Linear Algebra

Hrishikesh Belagali

April 22, 2025

1 Vector Spaces

1.1 $\mathbb{R}^n$ and $\mathbb{C}^n$

Definition: Complex Numbers

A complex number is an ordered pair of the form (a, b) denoted as $a + bi$. $i = \sqrt{-1}$

$$\mathbb{C} = \{(a, b) : a, b \in \mathbb{R}\}$$

$(+, \cdot)$ are defined on complex numbers as follows-

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

Any real number $a \in \mathbb{R}$ can be defined as a complex number of the form $a + 0i$. Hence, it is clear that $\mathbb{R} \subset \mathbb{C}$.

The operators of complex numbers are well-defined.
For $\lambda, \alpha, \beta \in \mathbb{C}$,

commutativity

$$\alpha + \beta = \beta + \alpha$$

$$\alpha \cdot \beta = \beta \cdot \alpha$$

associativity

$$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$$

$$(\alpha\beta)\lambda = \alpha(\beta\lambda)$$

identities

$$\lambda + 0 = \lambda \text{ and } \lambda * (1) = \lambda$$

additive inverse

$$\text{For every } \alpha \in \mathbb{C}, \text{ there exists a unique } \beta : \alpha + \beta = 0$$

The additive inverse of α is denoted by $-\alpha$.
multiplicative inverse

$$\text{For every } \alpha \in \mathbb{C}\backslash\{0\}, \text{ there exists a unqiue } \beta : \alpha \cdot \beta = 1$$

The multiplicative inverse of α is denoted by α^{-1}.
distributivity

$$\lambda(\alpha + \beta) = \lambda \cdot \alpha + \lambda \cdot \beta$$

Example: Proof of commutativity

To show that $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbb{C}$ Let $\alpha = a + bi$ and $\beta = c + di$, $a, b, c, d \in \mathbb{R}$

$$\alpha\beta = ac - bd + (ad + bc)i$$

$$\beta\alpha = ca - db + (da + cb)i$$

We know that multiplication is commutative over real numbers.

$$\therefore \alpha\beta = ca - db + (da + cb)i = ac - bd + (ad + bc)i = \beta\alpha$$

Hence, the commutativity of complex number multiplication is proven.

Additive and multiplicative inverses of complex numbers can be defined. By doing so, we can define subtraction and division.

additive inverse

$$\text{For every } \alpha \in \mathbb{C}, \text{ there exists an additive inverse } \beta : \alpha + \beta = 0$$

It can be proven that $\beta = -\alpha$

$$\alpha + (-\alpha) = (1 - 1)\alpha = 0 \cdot \alpha = 0$$

subtraction

Hence, the subtraction of two complex numbers α and β is defined as the addition of α and the additive inverse of β.

$$\alpha - \beta = \alpha + (-\beta)$$

multiplicative inverse

$$\text{For every } \alpha \neq 0 \in \mathbb{C}, \text{ there exists a unique multiplicative inverse } \beta : \alpha\beta = 1$$

It can be proven that $\beta = \frac{1}{\alpha}$.

$$\alpha\beta = \alpha\frac{1}{\alpha} = 1$$

division

Hence, the division of two complex numbers α and β is defined as the multiplication of α and the multiplicative inverse of β.

Definition: Field

Any set X is defined with closed operations $(+, .)$ and follows the above-mentioned axioms, $X, +, .$ is called a **field**. An arbitrary field is denoted by $\mathbb{F}$. A field obeys all the abovementioned properties, as $\mathbb{C}$ is a field itself.

Definition: Lists

A list (also called an **n-tuple**) is a collection of objects in a particular order. Two lists are equal if and only if they have the same size and the same elements in the same order. Lists are commonly written in the following manner-

$$\mathbf{v} = (v_1, v_2, v_3 \cdots v_n)$$

Definition: $\mathbb{F}^n$

$\mathbb{F}^n$ is defined as follows-

$$\mathbb{F}^n = \{(x_1, x_2, x_3 \cdots x_n) : x_k \in \mathbb{F} \text{ for } k = 1, \cdots n\}$$

Definition: Vector Space

A vector space $\mathbf{V}$, $+ : \mathbf{V} \times \mathbf{V} \mapsto \mathbf{V}$, $\cdot : \mathbb{F} \times \mathbf{V} \mapsto \mathbf{V}$ is defined as a collection of elements closed under the operations $+$ and $\cdot$, adhering to the following axioms-

- **commutativity of addition:** $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \ \forall \ \mathbf{u}, \mathbf{v} \in \mathbf{V}$
- **associativity of addition:** $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \ \forall \ \mathbf{v}, \mathbf{u}, \mathbf{w} \in \mathbf{V}$
- **additive identity:** $\mathbf{v} + \mathbf{0} = \mathbf{v} \ \forall \ \mathbf{v}, \mathbf{0} \in \mathbf{V}$
- **additive inverse:** $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} \ \forall \ \mathbf{v} \in \mathbf{V}$
- **associativity of scalar multiplication:** $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v}) \ \forall \ \alpha, \beta \in \mathbb{F}, \mathbf{v} \in \mathbf{V}$
- **multiplicative identity:** $1\mathbf{v} = \mathbf{v} \ \forall \ \mathbf{v} \in \mathbf{V}$
- **distributivity**:

$$(\alpha + \beta)(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w} + \beta\mathbf{v} + \beta\mathbf{w}$$

Example: $\mathbf{V} = \mathbb{R}^n$

This space is an n-dimensional real-valued vector space of the form

$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_n \end{bmatrix}, x_1, x_2, \cdots x_n \in \mathbb{R}$$

$\mathbf{V}$ is closed under scalar multiplication and addition.

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \cdots \\ x_n + y_n \end{bmatrix} \in \mathbf{V}$$

$$\alpha \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \cdots \\ \alpha x_n \end{bmatrix} \in \mathbf{V}$$

commutativity

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \cdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} y_1 + x_1 \\ y_2 + x_2 \\ \cdots \\ y_n + x_n \end{bmatrix} = \mathbf{y} + \mathbf{x}$$

associativity

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \left(\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} \right) + \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \cdots \\ x_n + y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 + z_1 \\ x_2 + y_2 + z_2 \\ \cdots \\ x_n + y_n + z_n \end{bmatrix} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

additive identity For any $\mathbf{x} \in \mathbb{R}$,

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + 0 \\ x_2 + 0 \\ \cdots \\ x_n + 0 \end{bmatrix} = \mathbf{x}$$

Example: $V = \mathcal{P}_n$

Let $n \in \mathbb{Z}_{>0}$.

$$V = \mathcal{P}_n = \{p : \mathbb{R} \to \mathbb{R} : \exists a_0, a_1 \cdots a_n \in \mathbb{R} \text{ s.t } p(x) = a_0 + a_1 x + \cdots a_n x^n\}$$

This forms the set of polynomials degree less than or equal to n. It is easy to show that the set is closed under addition -

$$(p+q)(x) = (a_0 + b_0) + (a_1 + b_1)x \cdots (a_n + b_n)x^n, p, q \in \mathcal{P}_n$$

For $\alpha \in \mathbb{R}, p \in \mathcal{P}_n$,

$$\alpha p = (\alpha p)(x) = \alpha a_0 + \alpha a_1 x + \cdots \alpha a_n x^n$$

Hence, the set is closed under scalar multiplication.

