to f', V := \mathcal{P}(x), W := \mathcal{P}(x)$. Null $(T) = \mathcal{P}_0(x)$, Range $(T) = \mathcal{P}_2(x)$.
Let $T : f \to f'', V := \mathcal{P}(x), W := \mathcal{P}(x)$. Null $(T) = \mathcal{P}_1(x)$, Range $(T) = \mathcal{P}_1(x)$.

Example 3.8

Find a basis of $\mathcal{L}(V, W)$ given bases $\{u_1, u_2, \ldots, u_m\}$ and $\{w_1, w_2, \ldots, w_n\}$ of V and W. The basis consists of $m \times n$ vectors as follows:

$$T_{11} = T(u_1) = w_1, T(u_2) = 0, T(u_3) = 0, \ldots, T(u_m) = 0$$

$$T_{12} = T(u_1) = w_2, T(u_2) = 0, T(u_3) = 0, \ldots, T(u_m) = 0$$

$$\ldots$$

$$T_{mn} = T(u_1) = 0, T(u_2) = 0, T(u_3) = 0, \ldots, T(u_m) = w_n$$

Example 3.9

Let $\mathcal{U} = \{f : \mathbb{R} \to \mathbb{R} : f(x) = f(1 - x)\ \forall x\}$.

1. Show that $\mathcal{U}$ is a subspace of $f : \mathbb{R} \to \mathbb{R}$.

2. Find a complement.

$$\mathcal{W} = \{g : \mathbb{R} \to \mathbb{R} : g(x) = -g(1 - x)\ \forall x\}$$

Proof. We can see that the zero function $f(x) = 0$ satisfies the requirement since $0 = 0$ for all values of x.

Suppose $f(x), g(x) \in \mathcal{U}$, then we compute

$$\begin{aligned}(f + g)(x) &= f(x) + g(x) \\ &= f(1 - x) + g(1 - x) \\ &= (f + g)(1 - x)\end{aligned}$$

Therefore we can see that $\mathcal{U}$ is closed under addition.
Suppose $f(x) \in \mathcal{U}, \lambda \in \mathbb{R}$, then we compute

$$\begin{aligned}(\lambda \cdot f)(x) &= \lambda \cdot f(x)\\ &= \lambda \cdot f(1-x)\\ &= (\lambda \cdot f)(1-x)\end{aligned}$$

Therefore we can see that $\mathcal{U}$ is closed addition. Hence $\mathcal{U}$ is a vector space. ■

Proof. The proof for subspace is similar to part (i) and is omitted here.
We now want to show that $\mathcal{U} + \mathcal{W} = \mathbb{R}^{\mathbb{R}}$. We can see that for $f(x) \in \mathbb{R}^{\mathbb{R}}$, we can rewrite $f(x)$ as

$$f(x) = \frac{f(x) + f(1-x)}{2} + \frac{f(x) - f(1-x)}{2}$$

Clearly $\frac{f(x) + f(1-x)}{2} \in \mathcal{U}$ and $\frac{f(x) - f(1-x)}{2} \in \mathcal{W}$. For uniqueness, suppose that a nonzero $h(x) \in \mathcal{U} \cap \mathcal{W}$, therefore $h(x) = h(1-x) = -h(1-x)$, and the only solution is $f(x) = 0$, a contradiction, therefore $\mathcal{U} \cap \mathcal{W} = \{0\}$. Hence $\boxed{\mathbb{R}^{\mathbb{R}} = \mathcal{U} \oplus \mathcal{W}}$ ■

Theorem 3.10 (Rank-Nullity Theorem also known as the Fundamental Theorem of Linear Maps)
Let V, M be finite dimensional vector spaces, let $T \in \mathcal{L}(V, W)$. Then

$$\dim V = \dim \operatorname{Null} T + \dim \operatorname{Range} T$$

Proof. Let $u_1, u_2, \ldots, u_k$ to be the the basis for the basis for Null T. By the linear independent list extension theory, this list can be extended to a basis of V. Say $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l$ is asujc an extension to a basis of V. We can see that dim $= k + l$. We want to show that Range $T = l$. Consider $Tv_1, Tv_2, \ldots, Tv_l$. We want to show that $Tv_1, Tv_2, \ldots, Tv_l$ is basis for Range T. Notice that $v \in V$ can be written as a linear combination of $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_l v_l$. Then we compute

$$\begin{aligned}Tv &= \alpha_1 Tu_1 + \alpha_2 Tu_2 + \cdots + \alpha_k Tu_k + \beta_1 Tv_1 + \beta_2 Tv_2 + \cdots + \beta_l Tv_l\\ &= \beta_1 Tv_1 + \beta_2 Tv_2 + \cdots + \beta_l Tv_l\end{aligned}$$

hence $Tu \in \operatorname{span}(Tv_1, Tv_2, \ldots, Tv_l)$.
Suppose $\beta_1 Tv_1 + \beta_2 Tv_2 + \cdots + \beta_l Tv_l = 0$. Then $\beta_1 Tv_1 + \beta_2 Tv_2 + \cdots + \beta_l Tv_l \in$ Null T. So

$$\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_l v_l = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k$$

for some $\alpha_1, \alpha_2, \ldots, \alpha_k$ since $u_1, u_2, \ldots, u_k$ form a basis for Null T.
But $v_1, v_2, \ldots, v_l, v_1, v_2, \ldots, v_k$ form a basis for V, all of the coefficient has to be 0. Therefore $Tv_1, Tv_2, \ldots, Tv_k$ is indeed a basis for Range T. ■

Example 3.11 (Direct consequences of the Theorem)

Suppose $\dim W < \dim V$ (both finite), and $T \in \mathcal{L}(V, W)$. Then T cannot be injective.

Proof. T is injective implies that $\operatorname{Null} T = \{0\}$. So $\dim V = 0 + \dim \operatorname{Range} T \leq \dim W < \dim V$, a contradiction. ■

Example 3.12 (Direct consequences of the Theorem)

Suppose $\dim W > \dim V$ (both finite), and $T \in \mathcal{L}(V, W)$. Then T cannot be surjective.

