MA2102 - Linear Algebra I

Autumn 2021

Instructor: Somnath Basu

IISER KOLKATA

Birmd

Fields, Vector Spaces & Subspaces:

- **Field:** It is a set (F) with 2 binary operations, $+ : F\times F \to F$ (*addition*), and $\cdot : F\times F \to F$ (*multiplication*) defined on it, which satisfied the *field axioms*, which are as follows:
  1 Commutative Laws:
  - ◇ $x + y = y + x$
  - ◇ $x \cdot y = y \cdot x$

  2 Associative Laws:.
  - ◇ $x + (y + z) = (x + y) + z$
  - ◇ $x \cdot (y \cdot z) = (x \cdot y) \cdot z$

  3 Distributive Law:
  - ◇ $x \cdot (y + z) = x \cdot y + x \cdot z$

  4 Existence of Identity:
  - ◇ For every $x \in F$, $\exists\, 0 \in F \mid x + 0 = x$ (*Additive Identity*)
  - ◇ For every $x \in F \setminus \{0\}$, $\exists\, 1 \in F \mid x \cdot 1 = x$ (*Multiplicative Identity*)

  5 Existence of Inverse:
  - ◇ For every $x \in F$, $\exists\, y \in F$ such that $x + y = 0$. This y is represented as $-x$, and is called *additive inverse of* x or *negative of* x.
  - ◇ For every $x \in F \setminus \{0\}$, $\exists\, y \in F \setminus \{0\}$ such that $x \cdot y = 1$. This y is represented as x^{-1} or $\frac{1}{x}$, and is called *multiplicative inverse of* x or *reciprocal of* x.

  > *e.g.,* $\{0, 1\}$ (binary system), $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are examples of a field.

- **Vector Space:** It is a set (V) over the field F with 2 binary operations, $+ : V \times V \to V$ (*addition*), and $\cdot : F \times V \to V$ (*scaling*) defined on it, which satisfies the following axioms:
  1 For Addition:
  - ◇ $\boldsymbol{u} + \boldsymbol{v} \in V$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$ (*Closure*)
  - ◇ $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ (*Commutation*)
  - ◇ $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ (*Association*)
  - ◇ For every $\boldsymbol{u} \in V$, $\exists\, \mathbf{0} \in F \mid \boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$ (*Additive Identity*)
  - ◇ For every $\boldsymbol{u} \in V$, $\exists\, \boldsymbol{v} \in V$ such that $\boldsymbol{u} + \boldsymbol{v} = \mathbf{0}$. This $\boldsymbol{v}$ is represented as $-\boldsymbol{u}$ (*Additive Inverse* or *Negative*)

  2 For Scaling: $(a, b \in F$ and $\boldsymbol{u}, \boldsymbol{v} \in V)$
  - ◇ $a \cdot \boldsymbol{u} \in V$ for all $a \in F$, $\boldsymbol{u} \in V$ (*Closure*)
  - ◇ $(a \cdot b) \cdot \boldsymbol{u} = a \cdot (b \cdot \boldsymbol{u})$ for all $a, b \in F$ and $\boldsymbol{u} \in V$ (*Association*)
  - ◇ $a \cdot (\boldsymbol{u} + \boldsymbol{v}) = a \cdot \boldsymbol{u} + a \cdot \boldsymbol{v}$ for all $a \in F$ and $\boldsymbol{u}, \boldsymbol{v} \in V$ (*Distribution*)
  - ◇ $(a + b) \cdot \boldsymbol{u} = a \cdot \boldsymbol{u} + b \cdot \boldsymbol{u}$ for all $a, b \in F$ and $\boldsymbol{u} \in V$ (*Distribution*)
  - ◇ $\exists\, 1 \in F \mid 1 \cdot \boldsymbol{u} = \boldsymbol{u}$ for all $\boldsymbol{u} \in V$ (*Scaling Identity*)

  > *e.g.,* $\{\mathbf{0}\}$ over F (Null Space); $\mathbb{R}^n$ over $\mathbb{R}$, $\mathbb{C}$ over $\mathbb{R}$ are examples of vector space.

- **Vector Subspace:** Vector subspace W is a subset of *vector space* V, which follows the following axioms:
  - ◇ $\boldsymbol{u} + \boldsymbol{v} \in W$ for all $\boldsymbol{u}, \boldsymbol{v} \in W$
  - ◇ $a \cdot \boldsymbol{u} \in W$ for all $a \in F$, $\boldsymbol{u} \in W$
  - ◇ $\mathbf{0} \in W$, where $\mathbf{0}$ is zero element (*additive identity* of V)

- **Sum and Direct Sum:** Let W_1 and W_2 be two vector subspace of Vector Space V. Then

$$W_1 + W_2 := \{w_1 + w_2 \mid w_1 \in W_1 \text{ and } w_2 \in W_2\} \text{ is Sum; and}$$
$$W_1 \oplus W_2 := W_1 + W_2 \text{ when } W_1 \cap W_2 = \{\mathbf{0}\} \text{ is Direct Sum.}$$