Proof. $\mathcal{P}_n$ is a vector space.
We have already shown that the set is closed under its definition of addition and scalar multiplication.

additive commutativity:
Let $p, q \in \mathcal{P}_n$

$$p(x) = a_0 + a_1 x + \cdots a_n x^n$$

$$q(x) = b_0 + b_1 x + \cdots b_n x^n$$

$$(p+q)(x) = a_0 + b_0 + (a_1 + b_1)x + \cdots (a_n + b_n)x^n$$

Addition is commutative over the field of real numbers.

$$= (b_0 + a_0) + (b_1 + a_1)x + \cdots (b_n + a_n)x^n = (q+p)(x)$$

Additive associativity follows from the associativity of real numbers.

additive identity:
Define $p_0(x) = 0 + 0x + \cdots 0x^n$ Clearly, $p_0(x)$ belongs to $\mathcal{P}_n$ For some $p \in \mathcal{P}_n$,

$$(p + p_0)(x) = (a_0 + 0) + (a_1 + 0) + \cdots (a_n + 0)x^n$$

$$= p(x)$$

Hence, the zero vector is $0 + 0x \cdots 0x^n$.

additive inverse:
Let $p(x) = a_0 + a_1 x + \cdots a_n x^n$. Let $q(x) = -a_0 - a_1 x - \cdots - a_n x^n$. $p, q \in \mathcal{P}_n$

$$(p+q)(x) = (a_0 - a_0) + (a_1 - a_1)x + \cdots (a_n - a_n)x^n = \mathbf{0}$$

Hence, we have shown that for any arbitrary element of $\mathcal{P}_n$, there exists an additive inverse whose addition gives the additive identity.

scalar distributivity:
Let $\alpha, \beta \in \mathbb{R}, p \in \mathcal{P}_n, p(x) = a_0 + a_1 x + \cdots a_n x^n$ Then,

$$((\alpha + \beta)p)(x) = (\alpha + \beta)a_0 + (\alpha + \beta)a_1 x + \cdots (\alpha + \beta)a_n x^n$$

Using the distributivity of real numbers,

$$= \alpha a_0 + \beta a_0 + \alpha a_1 x + \beta a_1 x + \cdots + \alpha a_n x^n + \beta a_n x^n$$

$$= (\alpha p)(x) + (\beta p)(x)$$

Hence,

$$((\alpha + \beta)p)(x) = (\alpha p)(x) + (\beta p)(x)$$

Example: $V = \mathcal{P}_n^{=}$

Consider the polynomial space where the degree **must** be n. This does not form a vector space because addition and scalar multiplication are not closed.

Consider $p(x) = x^2 + 1$ and $q(x) = -x^2$. The element $(p+q)(x) = 1$ which is not a member of $\mathcal{P}_n^{=}$

Example

Let $V = (0, \infty) \subset \mathbb{R}$. The operations are defined by $(\oplus, \otimes)$ which are addition and scalar multiplication respectively. Let $\mathbf{v}, \mathbf{w} \in V$

$$\mathbf{v} \oplus \mathbf{w} = \mathbf{v} \cdot \mathbf{w}$$

$$\alpha \mathbf{v} = \mathbf{v}^{\alpha}$$

The additive identity is 1 and the scalar multiplicative identity is 1.

Example

$$V = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} x_1, x_2 \in \mathbb{R} \right\}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + 2y_2 \end{bmatrix}$$

$$\alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix}$$

It lacks additive commutativity and is hence, not a vector space.

Some fundamental results:

Proposition: The additive identity is unique.

Proof. Let V be a vector space with two additive identities 0 and $0'$. Then,

$$0 = 0 + 0' = 0' + 0 = 0'$$

Hence, proven.

Proposition: The additive inverse is unique.

Proof. Let V be a vector space. For each vector w, let there be two additive inverses v and v'.

$$v = v + 0 = v + (v' + w) = (v + w) + v' = v'$$

Hence, $v = v'$.

Proposition: $0v = 0$

Proof.

$$0v = (0 + 0)v = 0v + 0v$$

Adding the additive inverse of $0v$ to both sides,

$$0v + (-0v) = (0v + 0v) + (-0v)$$

Using the associativity of addition,

$$0 = 0v + (0v + (-0v))$$

$$0 = 0v$$

Hence, proven.

Proposition: $(-1)v = -v$.

Proof.

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

Clearly, $(-1)v$, when added to v, gives us the additive identity. Hence, $(-1)v = -v$.

1.2 Subspace

Definition: Subspace

If $\mathbf{V}$ is a vector space, the space $\mathbf{U}$ is called a subspace of $\mathbf{V}$ if:

- $\mathbf{U}$ is a vector space defined under the same addition and multiplication as $\mathbf{V}$,
- $\forall x \in \mathbf{U}, x \in \mathbf{V}$.
- $\vec{0}_V \in \mathbf{U}$.

Example: $\mathbb{R}^2$

$$\mathbf{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 - x_2 = 1 \right\} \subseteq \mathbb{R}^2$$

This set lacks the additive identity and is hence, not a subspace.

Example

Let $\mathbf{V}$ be any vector space.

- $\mathbf{U} = \mathbf{V}$ is a subspace of $\mathbf{V}$. That is, any space is a subspace of itself.
- $\{\vec{0}\}$, where $\vec{0}$ is the additive identity of $\mathbf{V}$ is the trivial subspace of $\mathbf{V}$.

Example: $\mathbb{R}^3$

$\mathbf{V} = \mathbb{R}^3$,

$$\mathbf{U} = \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix} | x_1, x_3 \in \mathbb{R} \right\}$$

is a subspace of $\mathbf{V}$

Theorem 1.1: Subspace Conditions

If $\mathbf{V}$ is a vector space and $\mathbf{U} \subseteq \mathbf{V}$, then $\mathbf{U}$ is a subspace if and only if

- $\vec{0} \in \mathbf{U}$, where $\vec{0}$ is the additive identity of $\mathbf{V}$.
- $\mathbf{U}$ is closed under vector addition and scalar multiplication, as defined in $\mathbf{V}$.

Proof. The biconditional statement consists of two distinct propositions.

Statement 1.1. If $\mathbf{U}$ is a subspace, then $\vec{0} \in \mathbf{U}$ and $\mathbf{U}$ is closed under vector addition and scalar multiplication.

If $\mathbf{U}$ is a subspace, it is also a vector space defined under the operations of the superspace. By virtue of being a vector space, it is closed under addition and scalar multiplication. This also implies the existence of the additive identity, however it may or may not be the same as the one from $\mathbf{V}$. Let $v \in \mathbf{U}$. If the additive inverse is $-v$,

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

Closure under addition implies $\vec{0} \in \mathbf{U} \implies \vec{0} \in \mathbf{V}$. The uniqueness of the additive identity ensures that $\vec{0}$ is the additive identity of $\mathbf{V}$.

Statement 1.2. Assume $\mathbf{U}$ satisfies closure and additive identity properties. Then, $\mathbf{U}$ is a subspace of $\mathbf{V}$.