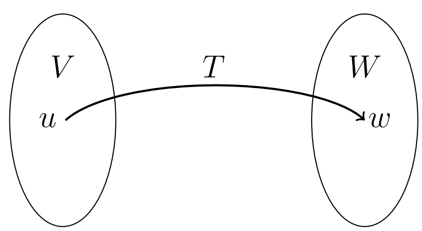
Proof. T is surjective implies that $\operatorname{Range} T = W$. So $\dim V = \dim \operatorname{Null} T + \dim \operatorname{Range} T \geq \dim W > \dim V$, a contradiction. ■

Example 3.13 (Fun Question)

Suppose that $p \in \mathcal{P}(\mathbb{R})$, prove that $\exists q \in \mathcal{P}(\mathbb{R})$ such that $5q'' + 3q' = p$.
[*This exercise can be done without linear algebra, but it's more fun to do it using linear algebra.*]

Proof. Let $d = \deg p$. Define linear transformation $T : \mathcal{P}_{d+1}(\mathbb{R}) \to P_d(\mathbb{R})$ as $T : q \to 5q'' + 3q'$. We can see that $\dim \operatorname{Null} T = 1$, by the rank nullity theorem, know that T must be surjective as $\dim \mathcal{P}_{d+1}(\mathbb{R}) = \dim \operatorname{Null} T + \dim \operatorname{Range} T = 1 + \dim \operatorname{Range} \implies \dim \operatorname{Range} T = \dim P_d(\mathbb{R})$. ■

3.3 Matrix Notation



Recall this diagram, we want to understand T "correctly". Pick a basis $v_1, v_2, \ldots, v_n$ for V and $w_1, w_2, \ldots, w_m$ for W. We can see that $\dim V = n$ and $\dim W = m$. We can define T as

$$Tv_j = A_{1,j}w_1 + A_{2,j}w_2 + \cdots + A_{m,j}w_m$$

Notice that A has the following form

$$\begin{array}{c} w_1 \to \\ w_2 \to \\ \vdots \\ w_m \to \end{array} \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}$$

This is called the matrix representation of T.

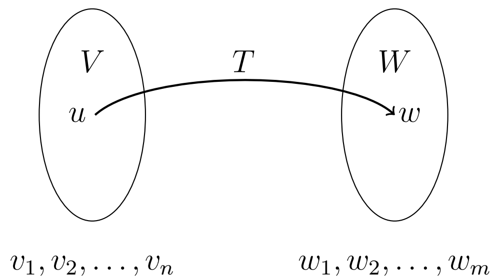
Example 3.14

Let $D : V \to W$ be defined as $D := p \to p'$. Let $V := \operatorname{span}(1, \cos x, \sin x, \cos 2x, \sin 2x) = W$. We can see that

$$\begin{aligned} 1 &\mapsto 0 \\ \cos x &\mapsto -\sin x \\ \sin x &\mapsto \cos x \\ \cos 2x &\mapsto -2\sin 2x \\ \sin 2x &\mapsto \cos 2x \end{aligned} \qquad A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}$$

3.4 Matrix Representation

Recall that if T is a linear transformation



$$Tv_u = \sum_{k=1}^{m} A_{i,k} w_k$$

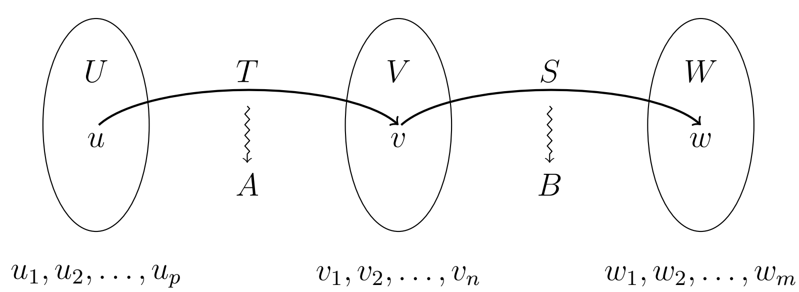
Note that Matrix $A = [A_{i,k}]$ has m rows n columns.

$$\begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}$$

Suppose $v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n, Tv = c_1 Tv_1 + c_2 Tv_2 + \cdots + c_n Tv_n$.

$$Tv = c_1 \sum_{i=1}^{m} A_{i,1} w_i + c_2 \sum_{i=1}^{m} A_{i,2} w_i + \cdots + c_n \sum_{i=1}^{m} A_{i,n} w_n = \sum_{i=1}^{m} \left(\sum_{i=1}^{n} A_{i,j} c_j \right) w_j$$

Notice that the operation is the equivalent as the matrix-vector multiplication.



$$STu_k = S(Tu_k) = S\left(\sum_{j=1}^{n} A_{j,k} v_j \right) = \sum_{j=1}^{n} A_{j,k} \left(Sv_j \right) = \sum_{j=1}^{n} A_{j,k} \sum_{i=1}^{m} B_{i,j} w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} B_{i,j} A_{j,k} \right) w_i$$

Use name $\mathcal{M}(S) \coloneqq B$, $\mathcal{M}(T) \coloneqq A$, $\mathcal{M}(ST) = BA = \mathcal{M}(S) \cdot \mathcal{M}(T)$. So matrix representation multiply as matrices to produce a composition map or product.

Remark 3.15 (Book Keeping)
$A_{*,j}$ denotes the jth column of A.
$A_{i,*}$ denotes the ith row of A.

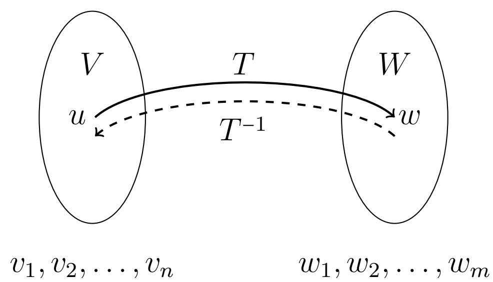
Notice that $\mathcal{M}$ is a linear map, $\mathcal{L}(V, W) \xrightarrow{\mathcal{M}} \mathbb{F}^{m,n}$.

Proposition 3.16
$\mathcal{M}$ is a linear map.

Proposition 3.17
$\mathbb{F}^{m,n}$ has a basis.

Proof. Consider $E_{i,j}$, the matrix consists of all zeros with the exception of 1 in position (i, j). This can be done for all $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. Also notice that $\dim \mathbb{F}^{m,n} = m \cdot n$ ∎

3.5 Invertibility and Isomorphism



Definition 3.18
$T \in \mathcal{L}(V, W)$ is invertible provided that there exists a mapping T^{-1} from W to V (not necessarily linear) such that

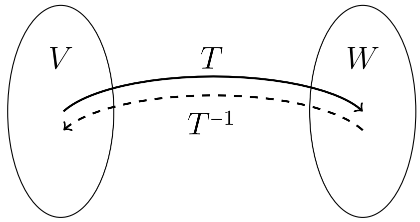
$$T^{-1} \circ T = \mathbb{I}_V$$

$$T \circ T^{-1} = \mathbb{I}_W$$

Where $\mathbb{I}_V, \mathbb{I}_W$ is the identity map on V and W.

Theorem 3.19
T is invertible if and only if T is both injective and surjective.

Proof. Suppose T is invertible, then $T(T^{-1}w) = w\ \forall w \in W$, so Range $T = W$. Also we know that $T^{-1}(Tv) = v$. Suppose $Tv_1 = Tv_2$, apply the left inverse and we have $T^{-1}(Tv_1) = T^{-1}(Tv_2) \implies v_1 = v_2$. Hence T is injective. Therefore T is bijective.
Now suppose T is bijective. We want to construct T^{-1}



We need to take $w \in W$, there is a $v \in V$ such that $Tv = w$ and such v is unique since T is injective. We declare $T^{-1}w$ to be v. So $T^{-1} \circ T = \mathbb{I}_V$. We compute

$$(T \circ T^{-1})w = T(T^{-1}w) = Tv = w\ \forall w \in W$$

So $T \circ T^{-1} = \mathbb{I}_W$ ■

Definition 3.20

If V, W are vector spaces, such that there exists a invertible linear map $T \in \mathcal{L}(V, W)$ then V, W are isomorphic.