- **Polynomial Vector Spaces:** A polynomial with real coefficients, and degree at most n, denoted as $\mathcal{P}_n(\mathbb{R})$, is a vector space over $\mathbb{R}$. Note that

$$\mathcal{P}_0(\mathbb{R}) \subset \mathcal{P}_1(\mathbb{R}) \subset \mathcal{P}_2(\mathbb{R}) \subset \cdots \subset \mathcal{P}_n(\mathbb{R}) \subset \mathcal{P}_{n+1}(\mathbb{R}) \subset \cdots$$

  Thus, we define

$$\mathcal{P}(\mathbb{R}) := \bigcup_{n \geq 0} \mathcal{P}_n(\mathbb{R})$$

  $\mathcal{P}(\mathbb{R})$ is a vector space over $\mathbb{R}$.
  Interesting Subspaces of $\mathcal{P}(\mathbb{R})$:
  - **Polynomial with at most n-degree:** $\mathcal{P}_n(\mathbb{R})$, is a vector subspace of vector space $\mathcal{P}(\mathbb{R})$.
  - **Even Polynomials:** $E(\mathbb{R}) := \left\{p(x) := \sum_{i=0}^{k} a_{2i}x^{2i} \,|\, a_i \in \mathbb{R}, k \in \mathbb{Z}_{\geq 0}\right\}$ is also a vector subspace of $\mathcal{P}(\mathbb{R})$
  - **Truncated Polynomials:** For $k \in \mathbb{N}$, the subset

$$T_k := \left\{p(x) := \sum_{i=k+1}^{n} a_i x^i \,|\, a_i \in \mathbb{R}, n \geq k+1\right\}$$

    is a vector subspace of polynomial vector space.
- **Matrix Vector Space:** A *matrix* is a vector and the vector space (set) of a general matrices is denoted as $M_{m\times n}(F)$, where F is the field.
  Interesting Subspaces of $M_{n\times n}(\mathbb{R})$, also denoted simply as $M_n(\mathbb{R})$:
  - **Symmetric Matrices:**

$$\mathrm{Sym}_n := \{A \in M_n(\mathbb{R}) \mid A = A^T\}$$

  - **Traceless Matrices:**

$$W_n := \{A \in M_n(\mathbb{R}) \mid \mathrm{trace}(A) := a_{11} + \cdots + a_{nn} = 0\}$$

  - **Diagonal Matrices:**

$$D_n = \{A \in M_n(\mathbb{R}) \mid a_{ij} = 0 \text{ if } i \neq j\}$$

  - **Scalar Matrices:** (Also a *Field*)

$$S_n = \{\lambda I_n \in M_n(\mathbb{R}) \mid \lambda \in \mathbb{R}\}$$

Span, Linear Dependence & Independence:

- **Span:** Given a subet S of vector space V, the span is defined as

$$\mathrm{span}(S) := \{\boldsymbol{v} \in V \mid \boldsymbol{v} = \sum_{i=1}^{k} c_i \boldsymbol{v_i} \text{ for some } \boldsymbol{v_i} \in S \text{ and } c_i \in F\}$$

  Any vector in the span(S) can be expressed as a finite linear combination of vectors in S.
  - If S be a subset of a vector space V, then span(S) is a vector subspace containing S. Conversely, if W is a subspace containing S, then W contains span(S).
  - $\mathrm{span}(\emptyset) = \{\mathbf{0}\}$ (*Convention*)
  - $\mathrm{span}(\mathrm{span}(S)) = \mathrm{span}(S)$
  - Let W be a subspace of vector space V, then $\mathrm{span}(W) = W$ and $\mathrm{span}(V) = V$
  - If $S_1 \subseteq S_2$, then $\mathrm{span}(S_1) \subseteq \mathrm{span}(S_2)$
  - $\mathrm{span}(S_1 \cup S_2) = \mathrm{span}(S_1) + \mathrm{span}(S_2)$
  - $\mathrm{span}(S_1 \cap S_2) \subseteq \mathrm{span}(S_1) \cap \mathrm{span}(S_2)$

- **Linear Dependence and Independence:** A subset S of vector space V over field F is called *linearly dependent* if for $\boldsymbol{v_1}, \ldots, \boldsymbol{v_n} \in S$ and $c_1, \ldots, c_n \in F$ (not all zero), such that

$$\sum_{i=1}^{n} c_i \boldsymbol{v_i} = \mathbf{0}$$

If set S is not linearly dependent, then they are *linearly independent*, *i.e.*, if S is linearly independent, then

$$\sum_{i=1}^{n} c_i \boldsymbol{v_i} = \mathbf{0} \quad \Rightarrow \quad c_i = 0 \quad \forall\, i \in \{j\}_{j=1}^{n}$$

  - ◇ Let S be a linearly dependent subset of a vector space V. Then $\exists\ \boldsymbol{v} \in S$ such that $\text{span}(S) = \text{span}(S \setminus \{\boldsymbol{v}\})$.
  - ◇ Let S be a linearly independent subset of a vector space V. Given $\boldsymbol{v} \in V \setminus S$, the set $S \cup \{\boldsymbol{v}\}$ is linearly independent if and only if $\boldsymbol{v} \notin \text{span}(S)$.
  - ◇ Any subset of a linearly independent set is linearly independent.