Let $u_1, u_2 \in \mathbf{U}$, then $u_1, u_2 \in \mathbf{V}$, so commutativity and associativity of $\mathbf{V}$ extends to $\mathbf{U}$. $\vec{0} \in \mathbf{U}$ by definition of the additive identity in $\mathbf{V}$. For any $u \in \mathbf{U}$,

$$u + \vec{0} = u$$

$\therefore \vec{0}$ is the additive identity of $\mathbf{U}$.
Let $u \in \mathbf{U} \subseteq \mathbf{V}$

$$0u = (1 - 1)u = \vec{0}$$

$(-1)u \in \mathbf{U}$ by closure under scalar multiplication.

$$u + (-u) = \vec{0}$$

Hence, for any vector in the subspace, the additive inverse is well-defined. All other properties follow for any vector in $\mathbf{U}$ as $v \in \mathbf{U} \implies v \in \mathbf{V}$.

$\therefore \mathbf{U} \subseteq \mathbf{V}$.

Example

$$\mathbf{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} | x_1 x_2 x_3 = 0 \right\} \subseteq \mathbb{R}^n$$

The additive identity is within the subspace as $0 \cdot 0 \cdot 0 = 0$ However, closure under vector addition fails. Consider $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{U}$

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \notin \mathbf{U}$$

Hence, $\mathbf{U} \nsubseteq \mathbb{R}^3$.

Example

$$\mathbf{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} | x_3 = 2x_1 - x_2 \right\}$$

The additive inverse is contained as $0 = 2 \cdot 0 - 0$. Let $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbf{U}$.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$$

$$2(x_1 + y_1) - (x_2 + y_2) = 2x_1 - x_2 + 2y_1 - y_2 = x_3 + y_3$$

Hence, for all $x, y \in \mathbf{U}$, $x + y \in \mathbf{U}$.

Let $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbf{U}$, $\alpha \in \mathbb{R}$

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{pmatrix}$$

$$2(\alpha x_1) - \alpha x_2 = \alpha(2x_1 - x_2) = \alpha(x_3)$$

Hence, $\forall x \in \mathbf{U}, \forall \alpha \in \mathbb{R}, \alpha x \in \mathbf{U}$.

Example

$$\mathbf{U} = \left\{ \alpha \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \alpha, \beta \in \mathbb{R} \right\}$$

2 Span and Linear Combinations

Definition: Linear Combination

Let V be a vector space such that $v_1, v_2, \cdots v_n \in V$. A linear combination of this collection of vectors is the vector given by

$$v = \sum_{k=1}^{n} \alpha_k v_k, \forall k \alpha_k \in \mathbb{R}$$

Example: $\mathbb{R}^2$

Assume two vectors $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $w = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$. One linear combination of these vectors is

$$x = 6 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

.

Example: $\mathbb{R}^3$

Is $\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ a linear combination of

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}$$

In other words, are there $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + 2\alpha_2 \\ \alpha_1 - 2\alpha_3 \\ \alpha_1 + \alpha_2 - \alpha_3 \end{pmatrix}$$

This turns into the system

$$\alpha_1 + 2\alpha_2 = 1 \quad (1)$$

$$\alpha_1 - 2\alpha_3 = 0 \quad (2)$$

$$\alpha_1 + \alpha_2 - \alpha_3 = 3 \quad (3)$$

The given system does not have any solutions. Hence, the given vector cannot be expressed as a linear combination of the others.

Definition: Span

Let V be a vector space. Given vectors $v_1, v_2 \cdots v_n$, the span is the subset of V defined by

$$\text{span}(v_1, v_2, \cdots v_n) = \left\{ \sum_{k=1}^{n} \alpha_k v_k | \alpha_1, \alpha_2 \cdots \alpha_n \in \mathbb{R} \right\}$$

Convention:

$$\text{span}(\phi) = \{\vec{0}\}$$

Terminology:
LADR refers to a group of vectors as a "list", which is synonymous with a set or collection of vectors.

Example

Is $\begin{pmatrix} 4 \\ -6 \\ -1 \end{pmatrix} \in \text{span}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\right)$ in $\mathbb{R}^3$?

$$\begin{pmatrix} 4 \\ -6 \\ -1 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

$$4 = \alpha_1 + 2\alpha_2 \tag{4}$$
$$-6 = \alpha_1 \tag{5}$$
$$-1 = \alpha_1 + \alpha_2 \tag{6}$$

This system is consistent and has the solution $\alpha_1 = -6, \alpha_2 = 5$. Therefore,

$$\begin{pmatrix} 4 \\ -6 \\ -1 \end{pmatrix} = -6 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 5 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

Yes, it belongs to the span of the other two vectors.

Definition: Spanning Set

If V is a vector space, then the collection $v_1, v_2 \cdots v_n$ is called a spanning set if any vector $v \in V$ can be represented as a linear combination of the collection. In other words,

$$\text{span}(v_1, v_2 \cdots v_n) = V$$

Example: $\mathbb{R}^n$

Let a collection of $\mathbb{R}^n$ be the collection

$$U = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \cdots \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix} \right\}$$

This set spans $\mathbb{R}^n$ and is called the **canonical basis** of $\mathbb{R}^n$.

Example

Claim: $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ span $\mathbb{R}^2$.

Let $V \ni v = \begin{pmatrix} x \\ y \end{pmatrix}$. To prove the claim, we show that

$$v = \alpha_1 v_1 + \alpha_2 v_2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

We get the system

$$x = \alpha_1 + 2\alpha_2 \tag{7}$$
$$y = 2\alpha_1 \tag{8}$$

We get $\alpha_1 = \frac{y}{2}$, $\alpha_2 = \frac{x}{2} - \frac{y}{4}$

Example

Claim: $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$ do not span $\mathbb{R}^2$. Let $\mathbb{R}^2 \ni v = \begin{pmatrix} x \\ y \end{pmatrix}$, s.t.

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

This gives us the system

$$2\alpha_1 + 4\alpha_2 = x \tag{9}$$
$$\alpha_1 + 2\alpha_2 = y \tag{10}$$

This imposes the condition $x = 2y$ on the components of v. Hence, there are vectors which do not belong to this span.

Proposition: Let V be a vector space. Then for any $v_1, v_2 \cdots v_n \in V$ span$(v_1, v_2, \cdots v_n)$ is a subspace of V.

Proof.

$$\vec{0} = \sum_{k=1}^{n} 0 v_k$$

Since $0v = \vec{0}$, the additive identity belongs to the span. Consider two elements of the span $a = \sum_{k=1}^{n} \alpha_k v_k$ and $b = \sum_{k=1}^{n} \beta_k v_k$.

$$a + b = \sum_{k=1}^{n} \alpha_k v_k + \sum_{k=1}^{n} \beta_k v_k$$

$$a + b = \sum_{k+1}^{n} (\alpha_k v_k + \beta_k v_k)$$

$$a + b = \sum_{k+1}^{n} (\alpha_k + \beta_k) v_k$$

Since the sum of any real numbers is a real number, $a + b$ belongs to the span. Consider a scalar $\lambda \in \mathbb{R}$. Using the distributive property,

$$\lambda a = \lambda \sum_{k=1}^{n} \alpha_k v_k = \sum_{k=1}^{n} (\lambda \alpha_k) v_k$$

Hence, λa belongs to the span.

Therefore, the span is a subspace of V.

Proposition: span$(v_1, v_2 \cdots v_n)$ is the smallest subspace of V that contains $v_1, v_2, v_3, \cdots v_n$.