Remark 3.21

Before we proceed, we want to check that T^{-1} is a linear map when $T \in \mathcal{L}(V, W)$ and T^{-1} exists.

Proof. Take $w_1, w_2 \in W, \lambda \in \mathbb{F}$. We compute $T^{-1}(\lambda w_1 + w_2)$. We know that $w_1 = Tv_1$ and $w_2 = Tv_2$. Then we know that $T(\lambda v_1 + \lambda v_2) = \lambda Tv_1 + Tv_2 = \lambda w_1 + w_2$. Subsitute this into T^{-1} and we get

$$T^{-1}(\lambda w_1 + w_2) = T^{-1} \circ T(\lambda v_1 + v_2) = \mathbb{I}(\lambda v_1 + v_2) = \lambda v_1 + v_2 = \lambda T^{-1}w_1 + T^{-1}w_2$$

Hence T^{-1} is linear. ■

Corollary 3.22

$\mathcal{M}$ is actually a bijection between $\mathcal{L}(V, W)$ and $\mathbb{F}^{m,n}$, therefore $\mathcal{L}(V, W)$ is isomorphic to $\mathbb{F}^{m,n}$.

Theorem 3.23

Suppose $T \in \mathcal{L}(V, W)$ is linear and invertible, and let $v_1, v_2, \ldots, v_m$ be a basis for V. Then $Tv_1, Tv_2, \ldots, Tv_n$ is a basis for W.

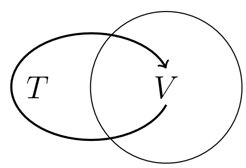
Proof. Suppose $\alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_n Tv_n = 0$. Then $T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = 0$. Since T is injective, this implies $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$. Therefore $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ since $v_1, v_2, \ldots, v_n$ is a basis. Take $w \in W$, then there exists a unique $v \in V$ such that $Tv = w$, and $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ for some $\alpha_1, \alpha_2, \ldots, \alpha_n$, so $Tv = w = \alpha_1 Tv_1 + \alpha_2 Tv_2 + \cdots + \alpha_n Tv_n$, hence span. ∎

Corollary 3.24

dim is invariant under isomorphism.

3.5.1 Linear Operators

We are dealing with a specific case where $\mathcal{L}(V, W)$ is replaced by $\mathcal{L}(V, V)$.



Recall that $\dim V = \dim \text{Null } T + \dim \text{Range } T$. This gives a better test for invertablity if $W = V$.

Theorem 3.25

Let $T \in \mathcal{L}(V, V)$. If V is finite dimensional vector space, then the following are equivalent:

(a) T is injective.

(b) T is surjective.

(c) T is invertible.

Proof.
(a) $\implies$ (c). Trivial by definition.
(b) $\implies$ (c). Suppose T is injective $\xRightarrow[T \text{ being linear}]{}$ Null $T = \{0\}$ $\iff$ $\dim \text{Null } T = 0$. Therefore

$$\dim V = \dim \text{Null } T = \dim \text{Range } T = 0$$

So $\dim V = \dim \text{Range } T$, hence Range $T = V$, so T is surjective.
(c) $\implies$ (a).

$$\dim V = \dim \text{Null } T + \dim \text{Range } T = \dim \text{Null } T + \dim V \implies \dim \text{Null } T = 0 \implies \text{Null } T = \{0\}$$

So T is injective. T is already known to be surjective, so T is bijective, or T is invertable. ∎

Example 3.26

The theorem does not hold for infinite dimension vector spaces, for example:
The differentiation map $T : f(x) \mapsto f'(x)$ is surjective but not invertible over $\mathcal{P}(\mathbb{R})$. From calculus we know that for every $f(x) \in \mathcal{P}(\mathbb{R})$, there exists $g(x) \in \mathcal{P}(\mathbb{R})$ such that $g'(x) = f(x)$, however, we can see that Null $T \neq \{0\}$ as $1 \in$ Null T. Hence T is not injective.
The integration map $T : f(x) \mapsto \int_0^x f(t)dt$ is injective but not invertible over $\mathcal{P}(\mathbb{R})$. We can see that Null $T = \{0\}$, hence T is injective, however, $1 \notin$ Range T.

3.6 Duality

Definition 3.27

Given a vector space V, we define its dual space as $V' = \mathcal{L}(V, \mathbb{F})$.

Remark 3.28

Objects in $\mathcal{L}(V, \mathbb{F})$ are also called linear functional.

Example 3.29

Linear Functional on $\mathbb{R}^3$: $(x_1, x_2, x_3) \mapsto \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 \ \forall \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$.

Definition 3.30

Suppose $v_1, v_2, \dots, v_n$ is a basis of V. The list of linear functionals $\varphi_1, \varphi_2, \dots, \varphi_n \in V'$ such that

$$\varphi_i(v_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

We claim that $\varphi_1, \varphi_2, \dots, \varphi_n$ is the dual basis of V'.

Lemma 3.31

A dual basis is a basis of V'.

Proof. Suppose there exists $\alpha_1, \alpha_2, \dots, \alpha_n$ such that

$$\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \cdots + \alpha_n \varphi_n = 0$$

We compute for v_1

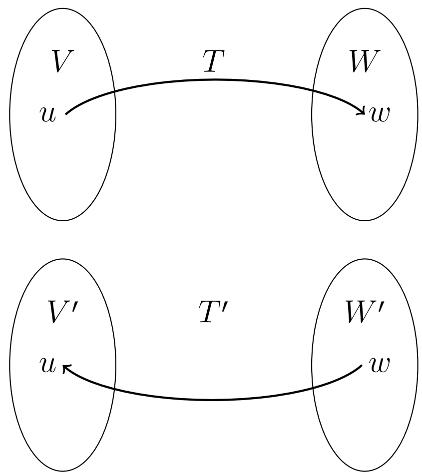
$$(\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \cdots + \alpha_n \varphi_n)(n) = \alpha_1 \cdot 1 + 0 \implies \alpha_1 = 0$$

Similarly, if we plug in an arbitrary v_j

$$(\alpha_1\varphi_1 + \alpha_2\varphi_2 + \cdots + \alpha_n\varphi_n)(v_j) \implies v_j = 0$$

This shows that $\varphi_1, \varphi_2, \ldots, \varphi_n$ is linear independent, and the count is n. So $\varphi_1, \varphi_2, \ldots, \varphi_n$ form a basis for V. ■

3.6.1 Dual Maps



"upside down"

Definition 3.32

Given $T \in \mathcal{L}(V, W)$, we define $T' : \varphi \mapsto \varphi \circ T$. $\phi \in W'$ i.e. $\varphi \in \mathcal{L}(W, \mathbb{F})$. Notice that $\varphi \circ T \in \mathcal{L}(V, \mathbb{F})$.