Proof. Let $v_1, \cdots v_n$ be a list of vectors in V.
The span of this list contains the zero vector because

$$\vec{0} = 0v_1 + 0v_2 \cdots 0v_n$$

It is closed under addition and scalar multiplication as well.

$$(a_1 v_1 + a_2 v_2 \cdots a_n v_n) + (c_1 v_1 + c_2 v_2 \cdots c_n v_n) = (a_1 + c_1) v_1 + (a_2 + c_2) v_2 \cdots (a_n + c_n) v_n$$

$$\lambda(a_1 v_1 + \cdots a_n v_n) = \lambda a_1 v_n + \cdots \lambda a_n v_n = \mu v_1 + \cdots \mu v_n$$

Each of the vectors of the list belong to the span of the list, by setting the corresponding coefficient to 1 and the rest to zero. Also, any subspace containing each of the vectors in the list also contains the entire span, since subspaces are closed under addition and scalar multiplication. Therefore, the span of any list of vectors is the smallest subspace containing them.

Definition: Dimensionality

A vector space V is finite-dimensional if there is a finite set $v_1, v_2, v_3 \cdots v_n \in V$ which span V. If V is not finite-dimensional, it is infinite-dimensional

Claim: Let $\mathcal{P}$ be the space of all polynomials. Then, $\mathcal{P}$ is infinite dimensional.

Proof. Suppose by contradiction that $\mathcal{P}$ is finite. Then

$$\exists p_1, p_2 \cdots p_k \in \mathcal{P} \text{ s.t. } \mathcal{P} = \text{span}(p_1, \cdots p_k)$$

Let $n_i := \deg(p_i)$. Then, the larget value of n_i is finite. Hence, for any $v \in \mathcal{P}$

$$v = \sum_{i=1}^{k} \alpha_i p_i$$

Notating each p_i as a polynomial,

$$p_i(x) = a_{i0} + a_{i1} + \cdots a_{in_i} x^{n_i}$$

Hence, the largest power that appears with nonzero coefficient in $\mathcal{P}$ has power less than or equal to the greatest value of n_i. Hence, for an arbitrary element of $\mathcal{P}$, $\deg(p) \leq \max n_i$. This contradicts the definition of $\mathcal{P}$, hence P is infinite-dimensional.

Definition: Linear Independence

A collection of vectors $v_1, \cdots v_n \in V$ is linearly independent if the only solution to

$$\vec{0} = \sum_{k=1}^{n} \alpha_k v_k$$

is $\alpha_1, \cdots \alpha_k = 0$.
The empty set is also deemed to be linearly independent.

Proposition: If V is a vector space, $v_1, v_2 \in V$ is linearly dependent if and only if

$$\exists \alpha \text{ such that } v_1 = \alpha v_2 \text{ or } v_2 = \alpha v_1$$

Proof. Let the second statement hold true. Without loss of generality, assume $v_1 = \alpha v_2$.

If $\alpha = 0$, $v_1 = \vec{0}$ and hence the collection is linearly dependent. If $\alpha \neq 0$, by field axioms there exists a number α^{-1} such that $\alpha\alpha^{-1} = 1$.

$$v_1 + v_2 = \vec{0}$$
$$-\alpha^{-1}\alpha v_2 + v_2 = \vec{0}$$
$$-(1)v_2 + v_2 = \vec{0}$$
$$\vec{0} = \vec{0}$$

Hence, for a nonzero value of α, the linear combination of the collection gives the zero vector. Therefore, the collection is linearly dependent.

Let the collection be linearly dependent. By extension, $v_1 \neq \vec{0}$ and $v_2 \neq \vec{0}$. For some nonzero α_1 or α_2,

$$\alpha_1 v_1 + \alpha_2 v_2 = \vec{0}$$

Without loss of generality, we can assume $\alpha_1 \neq 0$. Then,

$$\alpha_1^{-1}\alpha_1 v_1 + \alpha_1^{-1}\alpha_2 v_2 = \alpha_1^{-1}\vec{0}$$

Hence,

$$v_1 + \alpha_1^{-1}\alpha_2 v_2 = \vec{0}$$

$$v_1 = \beta v_2 : \beta = -\alpha_1^{-1}\alpha_2$$

In $\mathbb{R}^4$, is span $\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \\ 3 \end{pmatrix} \right\} = \mathbb{R}^4$?

No, the max number of linearly independent vector is the lower bound for the cardinality of any spanning set.

Lemma 1.2

Suppose $v_1, \cdots v_n$ is a linearly dependent list in V. Then, there exists $j \in \{1, 2, \cdots m\}$ such that $v_j \in \text{span}(v_1, \cdots v_{j-1}, v_{j+1}, v_m)$

If this vector is removed, the resulting list's span is unchanged.

Proof. Because the list is linearly dependent, there exists scalars $\alpha_1, \cdots \alpha_n \in \mathbb{F}$ not all zero such that

$$\vec{0} = \alpha_1 v_1 + \cdots \alpha_n v_n$$

If j is the largest index such that $a_j \neq 0$, then we can rewrite the above equation as

$$v_j = -\frac{\alpha_1}{\alpha_j} v_1 - \frac{\alpha_2}{\alpha_j} v_2 \cdots \frac{\alpha_n}{\alpha_j} v_n \text{ (1)}$$

Clearly, $v_j \in \text{span}(v_1, \cdots v_{j-1}, v_{j+1}, v_n)$

Consider an arbitrary vector u in the span of $v_1, \cdots v_n$. It can be written as

$$u = \alpha_1 v_1 + \cdots \alpha_j v_j \cdots + \alpha_n v_n$$

If we substitute (1) into the expression, we get u as a linear combination of $v_1, \cdots v_{j-1}, v_{j+1}, \cdots v_n$.

This works if the list has more than 1 vector. Note that if a list contains only 1 vector and is linearly dependent, that vector must be the zero vector. Since $\vec{0} \in \text{span}(\phi)$, the first statement is proven. Removing $\vec{0}$ from the list gives us $\text{span}(\phi)$, which is $\vec{0}$. Hence, the span is unchanged.

Proposition: In a finite-dimensional vector space, the length of every linearly independent list of vectors is less than or equal to the length of every spanning list of vectors.

Proof. Suppose $u_1, \cdots u_m$ is linearly independent in V and $w_1, \cdots w_n$ spans V. We need to show that $m \leq n$.

Since $w_1, \cdots w_n$ span V, adding another vector to the list does not change its span. So,

$$u_1, w_1, \cdots w_n$$

also spans V. By the linear dependence lemma, one of the vectors w can be removed and the resulting set will still span V. We can iteratively add u_j to the list, and each time it will stil span V. Since all vectors u_i are linearly independent, by the linear dependence lemma, the vector we remove will be one of the w's and not one of the u's. By the time all vectors u have been added, the set contains $u_1, \cdots u_m$. Since at each iteration of adding u we remove a w, there are at least as many w's as there are u's. Hence, the statement is proven that $m \leq n$.

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

Proof. Suppose V is finite dimensional and U is a subset of V. We need to show that U is also finite dimensional. If U is $\{\vec{0}\}$, then U is finite dimensional and we are done. Otherwise, we choose a vector v_1. If $U = \text{span}(v_1)$, then U is finite dimensional. If not, we can add another vector v_2 to our list that is not in the span of v_1. We again check if the span of these two vectors is U. We iteratively add vectors to the list that are not in its span. Eventually, this process has to terminate, because the number of linearly independent vectors is always less than or equal to the number of spanning vectors in a space.

Definition: Basis

If a list of vectors $v_1, \cdots v_n \in V$ can uniquely represent any arbitrary vector $v \in V$, then the list is known as a basis for the space.

Proof. We will first show that if a list $v_1, \cdots v_n \in V$ is a basis, then any vector $v \in V$ admits a unique reprsentation. Since the list spans V, for some scalars $a_1 \cdots a_n$

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

Let us assume that there is another set of scalars $c_1, \cdots c_n$ such that

$$v = c_1 v_1 + \cdots c_n v_n$$

Subtracting these two equations,

$$\vec{0} = (a_1 - c_1)v_1 + \cdots (a_n - c_n)v_n$$

Since the list is linearly independent, $a_1 - c_1 = \cdots a_n - c_n = 0$, so $a_1 = c_1, \cdots a_n = c_n$. Hence, the representation of any vector is unique in terms of the basis.

We will now show the other direction, if $v \in V$ admits a unique representation in terms of the list, then the list is linearly independent and a basis. Since any vector can be written in terms of the list, by definition the list spans V. To show linear independence, consider the vector $\vec{0}$ and the scalars $a_1, \cdots a_n$

$$\vec{0} = a_1 v_1 + a_n v_n$$

Since the representation is unique, and we know that $0v = \vec{0}$, all of the scalars have to equal zero. Hence, by definition the list is linearly independent.

Example

What are some bases for $\mathbb{R}^2$?

- **Canonical Basis:** $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ Let $\mathbb{R}^2 \ni x = \begin{pmatrix}x_1\\x_2\end{pmatrix}$ be an arbitrary vector. Then,

  $$x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix}$$

  So, the vectors span $\mathbb{R}^2$.

  Suppose $\exists \alpha_1, \alpha_2 \in \mathbb{R}$ s.t.t

  $$\alpha_1\begin{pmatrix}1\\0\end{pmatrix} + \alpha_2\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

  Therefore,

  $$\begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \leftrightarrow \alpha_1 = \alpha_2 = 0$$

  Hence, the vectors are linearly independent. Therefore, given vectors form a basis.
- $\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}$:

Corollary 1.3

If $v_1, \cdots v_n$ span V, and $\{v_1, \cdots v_n\} \subseteq \{w_1, \cdots w_n\}$, then $w_1, \cdots w_n$ spans V.

Show that $\left\{\begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}3\\1\\2\end{pmatrix}\right\}$ spans $\mathbb{R}^3$.

We can show that the canonical basis is a subset of the span of these vectors. Then, this set of vectors would also span the space by the corollary above.

$$\begin{pmatrix}1\\0\\0\end{pmatrix} = 0\begin{pmatrix}1\\2\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}2\\0\\0\end{pmatrix} + 0\begin{pmatrix}3\\1\\2\end{pmatrix}$$

$$\begin{pmatrix}0\\1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\2\\0\end{pmatrix} - \frac{1}{4}\begin{pmatrix}2\\0\\0\end{pmatrix} + 0\begin{pmatrix}3\\1\\2\end{pmatrix}$$

$$\begin{pmatrix}0\\0\\1\end{pmatrix} = -\frac{1}{4}\begin{pmatrix}1\\2\\0\end{pmatrix} + \frac{7}{8}\begin{pmatrix}2\\0\\0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}3\\1\\2\end{pmatrix}$$

Therefore, the given subset spans $\mathbb{R}^3$.

Proposition: Any spanning list can be reduced to a basis. Similarly, any linearly independent list can be extended to a basis.

Proof. Consider a list $v_1, \cdots v_n$ and the multistep process as follows:

Step 0:
If $v_1 = \vec{0}$, then remove it from the list, otherwise leave the list unchanged.

For $k \in \{2, \cdots n\}$:
If v_k is in the span of the list $v_1, \cdots v_{k-1}, \cdots v_n$, then remove it from the list. Otherwise, leave the list unchanged.

By the end of the sequence, we have a spanning list, because we have only removed vectors from the list which are in the span of the list. Also, by ensuring that none of the vectors are in the span of the other vectors of the list, we have reduced the list to a linearly independent list. Therefore, the resulting list after this "reduction algorithm" is a basis in V. Note that the order of removing elements matters for the vectors that form a basis, however the resulting list will always be a basis.

Corollary 1.4

Every finite dimensional vector space has a basis.

Proof. By definition if V is finite dimensional, there is a finite spanning set, which can always be reduced to a basis.

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

Proof. Consider a linearly independent list $w_1, \cdots w_m$ and a basis of V $e_1, \cdots e_n$. Then, the list

$$w_1, \cdots w_m, e_1 \cdots e_n$$

spans the space. Using the reduction algorithm, we can recover a basis from this spanning set, containing all elements of the w list and some elements of the basis list e. Note that the resulting list will contain all of the vectors of w because they are linearly independent.

Proposition: Any two bases of a finite-dimensional vector space have the same length.

Proof. Consider any two bases of a space B_1 and B_2. Since B_1 is linearly independent and B_2 is spanning, if we denote the size of B_1 as m and the size of B_2 as n, then we know that $m \leq n$. Since B_1 is spanning and B_2 is linearly independent, we know that $m \geq n$. Hence, we can conclude that $m = n$.

Definition: Definition

The dimension of a finite-dimensional vector space is the length of any basis of the vector space.
The dimension of V is denoted by $\dim V$.

Proposition: If V is a finite dimensional space and U is a subspace of V, then $\dim U \leq \dim V$.

Proof. We can consider a basis of U to be a linearly independent list of V, and a basis of V to be a spanning set of V. Then, by definition, $\dim U \leq \dim V$.

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

Proof. Let $n = \dim V$ and let $v_1, \cdots v_n$ be a linearly independent list in V. Then, we can extend the list to a basis. However, since any basis of V has exactly n elements, we do not need to add any vectors to the list, it is alraedy a basis.

Proposition: Suppose V is finite-dimensional. Then every spanning list of size $\dim V$ is a basis of V.

Proof. Let $m = \dim V$ and $v_1, \cdots v_m$ span V. We can reduce the list to a basis of V. However, since any basis of V is of size m, we do not need to remove any vectors, as the list is already a basis.

3 Linear Maps/Transformations

Definition: Linear Map

Suppose V, W are vector spaces. $T: V \to W$ is called a linear map from V to W if every element of V is related to some element in W, such that for two vectors $v_1, v_2 \in V$

$$T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2), \alpha_1, \alpha_2 \in \mathbb{R}$$

$T(v_1), T(v_2) \in W$.

Example: $\mathbb{R}^2 \to \mathbb{R}^2$

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 \end{pmatrix}$$

Let $v_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, v_2 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$T(v_1 + v_2) = T \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 + 2(x_2 + y_2) \\ 3(x_1 + y_1) \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 \end{pmatrix} + \begin{pmatrix} y_1 + 2y_2 \\ 3y_1 \end{pmatrix} = Tv_1 + Tv_2$$

$$T(\lambda v_1) = T \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 + 2\lambda x_2 \\ 3\lambda x_1 \end{pmatrix} = \lambda \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 \end{pmatrix} = \lambda T(v_1)$$

Proposition: If $T: V \to W$ is linear, then $T(\vec{0}_V) = \vec{0}_W$

Proof. By definition of a linear map,

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

Adding $-T(\vec{0}_V)$ to both sides,

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

Contrapositive: If $T: V \to W$ is a transformation and $T(\vec{0}_V) \neq \vec{0}_W$, then T is nonlinear.