Example 3.33

Define V, W, φ, T as $\varphi : f \mapsto \int_0^1 f(t)dt, V = \mathcal{P}_3(\mathbb{R}), W = \mathcal{P}_2(\mathbb{R}), T : f \mapsto f'$. What is $\varphi \circ T$?

$$(\varphi \circ T)(f) = \varphi(T(f)) = \varphi(f') = \int_0^1 f'(t)dt = f(1) - f(0)$$

Remark 3.34 (Algebraic Property of dual maps)

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

   Proof. For any $S, T \in \mathcal{L}(V,W), S', T' \in \mathcal{L}(W', V')$, we compute

   $$(S+T)'(\varphi) = \varphi \circ (S+T) = \varphi \circ S + \varphi \circ T = S'(\varphi) + T'(\varphi)$$

   ■

2. $(\lambda S)' = \lambda S'$

   Proof.
   $$(\lambda S)'(\varphi) = \varphi \circ (\lambda S) = \lambda(\varphi \circ S) = \lambda \cdot S'(\varphi)$$

   ■

3. $(ST)' = T'S'$

   Proof. Sanity check: $S \in \mathcal{L}(V,W), T \in \mathcal{L}(U,V)$

   $$(ST)' : \varphi \mapsto \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = T' \circ S'$$

   ■

Definition 3.35 (Annihilators)
Let a set S to be a subset of a vector space V. We can define S^0 as

$$S^0 := \{\varphi \in V' : \varphi(v) = 0 \ \forall v \in S\}$$

Example 3.36
Consider $\mathbb{R}^3$, let $S := \{(1,0,0),(1,1,0)\}$. We know that any $\varphi \in \mathbb{R}^3$ will have the form of $(x_1, x_2, x_3) \mapsto a_1x_1 + a_2x_2 + a_3x_3$ for some constant a_1, a_2, a_3. Plug in $(1,0,0)$ and $(1,1,0)$ and we get

$$a_1 \cdot 1 + a_2 \cdot 0 + a_3 \cdot 0 = 0 \implies a_1 = 0$$

$$a_1 \cdot 1 + a_2 \cdot 1 + a_3 \cdot 0 = 0 \implies a_2 = 0$$

We can see that $S^0 = \{\varphi(x_1, x_2, x_3) = a_3x_3 \text{ for some } a_3 \in \mathbb{R}\}$ and forms a subspace for $\mathbb{R}^{3}$'.

Lemma 3.37
Regardless of the nature of S, S^0 is always a subspace.

Proof.
1. The zero functional is clearly in S^0.
2. Suppose $\varphi \in S^0$, take $\lambda \in \mathbb{F}$, then $(\lambda\varphi)(v) = \lambda \cdot \varphi(v) = 0$ for all $v \in S$. So $\lambda\varphi \in S$.

3. Suppose $\varphi, \psi \in S^0$, then $(\varphi + \psi)(v) = \varphi(v) + \psi(v) = 0$ for all $v \in S$. Therefore $\varphi + \psi \in S^0$. ■

Theorem 3.38

Suppose $S = U$, where U is a subspace of V, then

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

Proof. Consider the inclusion map:

$$i : U \to V : u \to u \ \forall u \in U$$

Take a look at the dual of U: $i' \in \mathcal{L}(V', U')$. Apply Rank-Nullity to the dual map and we can see that $\dim V' = \dim \operatorname{Null} i' + \dim \operatorname{Range} i'$. We also know that

$$\operatorname{Null} i' = \{\varphi \in V' : \varphi \circ i = 0\}$$

Notice that $\varphi \circ i = 0$ as a functional implies

$$(\varphi \circ i)(u) = 0 \implies \varphi(i(u)) = 0 \implies \varphi(u) = 0 \ \forall u \in U$$

Therefore we can see that $\operatorname{Range} i' = \{\varphi \circ i : \varphi \in V'\} = U'$ since any linear functional on U extends to V. I have a clever proof for this but it does not fit in the margin of the page and is left as an exercise for the reader. ■

Theorem 3.39

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

(a) $\operatorname{Null} T' = (\operatorname{Range} T)^0$

(b) $\dim \operatorname{Null} T' = \dim \operatorname{Null} T + \dim W - \dim V$

Proof.

(a) $\varphi \in \operatorname{Null} T' \iff \varphi \circ T = 0 \iff (\varphi \circ T)(v) = 0 \ \forall v \in V \iff \varphi(Tv) = 0 \ \forall v \in V \iff \varphi \in (\operatorname{Range} T)^0$

(b) $\dim \operatorname{Null} T' = \dim(\operatorname{Range} T)^0 = \dim W - \dim \operatorname{Range} T = \dim W - (\dim V - \dim \operatorname{Null} T) = \dim W - \dim V + \dim \operatorname{Null} T$

■

Corollary 3.40

T' is injective if and only if T is surjective.

Theorem 3.41

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

(a) $\dim \text{Range } T' = \dim \text{Range } T$

(b) $\text{Range } T' = (\text{Null } T)^0$

Proof.

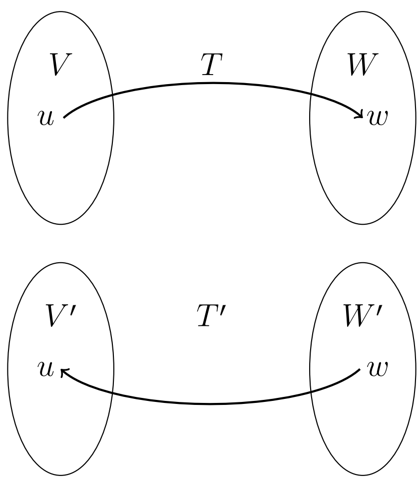
(a) $\dim \text{Range } T' = \dim W' - \dim \text{Null } T' = \dim W - (\dim W - \dim V + \dim \text{Null } T) = \dim V - \dim \text{Null } T = \dim \text{Range } T$

(b) $\psi \in \text{Range } T' \iff \exists \varphi : \varphi \circ T = \psi \iff \varphi \circ T(v) = \psi(V)\ \forall v \in V \iff \varphi(Tv) = \psi(v) \forall v$. So $Tv = 0 \implies \psi(v) = 0$. This shows $\text{Range } T' \subseteq (\text{Null } T)^0$. But $\dim \text{Range } T' = \dim \text{Range } T = \dim V - \dim \text{Null } T = \dim(\text{Null } T)^0$. Hence $\text{Range } T = \text{Null } T$.

■

3.6.2 Matrix Representation of the dual map

Recall that



"upside down"

We also recall that $T' : \varphi \mapsto \varphi \circ T$.

Question 3.42

How do we get $\mathcal{M}(T')$ given $\mathcal{M}(T)$?

Answer 3.43. For $\mathcal{M}(T)$, we need 2 bases $v_1, v_2, \dots, v_n$ for V and $w_1, w_2, \dots, w_m$ for W.
Take $\varphi_1, \varphi_2, \dots, \varphi_n$, a dual basis to $v_1, v_2, \dots, v_n$, it is a basis for V'.
Tale $\psi_1, \psi_2, \dots, \psi_m$, a dual basis to $w_1, w_2, \dots, w_m$, it is a basis for W'.
Given $\mathcal{M}(T)$, we want to construct / understand $\mathcal{M}(T')$with regard to the basis $\psi_1, \psi_2, \dots, \psi_m$ of W' and $\varphi_1, \varphi_2, \dots, \varphi_n$ of V'.
We know $\mathcal{M}(T)$ has m rows n columns, and $\mathcal{M}(T')$ has n rows and m columns.
Suppose $\mathcal{M}(T) = A, \mathcal{M}(T') = C$, we then know

$$Tv_j = \sum_{i=1}^{m} A_{i,j} w_i \ \ \forall j = 1, 2, \dots, n, T'\psi_l = \sum_{l=1}^{n} C_{l,k}\varphi_l \ \ \forall l = 1, 2, \dots, m$$

$$T'\psi_k = \psi_k \circ T \implies (\psi_k \circ T)(v_j) = \psi_k(Tv_j) = \psi_k\left(\sum_{i=1}^{m} A_{i,j} w_j\right) = \sum_{i=1}^{m} A_{i,j}\psi_k(w_i) = \sum_{i=1}^{m} A_{i,j}\delta_{ki} = A_{k,j}$$

$$(T'\psi_k)(v_j) = \left(\sum_{l=1}^{n} C_{l,k}\varphi_l\right)(v_j) = \sum_{l=1}^{n} C_{l,k}\varphi(v_j) = \sum_{l=1}^{n} C_{l,k}\delta_{l,j} = C_{j,k}$$

Notice that $A_{k,j} = C_{j,k} \ \ \forall j, k$.

Conjecture 3.44. So we obtained that

$$M(T') = M(T)^T$$

provided that the basis of V' and W' are chosen to be the dual to the bases of V and W, respectively.

Example 3.45

Let $T : p \mapsto p'$ for $V = \mathcal{P}_3(\mathbb{R})$ with basis $1, x, x^2, X^3$, and $W = \mathcal{P}_2(\mathbb{R})$ with $1, x, x^2$.
We can see that the dual basis for $1, x, x^2, x^3$ is

$$\varphi_0 : p \mapsto p(0), \varphi_1 : p \mapsto p'(0), \varphi_2 : p \mapsto \frac{p''(0)}{2}, \varphi_3 : p \mapsto \frac{p'''(0)}{3!}$$

Dual basis for $1, x, x^2$ is

$$\psi_0 : p \mapsto p(0), \psi_1 : p \mapsto p'(0), \psi_2 : p \mapsto \frac{p''(0)}{2},$$

Notice that

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \text{ and } \mathcal{M}(T') = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

4 Polynomials

Recall we call consider polynomials over $\mathbb{C}$ or $\mathbb{R}$.

Theorem 4.1

For any $z_1, z_2 \in \mathbb{C}$, we define $|z| = \sqrt{a^2 + b^2}$, we know that

1. $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
2. $|z_1 \cdot z_2| \leq |z_1| + |z_2|$

Proof. Left as an execise. ■

4.1 Axler's Recap on Polynomial

Theorem 4.2

Suppose $p(x) \in \mathcal{P}(\mathbb{F})$, is identically zero. Then all of its coefficient must be 0.

Proof. If $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, then $a_j = \frac{p^{(j0}(0)}{j!}$, If $p(x) \equiv 0$, then $p^{(j)}(x) = 0$, so $a_j = \frac{0}{j!} = 0 \forall j$ ■

Corollary 4.3

Suppose $p(x) \equiv q(x)$ for $p, q \in \mathcal{P}(\mathbb{F})$, then all coefficients of p are the same as all coefficients of q.

4.2 Zero of polynomials and their algebraic manifestations

Algorithm 4.4 (Euclidean Algorithm for polynomials)

Given $p(x), s(x)$, without the loss of generality, $\deg p(x) > \deg s(x)$, otherwise it's boring; we can always find $q(x), r(x)$ such that $p(x) = s(x)q(x) + r(x)$, where $\deg r(x) < \deg s(x)$.

Corollary 4.5

$p(a) \iff p(x) = (x-a)q(x)$ for some $a \in \mathbb{F}$.

Proof. If $p(a) = (x-a)q(x)$, then $p(a) = 0 \cdot q(a) = 0$.
Conversely, suppose $p(a) = 0$, by division algorithm we have $p(x) = (x-a)q(x) + r(x)$, where $\deg r \leq \deg(x-a)$, therefore $r(x) = c$ for some $c \in \mathbb{F}$. Plug in a and we get $(a-a)q(a) + c = 0 \implies 0 + c = 0 \implies c = 0$. Therefore $p(x) = (r-a)q(x)$. ■

Theorem 4.6

Let $p(x)$ be a nonzero polynomial with coefficients in $\mathbb{F}$ have degree n. Then p has at most n zeros in $\mathbb{F}$.

Proof.
Base case: $\deg p = 1$, i.e. $p(x) = a_1 x + a_0$ for some $a_1 \in \mathbb{F}^\times, a_0 \in \mathbb{F}$. Then $p\left(\frac{-b}{a}\right) = 0$, so p has exactly one zero.
Inductive Hypothesis: Suppose the statement is true for all polynomials for all polynomials of degree less than m.
Inductive Step: Take $p(x)$ to be a degree m polynomial. If p has no zeros in $\mathbb{F}$, we are done. If p has a zero, by corollary we have $p(x) = (x-a)q(x)$, where $\deg q = m-1$. So the inductive hypothesis applies and q at most $n-1$ distinct zeros in $\mathbb{F}$. ∎

Theorem 4.7 (Fundamental Theorem of Algebra)

Every nonconstant polynomial with complex coefficients has a zero.

Proof with "Black Box" from Complex Analysis.
Assume $\deg p \geq 1$. Assume that $p(a) \neq 0\ \forall a \in \mathbb{C}$. Consider the function $\frac{1}{p(x)}$ is well-defined $\forall x \in \mathbb{C}$ and is analytic in $\mathbb{C}$, more over $\lim_{|z|\to\infty} \frac{1}{p(z)} = 0$. We know that

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ &= x^n \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right) \\ \frac{1}{p(x)} &= \frac{1}{x^n \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right)} \end{aligned}$$

As $|x| \to \infty$, $\dfrac{1}{x^n} \to 0$. Since $\left|\dfrac{1}{x^n}\right| = \dfrac{1}{|x|^n} \to 0$. But $\dfrac{a_0}{x^n} + \dfrac{a_1}{x^{n-1}} + \cdots + a_n \to a_n \neq 0$. Hence $\dfrac{1}{p(x)} \to 0$ as $|x| \to -\infty$.
By Louisville's theorem, any analytic function with this property has to be constant. But $\frac{1}{p(x)}$ is non-constant, so p must have at least 1 zero in $\mathbb{C}$. ∎

Corollary 4.8

Any polynomial $p(x)$ with coefficients in $\mathbb{C}$ factors as follows

$$p(x) = c(x-a_1)(x-a_2)\cdots(x-a_m), c \neq 0$$

Proof. By Induction it's clear for degree 1 and if $\deg p = m$ then factor $p(x) = (x-a)q(x)$ and repeat the process for q. ∎

Question 4.9

What happens over $\mathbb{R}$?