Proof. This is a contrapositive and logically equivalent to the statement proven above.

Example

zero
The zero map takes a vector and maps it to the additive inverse of the codomain space. If we have $0 \in \mathcal{L}(V, W)$, then for any vector $v \in V$

$$0v = \vec{0}_W$$

identity
The identity map $I \in \mathcal{L}(V, V)$ takes a vector and returns the same vector.

$$Iv = v$$

differentiation The differential operator $D \in \mathcal{L}(\mathcal{P}_n, \mathcal{P}_n)$ is defined as

$$Dp = p'$$

Proposition: Suppose $v_1, \cdots v_n$ is a basis of V and $w_1, \cdots w_n \in W$ is a list of vectors, not necessarily a basis. Then there exists a unique linear map $T: V \to W$ such that

$$Tv_j = w_j \text{ for each } j = 1, \cdots n$$

Proof. First, we will show that such a linear map exists. Consider $T : V \to W$ such that

$$T(c_1v_1 + \cdots c_nv_n) = c_1w_1 \cdots c_nw_n$$

This map maps each vector as desired, since we can set the arbitrary coefficients of the desired vector to 1, and the rest to 0. We will now show that this is a linear map.

Let $z_1, z_2 \in V$ such that $z_1 = \sum_{k=1}^{n} \alpha_k v_k$ and $z_2 = \sum_{k=1}^{n} \beta_k v_k$ Then

$$T(z_1 + z_2) = T\left(\sum_{k=1}^{n}(\alpha_k + \beta_k)v_k\right) = \sum_{k=1}^{n} T(\alpha_k + \beta_k)v_k = \sum_{k=1}^{n} T(\alpha_k v_k) + \sum_{k=1}^{n} T(\beta_k v_k) = T(z_1) + T(z_2)$$

$$T(\lambda z_1) = T\left(\sum_{k=1} \lambda\alpha_k v_k\right) = \sum_{k=1} \lambda\alpha_k w_k = \lambda T\left(\sum_{k=1}^{n} \alpha_k v_k\right) = \lambda T z_1$$

If we assume there is another map $S : V \to W$ is linear and $Sv_i = w_i$ for each $i = 1, \cdots n$. Let $v \in V$ be an arbitrary vector $v = \sum_{i=1}^{n} \alpha_i v_i$

$$Sv = S\left(\sum_{i=1}^{n} \alpha_i v_i\right)$$

$$Sv = \sum_{i=1} S(\alpha_i v_i)$$

$$Sv = \sum_{i=1} \alpha_i S v_i$$

$$Sv = \sum_{i=1} \alpha_i w_i = Tv$$

Therefore, T is unique.

Definition: $\mathcal{L}(V, W)$

The set of all linear maps from $V \to W$ is denoted as $\mathcal{L}(V, W)$. It always contains the trivial map $Tv = \vec{0}$.

Proposition: $\mathcal{L}(V, W)$ is a vector space.

Proof. **Operator Definitions**
Consider two elements of $\mathcal{L}(V, W)$, $T : V \to W$ and $S : V \to W$. Then, the addition of these two maps is defined as

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

Definition: Composite Maps

Suppose there are three vector spaces V, W, Z. Let $S : W \to Z$ and $T : V \to W$. Then $ST : V \to Z$ such that $ST(v) = S(Tv)$. It is a composition of the two maps, $S \circ T$

Proof. We will show that ST is linear. Consider $v_1, v_2 \in V$.

$$ST(v_1 + v_2) = S(T(v_1 + v_2)) = S(T(v_1) + T(v_2)) = S(T(v_1)) + S(T(v_2)) = ST(v_1) + ST(v_2)$$

Consider $\lambda \in \mathbb{F}$, then

$$ST(\lambda v_1) = S(T(\lambda v_1)) = S(\lambda T(v_1)) = \lambda ST(v_1)$$

Definition: Injective

If V and W are any sets, a function $T : V \to W$ is injective if for any $v_1, v_2 \in V$, $Tv_1 = Tv_2 \implies v_1 = v_2$. The contrapositive is also commonly used for proof based methods.

Definition: Surjective

If V and W are any sets, a function $T : V \to W$ is surjective if for any vector $w \in W$, there is some $v \in V : Tv = w$.

Definition: Null space/Kernel

The null space of $T : V \to W$ is the subset $N(T)$ of V such that

$$v \in N(T) \implies T(v) = \vec{0}_W$$

Definition: Range

The range of T is the subset of W such that each vector of the subset has a preimage in V.

Note: If $U \subseteq V$, any subset and $T : V \to W$ is a function, the image of U under T is

$$T(U) = \{Tv : v \in U\}$$

Proposition: A linear map is injective if and only if its kernel is $\{\vec{0}\}$.

Proof. Let $T \in \mathcal{L}(V, W)$ be an arbitrary linear map. We will first show that if $\text{null}(T) = \{0\}$, then the map is injective.
Consider two vectors $v, w \in V$. Then, to show that T is injective,

$$T(v) = T(w) \implies v = w$$

Starting from the hypothesis,

$$T(v) = T(w)$$

Using the linearity of T,

$$T(v - w) = \vec{0}$$

Clearly, $v - w \in \text{null}(T)$. But, we started with the assumption that $\{\vec{0}\}$ is the only element of $\text{null}(T)$. Therefore,

$$v - w = \vec{0}$$
$$v = w$$

Hence, proven.
We will now assume that T is injective, and show that $\text{null}(T) = \{\vec{0}\}$. By the linearity of T, we know that $T\vec{0} = \vec{0}$. Therefore, $\vec{0} \in \text{null}(T)$. If any other vector v exists such that

$$Tv = \vec{0}$$

then T is not injective by definition. Hence, if T is injective, then $\text{null}(T) = \{\vec{0}\}$.

Definition: Rank Nullity Theorem

For any linear map $T \in \mathcal{L}(V, W)$,

$$\dim V = \dim \text{null}(T) + \dim \text{range}(T)$$

Corollary 1.5

There is no injective linear map from a higher dimensional space to a lower dimensional space.

$$\dim \mathsf{null}(T) = \dim V - \dim \mathsf{range}(T)$$

The largest possible dimension of $\mathsf{range}(T)$ is $\dim W$, therefore

$$\dim \mathsf{null}(T) = \dim V - \dim W > 0$$

Since the dimension of $\mathsf{null}(T)$ is greater than zero, T is not injective.

Corollary 1.6

There is no surjective map from a lower dimensional space to a higher dimensional space.

$$\dim \mathsf{range}(T) = \dim V - \dim \mathsf{null}(T)$$

Since $\dim W > \dim V$,

$$\dim \mathsf{range}(T) = \dim V - \dim \mathsf{range}(T) < \dim W$$

Therefore, the map can never be surjective.