Theorem 4.10

If $p(x)$ has coefficient in $\mathbb{R}$, and $c \in \mathbb{C}$ is a zero of p, then $\bar{c}$ is also a zero of p.

Proof. $p(c) = 0$ means

$$a_0 + a_1 c + a_2 c^2 + \cdots + a_n c^n = 0$$

We then can see

$$\begin{aligned}
\overline{a_0} + \overline{a_1 c} + \overline{a_2 c^2} + \cdots + \overline{a_n c^n} &= \overline{0} = 0 \\
\overline{a_0} + \overline{a_1}\overline{c} + \overline{a_2}\overline{c^2} + \cdots + \overline{a_n}\overline{c^n} &= 0 \\
a_0 + a_1 \overline{c} + a_2 \overline{c^2} + \cdots + a_n \overline{c^n} &= 0
\end{aligned}$$

Hence $p(\bar{c}) = 0$ as well. ∎

So over $\mathbb{C}$, a polynomial with real coefficient factors as follows

$$p(x) = (x - a_1)(x - a_2)\cdots(x - a_n)(x - \lambda_1)(x - \overline{\lambda_1})(x - \lambda_2)(x - \overline{\lambda_2})\cdots(x - \lambda_m)(x - \overline{\lambda_m})$$

For some $c \in \mathbb{R}, a_1, a_2, \ldots, a_n \in \mathbb{R}, \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{C}$.
To translate this into a factorization over $\mathbb{R}$, we can see that $x^2 - (\lambda + \bar{\lambda}) + |\lambda|^2$. These are quadratic with $\Delta < 0$. Indeed,

$$(\lambda + \bar{\lambda})^2 - 4|\lambda|^2 = \lambda^2 - 2|\lambda|^2 + \bar{\lambda}^2 = 2\mathrm{Re}\lambda^2 - 2|\lambda|^2$$

Notice that $\mathrm{Re}\lambda^2 \le |\lambda|^2$ and $\mathrm{Re}\lambda^2 = |\lambda|^2$ iff $\lambda \in \mathbb{R}$, therefore $\Delta < 0$.

Question 4.11

Why do we study polynomials?

Answer 4.12.

1. We will form polynomials in linear operators
2. We will associate special polynomials with linear operators.

Remark 4.13

An operator has he same co-domain as its domain.

5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.1 Invariant Subspaces

Definition 5.1

Let $T \in \mathcal{L}(V,V)$ on a vector space $V \neq \{0\}$. A subspace $U \subseteq V$ is called an invariant subspace is invariant under T if $Tu \in U$ $\forall u \in U$.

Example 5.2

For any $T \in \mathcal{L}(V,V)$, the following subspaces are invariant:

1. $\{0\}$
2. V
3. Null $T = \{v \in V : Tv = 0\}$
   If $Tv \in$ Null T, then $Tv = 0 \in$ Null T.
4. Range $T = \{w \in W : w = Tv \text{ for some } v \in V\}$
   So $Tw \in$ Range T.

Question 5.3

What are 1-dimensional invariant subspaces?

Answer 5.4. Then $U = \text{span}(u)$ for some $u \neq 0$. Invariant means $Tu = \lambda u$ for some $\lambda \in \mathbb{F}$, where u is the eigenvector of T and λ is the eigenvalues.

Remark 5.5

$u \neq 0$ if u is a eigenvector is T. $\lambda = 0$ is possible.

Proposition 5.6

Let T be a linear operator in V, then the following are equivalent

1. λ is a eigenvalue of T.
2. $T - \lambda\mathbb{I}$ is not invertible.
3. $T - \lambda\mathbb{I}$ is not injective.
4. $T - \lambda\mathbb{I}$ is not surjective.

We have already proven that statement $2, 3, 4$ are logically equivalent.

Theorem 5.7

Suppose $v_1, v_2, \ldots, v_m$ are eigenvectors of $T \in \mathcal{L}(V)$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ will be linearly independent.

Proof. Suppose $v_1, v_2, \ldots, v_m$ are linearly independent. By linear dependence lemma, we find a the minimum index $k \leq m$ such that $v_k \in \operatorname{span}(v_1, v_2, \ldots, v_{k-1})$. i.e.

$$v_k = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k-1} v_{k-1} \tag{1}$$

Apply linear transformation on both sides

$$Tv_k = T\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k-1} v_{k-1} \tag{2}$$

$$\lambda v_k = \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \cdots + \alpha_n \lambda_n v_n \tag{3}$$

We multiply by equation 1 by λ_m and subtract by from 3 and we get

$$0 = \alpha_1(\lambda_1 - \lambda_k)v_1 + \alpha_2(\lambda_2 - \lambda_k)v_2 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}$$

A contradiction since k is not the minimum index with the property chosen above. Therefore the list $v_1, v_2, \ldots, v_m$ must be linearly independent. ■

Corollary 5.8

An operator $T \in \mathcal{L}(V)$ has at most $\boxed{\dim V}$ distinct eigenvalues.

5.1.1 Restriction Operators

Definition 5.9

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. Then the restriction operator $T|_U \in \mathcal{L}(U)$ is deifned by $T|_U(u) = Tu$ for all $u \in U$.

5.2 Eigenvectors and Upper-Triangular Matrices

5.2.1 Polynomials in T

Definition 5.10

Suppose $T \in \mathcal{L}(V)$, then T^k is defined as

$$T^k := \underbrace{k \circ k \circ \cdots \circ k}_{k \text{ times}}$$

Notice that $T^0 = \mathbb{I}, T^1 = T$.

Definition 5.11

If $p(x) = a_0 + a_1x + \cdots + a_nx^n$, then we can define $p(T)$ as $a_o\mathbb{I} + a_1T + a_2T + \cdots + a_nT^n$.

Example 5.12

Let $V := \mathcal{P}(\mathbb{R}), S : p \mapsto 3p'' + 2p' + p, D : p \mapsto p'$. We can see that S can be expressed as $S = D^0 + 2D + 3D^2$. Therefore

$$\mathcal{M}(S) = 3\mathcal{M}^2(D) + 2\mathcal{M}(D) + M(\mathbb{I})$$

we need to have to take the same basis for inputs and output when forming $\mathcal{M}(\cdot)$. Let's use our favorite basis $1, x, x^2, x^3$. We then can see

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathcal{M}(S) = \begin{bmatrix} 1 & 2 & 6 & 0 \\ 0 & 1 & 4 & 18 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Question 5.13

What is the best matrix representation for an operator?

Question 5.14

What information about eigenvalues/eigenvectors can be read off from a matrix representation?

Theorem 5.15

Suppose $T \in \mathcal{L}(V)$ and $v_1, v_2, \ldots, v_n$ is a basis of V. Then the following are logically equivalent:

1. $\mathcal{M}(T)$ is upper triangular.
2. $Tv_j \in \operatorname{span}(v_1, v_2, \ldots, v_j)\ \forall j = 1, 2, \ldots, n$.
3. $\operatorname{span}(v_1, v_2, \ldots, v_j)$ is invariant under $T\ \forall j = 1, 2, \ldots, n$.

Proof. 1) $\implies$ 2)

$$\begin{bmatrix} * & * & * & * & \cdots & * \\ & * & * & * & \cdots & * \\ & & * & * & \cdots & * \\ & & & * & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & * \end{bmatrix}$$

We can see that 2) holds true by inspection.
2) $\implies$ 3) Consider Tv_h for $h \leq j$, by 2) we have $Tv_k \in \operatorname{span}(v_1, v_2, \ldots, v_h) \subseteq \operatorname{span}(v_1, v_2, \ldots, v_j)$. So $\operatorname{span}(v_1, v_2, \ldots, v_j)$ is invariant under T.
3) $\implies$ 2) Consider Tv_j, by 3) it is a linear combination of $v_1, v_2, \ldots, v_j$ because $Tv_j \in \operatorname{span}(v_1, v_2, \ldots, v_j)$ so $\mathcal{M}(T)(i,j) = 0$ if $i > j$. ∎

Question 5.16

What about conditions for lower-triangular matrices?

Lemma 5.17

Over $\mathbb{C}$, every linear operator has at least one eigenvalue.

Proof. Take $v \in V \setminus \{0\}$, and consider the list $v, Tv, T^2v, \ldots, T^nv$ where $n = \dim V$. There is a nontrivial linear combination of these vectors which is 0. Suppose the equation

$$a_0 v_1 + a_1 Tv + a_2 T^2 v + \cdots + a_n T^n v = 0$$

i.e. $p(T)v = 0$ for nonconstant $p(x) \coloneqq a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$. By the fundamental theorem of algebra p splits into linear factors over $\mathbb{C}$.

$$p(x) = c(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_m)$$

for some $m \leq n$. Therefore

$$p(T)v = c(T - \lambda_1 \mathbb{I})(T - \lambda_2 \mathbb{I})\cdots(T - \lambda_m \mathbb{I})$$

Therefore at least one of these factors is not injective. This shows that T has at least 1 eigenvalue. ∎

Theorem 5.18

For any $T \in \mathcal{L}(V)$, V is finite dimensional vector space over $\mathbb{C}$, there exists its matrix representation $\mathcal{M}(T)$ which is upper-triangular.

Proof. We can induct on the dimension of V. *Base Step.* $n = 1$ is trivially true.
Inductive Hypothesis. Suppose Theorem holds for all vector spaces of dimension less than $\dim V$.
Inductive Step. Consider $\lambda \in \mathbb{C}$ an eigenvalue of T by lemma. We can define

$$U := \text{Range } (T - \lambda\mathbb{I})$$

U is a subspace of V. By the characterization of eigenvalues, $T - \lambda\mathbb{I}$ is not surjective, hence Range $T - \lambda\mathbb{I} \notin V$, hence $\dim \text{Range } (T - \lambda\mathbb{I}) < \dim V$. We want to show that U is invariant under T. Suppose $v \in U$, then

$$Tv = \underbrace{(T - \lambda\mathbb{I})v}_{\in U} + \underbrace{\lambda v}_{\in U}$$

therefore we know that U is invariant under T. Consider

$$T|_U \in \mathcal{L}(U) : (T|_U)(v) := Tv \forall v \in U$$

If $U \neq \{0\}$, then there is a basis $u_1, u_2, \ldots, u_m$ of U ($m < n$) such that the matrix representation of T/U with respect to $u_1, u_2, \ldots, u_m$ is upper triangular by the inductive hypothesis. Extend $u_1, u_2, \ldots, u_m$ to a basis of V, $u_1, u_2, \ldots, u_m, v_1, v_2, \ldots, v_k$. We compute

$$Tv_j = \underbrace{(T - \lambda\mathbb{I})v_j}_{\in U = \text{span}(u_1, u_2, \ldots, u_m)} + \lambda v_j$$

We also know that $Tu_l \in \text{span}(u_1, u_2, \ldots, u_{l-1})$. We can see the matrix representation and hence we are done

$$\begin{array}{cc} & \\ & \\ & \\ & \\ m & \\ & \\ & \end{array}\left[\begin{array}{cccc|ccc} * & * & \cdots & * & * & * & * \\ 0 & * & \cdots & * & * & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & * & * & * \\ \hline 0 & 0 & \cdots & 0 & \lambda & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \lambda & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \lambda \end{array}\right]$$

■

Question 5.19

What about eigenvalues of a upper-triangular matrix?

Theorem 5.20

An upper triangular matrix is invertible if and only if all its diagonal entries are nonzero.

Proof. Suppose all diagonal entries are nonzero. Prove surjectivity.

$$\begin{aligned} Tv_1 &= A_{1,1}v_1, A_{1,1} \neq 0 \implies v_1 \in \text{Range } T \\ Tv_2 &= A_{1,2}v_1 + A_{2,2}v_2, A_{2,2} \neq 0 \implies v_2 \in \text{Range } T \\ &\vdots \qquad\qquad\qquad\qquad \implies \\ Tv_n &= A_{1,n}v_1 + A_{2,n}v_2 + \cdots + A_{n,n}v_n, A_{n,n} \neq 0 \implies v_n \in \text{Range } T \end{aligned}$$

Therefore Range $T = V$, so T is surjective, hence T is invertible. Suppose at least one one diagonal entry is 0 we want to show that T is not invertible. Say $A_{j,j} = 0$ for some j and upper triangular matrix A. If $j = 1$, then $v_1 \in T$, hence T is not invertible, and we are done. If $j > 1$, consider $U := \text{span}(v_1, v_2, \ldots, v_j)$. T maps U to $\text{span}(v_1, v_2, \ldots, v_{j-1})$. This shows T/U us not surjective, then we know that T/U is not injective and there exists $u \in U$ such that $u \in T/U \implies u \in T$. Therefore T is not injective. Hence T is not invertible. ∎

Corollary 5.21

An upper triangular matrix / operator in upper triangular form has the diagonal elements / entries as its eigenvalues.

Example 5.22

The matrix

$$A = \begin{bmatrix} 5 & * & * & * & * \\ 0 & 9 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 8 & * \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix}$$

has eigenvalue $1, 5, 9, 8, 10$.

Example 5.23

$T : \mathcal{P}_n(\mathbb{R}) \to \mathcal{P}_n(\mathbb{R}) : p \mapsto 3p'' - 5'p' + 7p$ has eigenvalues $3, -5, 7$.