Corollary 1.7

For all maps $T \in \mathcal{L}(V, W)$, if $\dim V = \dim W$, then T is injective iff it is surjective.
T is injective implies $\dim \mathsf{null}(T) = 0$. Then, $\dim V = \dim \mathsf{range}(T) = \dim W$.

Definition: Matrix

Let m and n denote positive integers. An $m \times n$ matrix A is a rectanglar array of elements of $\mathbb{F}$ with m rows and n columns.

$$\begin{pmatrix} A_{1,1} & \cdots & \cdots & A_{1,n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ A_{m,1} & \cdots & \cdots & A_{m,n} \end{pmatrix}$$

Definition: Matrix of a linear map

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \cdots v_n$ is a basis of V, and $w_1, \cdots w_m$ is a basis of W. The matrix of T is the $m \times n$ matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

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

If the bases are not clear, we can use the notation

$$[T]_{\mathcal{B}_W, \mathcal{B}_V}$$

4 Invertibility and Isomorphisms

Definition: inverse, invertible

A linear map $S : V \to W$ is defined to be invertible if there exists a map $T : W \to V$ such that $ST = I_V$ and $TS = I_W$, where I_W and I_V are the identity maps on W and V respectively.

In such a case, $T = S^{-1}$ and is called the **inverse** of S.

Theorem 1.8

Every linear map has a unique inverse.

Proof. Let S_1 and S_2 be the inverses of T.

$$S_1 I = S_1(TS_2) = (S_1T)S_2 = S_2$$

Hence, every linear map has a unique inverse.

Theorem 1.9

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

Proof. Let $T : V \to W$ be invertible. To show that it is injective, let $Tu = Tv$ for some $u, v \in V$

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

Since $Tu = Tv \implies u = v$, T is injective.

Consider $w \in V$

$$w = T^{-1}Tw$$

which implies range $T = W$. Hence, T is surjective

Now suppose T is injective and surjective. We want to prove that T is invertible. Consider for every element $w \in W$ there is some element $Sw \in V$ such that $T(Sw) = w$. This exists because the map is bijective.

$$T \circ S(w) = w = Iw$$

We will now show that $S \circ T = I$

$$T \circ (S \circ T)(v) = (T \circ S) \circ T(v) = ITv$$

Because T is injective, this implies that $S \circ T = I$. To complete the proof, we need to show that S is linear.

$$T(Sw_1 + Sw_2) = T(Sw_1) + T(Sw_2) = w_1 + w_2$$

$Sw_1 + Sw_2$ is the unique element mapping to $w_1 + w_2$. By the definition of S, this means $S(w_1 + w_2) = Sw_1 + Sw_2$

For some scalar $\lambda \in \mathbb{F}$,

$$T(\lambda S(w_1)) = \lambda T(S(w_1)) = \lambda w_1$$

λSw_1 is the unique element mapping to λw_1. Therefore, S is linear.

Definition: Isomorphism

Two vector spaces V and W are isomorphic if there is an invertible linear map $T \in \mathcal{L}(V, W)$ T is an isomorphism from V to W.

Proposition: Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.

Proof. We will first show that if two spaces are isomorphic, they have the same dimension.

Consider two finite-dimensional isomorphic vector spaces V and W defined over $\mathbb{F}$. By definition, there exists an isomorphism $T : V \to W$. Because T is invertible, we know it's bijective. Therefore, $\dim \text{range}(T) = \dim W$ and $\dim \text{null}(T) = 0$. Therefore, $\dim V = \dim W$.

We will now show that if two spaces have the same dimension, they are isomorphic. Consider the bases

$$\mathcal{B}_V = v_1, \cdots v_n$$

$$\mathcal{B}_W = w_1 \cdots w_n$$

Then, for scalars $\alpha_1, \cdots \alpha_n \in \mathbb{F}$, let T be defined by

$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 w_1 + \cdots + \alpha_n w_n$$

All possible linear combinations of $\mathcal{B}_W$ span W because $\mathcal{B}_W$ is a basis. Therefore, $\text{range}(T) = W$. Since $\mathcal{B}_V$ is a basis, it is linearly independent. Therefore, the only solution of $Tv = \vec{0}_W$ is $v = \vec{0}_V$. Hence, T is bijective. By extension, it is an isomorphism between V and W.

Definition: Operator

An operator is a linear map from a space to itself. It is a special kind of isomorphism, called an automorphism.

Definition: Pivot

The first nonzero entry of each row is called the pivot.

Definition: Echelon Form

All zero rows are below all nonzero rows
The pivot in each row is strictly to the right of the pivot in the row above it.

Definition: Reduced Echelon Form

If a matrix is in reduced echelon form, it is in echelon form and:

All pivot entries are 1.
All entries above the pivot are 0.

A system is inconsistent (does not have a solution) if and only if there is a pivot in the last column of the echelon form of the augmented matrix.

Pivot Analysis

1. A system has at most 1 solution if there is a pivot in every column of the coefficient matrix. Since there are no columns without a pivot, there is a leading variable in each row. Therefore, there are no free variables and the finitely-many solutions.
2. A system has at least 1 solution if there is a pivot in every row of the coefficient matrix.
3. A system has exactly 1 solution if there is a pivot in every row and column of the coefficient matrix.

Corollary 1.10

Consider the system of vectors $v_1, \cdots v_n \in \mathbb{F}^n$, and let $A = [v_1, \cdots v_n]$ be an $n \times n$ matrix with columns $v_1, \cdots v_n$. Then,

1. The system of vectors are linearly independent if the echelon form of A has a pivot in every column.
2. The system of vectors is spanning if the echelon form of A has a pivot in every row.
3. The system of vectors forms a basis if there is a pivot in every row and every column.

Corollary 1.11

A matrix is invertible if and only if there is a pivot in every row and column.

Determinant

The determinant is an alternating multilinear form $V^n \to \mathbb{F}$ where V is a vector space defined over the field $\mathbb{F}$. It has the following properties:

1. Multilinearity: For vectors $v_1, \cdots v_n \in V$, and $\lambda \in \mathbb{F}$,
$$D(\alpha v_1, \cdots v_n) = \alpha D(v_1, \cdots v_n)$$
$$D(v_1, \cdots v_k + v_{k+1}, \cdots, v_n) = D(v_1, \cdots v_k, \cdots v_n) = D(v_1, \cdots v_{k+1}, v_n)$$
2. Antisymmetry
$$D(v_1, \cdots v_i, \cdots v_j, \cdots v_n) = -D(v_1, \cdots v_j, \cdots v_i, \cdots v_n)$$
3. Normalization: The determinant of the canonical basis of any space and the identity matrix is 1.

For a square matrix A,
$$\det(A) = \det\left(A^T\right)$$
Therefore, any property on the columns of A also applies to the rows of A. The following properties hold for a square matrix A:

1. If any row or column in A is completely zeros, then $\det A = 0$

   Proof. Using multilinearity,
$$\det\left(v_1, \cdots \vec{0}, \cdots v_n\right) = \det\left(v_1, \cdots 0 \cdot \vec{0}, \cdots v_n\right) = 0 \det\left(v_1, \cdots \vec{0}, \cdots v_n\right) = 0$$
2. If any rows or columns are equal, then $\det A = 0$.

   Proof. If $v_i = v_j$
$$\det(v_1, \cdots v_i, \cdots v_i, \cdots v_n) = -\det(v_1 \cdots v_i, \cdots v_i, \cdots v_n) \therefore \det(v_1, \cdots v_n) = 0$$
3. If one column is a multiple of another, then $\det A = 0$.

   Proof. This uses the same logic as the two properties above - pull out the multiple using multilinearity, then use antisymmetry to prove that $\det A = 0$.
4. If the matrix is not invertible, (the columns of the matrix are linearly dependent), then the determinant is zero.

The determinant of a triangular matrix is the product of its diagonals.

Theorem 1.12

If A and B are square matrices, then
$$\det(AB) = \det(A)\det(B)$$

Corollary 1.13

For square matrices $A_1, \cdots A_n$, $\det(A_1 A_2 \cdots A_n) = \det(A_1)\det(A_2)\cdots\det(A_n)$

Corollary 1.14

For square matrices A and B,
$$\det(AB) = \det(BA)$$

$$\det(aA) = a^n \det(A)$$

5 Spectral Theory

Definition: Eigenvector/Eigenvalue

Consider a linear operator $T \in \mathcal{L}(V)$ for a finite dimensional space V defined over $\mathbb{C}$. If $Tv = \lambda v; \lambda \in \mathbb{C}, v \neq \vec{0}$, then v is called an eigenvector and λ is the corresponding eigenvalue.

Definition: Spectrum

The set of all eigenvalues of A is called the spectrum of A.

Consider the equation

$$Av = \lambda I_n v$$
$$(A - \lambda I_n)v = \vec{0}$$

So, $v \in$ null$(A - \lambda I_n)$. Since we are given that $v \neq \vec{0}$, the map corresponding to $A - \lambda I_n$ is not injective. Therefore, $A - \lambda I_n$ is not invertible, and

$$\det(A - \lambda I_n) = 0$$

Definition: Characteristic Polynomial

The characteristic polynomial of a given square matrix A is the equation

$$p(\lambda) = \det(A - \lambda I_n)$$

Note that this means any eigenvalue of A is a root of $p(\lambda)$.

Theorem 1.15: Fundamental Theorem of Algebra

A polynomial of degree n with coefficients in $\mathbb{R}$ have at most n roots, all in $\mathbb{C}$.

From the theorem above, the characteristic polynomial can be written as

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

for $c, \lambda_1, \cdots \lambda_n \in \mathbb{C}$.

Definition: Algebraic Multiplicity

The number of times a particular eigenvalue appears as a root in the characteristic equation is called its algebraic multiplicity. If it has multiplicity one, it is called a simple root.

Definition: Eigenspace

If λ is an eigenvalue of A, then the eigenspace of λ is the set of all eigenvectors corresponding to that particular eigenvalue, as well as the zero vector. Note that the eigenspace corresponds to the nullspace of $A - \lambda I_n$. Therefore, eigenspaces are subspaces of $\mathbb{C}^n$.

Definition: Geometric Multiplicity

The geometric multiplicity of a particular eigenvalue λ is the dimension of its corresponding eigenspace.

Definition: Similarity

Two matrices are similar if there exists an invertible matrix Q such that

$$A = Q^{-1}BQ$$

Proposition: If A and B are similar, $\det A = \det B$.

Proof. If the two matrices A and B are similar, then

$$A = Q^{-1}BQ$$

Applying the determinant on both sides,

$$\det(A) = \det\left(Q^{-1}\right)\det(B)\det(Q)$$

Lemma. $\det\left(Q^{-1}\right) = \frac{1}{\det(Q)}$
Proof. We know that

$$QQ^{-1} = I$$

Then,

$$\det(Q)\det\left(Q^{-1}\right) = \det(I)$$

By the normalization property of the determinant,

$$\det\left(Q^{-1}\right) = \frac{1}{\det(Q)}$$

Note that this is always well defined, because if $\det(Q)$ is 0, then Q^{-1} cannot exist. Therefore, from the given lemma,

$$\det(A) = \det(B)\det\left(Q^{-1}\right)\det(Q)$$
$$\det(A) = \det(B)$$

Hence, proven.

Proposition: Similar matrices have the same eigenvalues.
Proof. Consider two similar matrices A and B. Then, for some invertible matrix Q,

$$A = Q^{-1}BQ$$

We subtract λI from each side, where $\lambda \in \mathbb{C}$.

$$A - \lambda I = Q^{-1}BQ - \lambda I$$

Note that $I = Q^{-1}Q$, so

$$A - \lambda I = Q^{-1}(BQ - \lambda Q)$$
$$A - \lambda I = Q^{-1}(B - \lambda I)Q$$

Therefore, if A and B are similar, then $A - \lambda I$ and $B - \lambda I$ are also similar. Therefore,

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

Since A and B have the same characteristic polynomial, they have the same eigenvalues. Hence, proven.

Proposition. Geometric multiplicity is always less than algebraic multiplicity for any eigenvalue.
Proof. Consider a map $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Let the geometric multiplicity of λ^* be k. Consider an eigenbasis $v_1, \cdots v_k$ of the eigenspace of λ^*.

Corollary 1.16

If the algebraic multiplicity is 1, then the geometric multiplicity is 1.

If $A \in \mathcal{M}_{n\times n}(\mathbb{F})$,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} A_{ii}$$

Proposition. If $A \in \mathcal{M}_{n\times n}(\mathbb{C})$, and $\lambda_1, \cdots \lambda_n$ are its eigenvalues, then

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$$

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

Proof. look at textbok lol

Theorem 1.17

If A is triangular, then the eigenvalues are the diagonal entries of A.

5.1 Diagonalization

Definition: Diagonalizable

A linear map $T \in \mathcal{L}(V, V)$ is called diagonalizable if there exists a basis $\mathcal{B}$ which makes $[T]_{\mathcal{B},\mathcal{B}}$ is diagonal.

Definition: Eigenbasis

An eigenbasis is a basis of a space consisting only of eigenvectors.

Theorem 1.18

Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. The following are equivalent-

1. A is diagonalizable.
2. There exists an eigenbasis.
3. There exists an invertible matrix P such that

$$A = P^{-1}DP$$

where D is the matrix representation of A in terms of the eigenbasis.

Proof. Write $T_A v = Av \forall v \in V$.

$(2 \implies 1)$ Assume $B_E = v_1, \cdots v_n$ is an eigenbasis of V. Let λ_i be the eigenvalue associated with v_i. Then,

$$\begin{aligned} T_A v_1 &= \lambda_1 v_1 + 0v_2 + \cdots + 0v_n \\ T_A v_2 &= 0v_1 + \lambda_2 v_2 + \cdots + 0v_n \\ &\vdots \\ T_A v_n &= 0v_1 + \cdots 0v_2 + \cdots \lambda_n v_n \end{aligned}$$

Therefore,

$$[T_A]_{B_E,B_E} = \begin{pmatrix} \lambda_1 & \cdots & \cdots & \text{\huge 0} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \text{\huge 0} & \cdots & \cdots & \lambda_n \end{pmatrix}$$

Clearly, A is diagonalizable.

$(1 \implies 3)$ Assume A is diagonalizable. FINISH THIS LATER

How to diagonalize:

1. Find the eigenvectors, check if they are an eigenbasis
2. $Q = [v_1, \cdots v_n]$
3. Then, D has the corresponding eigenvalues.