5.3 Eigenspaces and Diagonal Matrices

Definition 5.24

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

$$E(\lambda, T) := \{v \in V : Tv = \lambda v\} = \text{Null } (T - \lambda I)$$

Definition 5.25

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

Theorem 5.26

For $T \in \mathcal{L}(V)$, where V is a finite dimensional vector space, then the following are equivalent

1. $\mathcal{M}(T)$ is a diagonal matrix.
2. the corresponding basis for V consists of eigenvalue of T.
3. $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$ where $\dim U_j = 1$ and U_j is invariant under T for all j.
4. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, where $V/W_l = \lambda_l \mathbb{I}$ for all l and W_l is invariant under T.
5. $\dim V = \dim W_1 + \dim W_2 + \cdots + W_k$, where $W_e = (T - \lambda_e \mathbb{I})$.

Proof. Refer to Axler Page 157. ■

6 Inner Product Spaces

Motivation

Definition 6.1
In $\mathbb{R}^n$, the dot product of $\vec{x}$ and $\vec{y}$ is defined by

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

for $\vec{x} = (x_1, x_2, \ldots, x_n), \vec{y} = (y_1, y_2, \ldots, y_n)$.

6.1 Inner Product and Norms

Settings

V is a vector space over $\mathbb{F}$, we can define the following mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$.

Definition 6.2
$\langle \cdot, \cdot \rangle$ is called an inner product if it satisfying the following rules:

1. (additivity in the first slot) $\langle \vec{v} + \vec{u}, \vec{w} \rangle = \langle \vec{v} + \vec{w} \rangle + \langle \vec{u}, \vec{w} \rangle, \ \forall \vec{v}, \vec{u}, \vec{w} \in V$
2. (homogeneity in the first slot) $\langle \lambda \vec{v}, \vec{w} \rangle = \lambda \langle \vec{v} + \vec{w} \rangle, \ \forall \vec{v} \vec{u}, \vec{w} \in V, \lambda \in \mathbb{F}$
3. (conjugate symmetry) $\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}, \ \forall \vec{v}, \vec{w} \in V$
4. (positivity) $\langle \vec{v}, \vec{v} \rangle \geq 0, \ \forall \vec{v} \in V$
5. (definiteness) $\langle \vec{v}, \vec{v} \rangle = 0$ iff $\vec{v} = \vec{0}$.

Question 6.3
What about linearity in the second slot?

Answer 6.4. We can compute

$$\langle \vec{v}, \vec{u} + \vec{w} \rangle = \overline{\langle \vec{u} + \vec{w}, \vec{v} \rangle} = \overline{\langle \vec{u}, \vec{v} \rangle + \rangle \vec{w}, \vec{v}} = \overline{\langle \vec{u}, \vec{v} \rangle} + \overline{\langle \vec{w}, \vec{v} \rangle} = \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{w} \rangle$$

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

Not quite. ☹

Remark 6.5
If $\vec{v} \in V$ is fixed then the function $\langle \cdot, \vec{v} \rangle : \vec{u} \mapsto \langle \vec{u}, \vec{v} \rangle$ is a function functional.

Example 6.6

On $\mathbb{R}^n$, we could use any function of the type

$$c_1x_1y_1 + c_2x_2y_2 + \cdots + c_nx_ny_n$$

where all $c_j \in \mathbb{R}^+$.

Remark 6.7 (Generalization to $\mathbb{C}^n$)

The inner product of this form of the standard product to $\mathbb{C}^n$ can be defined as

$$\langle \vec{x}, \vec{y} \rangle = x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_n\overline{y}_n$$

Remark 6.8 (Generalization to any function space)

$$\langle f, g \rangle := \int_D f(t)\overline{g(t)}dt$$

or generally

$$\langle f, g \rangle := \int_D f(t)\overline{g(t)}w(t)dt$$

where $w(t)$ is the positive weight function. e.g. if $V = \mathcal{P}(\mathbb{R})$, or $V = \mathcal{P}(\mathbb{C})$, then

$$\langle f, g \rangle := \int_0^\infty f(t)\overline{g(t)}e^{-t}dt$$

Definition 6.9

For $v \in V$, the (Euclidean) Norm is defined as

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

Theorem 6.10

(Properties of Norms)

1. $||\lambda v|| = |\lambda|\ ||v||\ \forall v \in V, \forall \lambda \in \mathbb{F}$
2. $||v|| > 0$ for all $v \in V$
3. $||v|| = 0$ if and only if $v = 0$

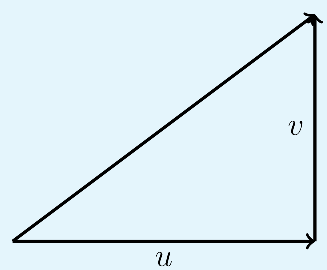
Definition 6.11

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

Definition 6.12

For $u, v \in V$, we say u and v is orthogonal if $\langle u, v\rangle = 0$

Theorem 6.13 (Pythagorean Theorem)



$$||u+v||^2 = ||u||^2 + ||v||^2 \iff \langle u+v, u+v\rangle = \langle u,u\rangle + \langle v,v\rangle$$

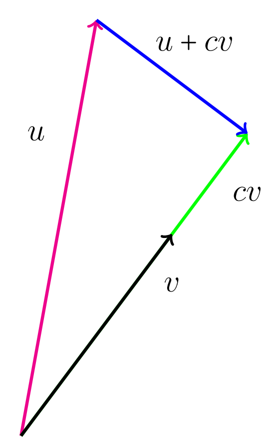
Proof. We compute

$$\langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = \langle u,u\rangle + 0 + 0 + \langle v,v\rangle = ||u|| + ||v||$$

■

Obeservation

Given $u, v \in V$ such that $v \neq 0$, we want to modify u such that $u + cv$ is orthogonal to v. We know that $\langle v + cv, v\rangle = 0$, solve for c gives $c = \dfrac{-\langle u, v\rangle}{\langle v, v\rangle}$.



An orthogonal decompostion

Theorem 6.14 (Cauchy-Schwarz Inequality)

For any $u, v \in V$ where V is a inner product space, the following holds

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

Proof. Given $u, v \in V$, we can assume without the loss of generality that $v \neq 0$. So we can consider vectors $u + cv$ and v that are orthogonal for the choice that

$$c := \frac{-\langle u, v\rangle}{\langle v, v\rangle}$$

By Pathgrathrean theorem, $||u + cv||^2 + ||cv||^2 = ||u||^2$. But $||cv||^2 = |c|^2||v||^2$ and recall

$$c = \frac{-\langle u, v\rangle}{\langle v, v\rangle}, \text{ so } c^2 = \frac{|\langle u, v\rangle|^2}{\langle v, v\rangle^2} = \frac{|\langle u, v\rangle|^2}{||v||^4}, \text{ therefore } ||cv||^2 = \frac{|\langle u, v\rangle|^2}{||v||^4}||v||^2 = \frac{|\langle u, v\rangle|^2}{||v||^2}$$

So by dropping $||u + cv||^2 > 0$, we obtain $||cv||^2 \leq ||u||$, i.e,

$$\frac{|\langle u, v\rangle|^2}{||v||^2} \leq ||u||^2 \implies |\langle u, v\rangle|^2 \leq ||u|^2 \cdot ||v||^2 \implies |\langle u, v\rangle|^2 \leq ||u|| \cdot ||v||$$

■

Last updated: April 2, 2019