Basis & Dimensions, Replacement Theorem:

- **Basis:** A basis β of a vector space V is a subset such that β is linearly independent and spans V.
  - ◇ Let $\beta = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$ be a basis for V. Then every vector $\boldsymbol{v} \in V$ can be expressed uniquely as $\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{v_i}$ for scalars a_i.
  - ◇ Let V be a vector space that admits a finite subset S that spans V. Then $\exists\ \beta \subseteq S$ which is a basis for V.

Note: *Standard Basis* for *Euclidean Space* ($\mathbb{R}^n$)

$$\beta = \{\, \underbrace{(1,0,0,\ldots,0)}_{\boldsymbol{e_1}}, \underbrace{(0,1,0,\ldots,0)}_{\boldsymbol{e_2}}, \ldots, \underbrace{(0,0,0,\ldots,1)}_{\boldsymbol{e_n}} \,\}$$

- **Dimensions:** If a vector space V have finite elements in *basis*, then the vector space is said to have *finite dimensions*. The size (cardinality) of this finite basis is called it's dimension, denoted as $\dim_F(V)$.
  Vector space V have infinite *dimensions*, if it do not have finite *basis*.

- **Replacement Theorem:** Let V be a vector space that is spanned by S of size n. Let L be a linearly independent set of size m. Then
  1. $m \leq n$;
  2. $\exists\ T \subseteq S$ of size $n - m \mid T \cup L$ spans V

  Corollaries:
  - ◇ Let V be a vector space with a *finite basis*. Then any two *basis* have same size. This common integer is called the dimensions of V (denoted as $\dim_F(V)$).
  - ◇ Let V be a vector space with $\dim_F(V) = n$. Then
    1. Any $S \subseteq V \;\big|\; \text{span}(S) = V$, then $|S| \geq n$.[1] If $|S| = n$, then S is *basis*.
    2. Any $S \subseteq V \;\big|\; |S| = n$ and S is linearly independent, then S is a *basis*.
    3. Any *linearly independent* set can be extended to basis.
  - ◇ Let V be a finite dimensional vector space and W be a subspace. Then
    1. $\dim_F(W) \leq \dim_F(V)$.
    2. If $\dim_F(W) = \dim_F(V)$, then $W = V$.

[1] $|S|$ is *cardinality* of set S

Note: Let W_1 and W_2 be vector subspaces of V. Then

- $\dim_F(W_1 + W_2) = \dim_F(W_1) + \dim_F(W_2) - \dim_F(W_1 \cap W_2)$
  If $W_1 \cap W_2 = \{\mathbf{0}\}$, then $W_1 + W_2 \equiv W_1 \oplus W_2$ and
  $$\dim_F(W_1 \oplus W_2) = \dim_F(W_1) + \dim_F(W_2)$$
- Conversely, if $\dim_F(W_1 + W_2) = \dim_F(W_1) + \dim_F(W_2)$, then $W_1 \cap W_2 = \{\mathbf{0}\}$

Quotient Space:

- **Quotient Space:** Let V be a vector space over field F, and W be a subspace of V. Now, lets define a relation $\sim_W$ on V as:
  $$\boldsymbol{v_1} \sim_W \boldsymbol{v_2} \Leftrightarrow \boldsymbol{v_1} - \boldsymbol{v_2} \in W$$
  Note that $\sim_W$ is an *equivalence relation* (Check). Now, the *equivalence class* of $\sim_W$ is
  $$[\boldsymbol{v}] = \boldsymbol{v} + W := \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{w} \in W\}$$
  And, the set of all equivalence classes (*Quotient Set under* $\sim_W$) is
  $$V/W := \{[\boldsymbol{v}] \mid \boldsymbol{v} \in V\}$$
  - ◇ **Theorem:** The set V/W (*pron.*, "V mod W") is a vector space over F, with *addition* and *scaling* operation defined as
    1. $+ : V/W \times V/W \to V/W, \quad [\boldsymbol{v}_1] + [\boldsymbol{v_2}] = [\boldsymbol{v_1} + \boldsymbol{v_2}]$
    2. $\cdot : F \times V/W \to V/W, \quad \lambda \cdot [\boldsymbol{v}] = [\lambda \boldsymbol{v}]$
  - ◇ $[\mathbf{0}] \equiv W$ (*Additive Identity* of V/W)
  - ◇ $\dim_F(V/W) = \dim_F(V) - \dim_F(W)$

Linear Maps, Rank-Nullity Theorem, Linear Isomorphism:

- **Linear Maps:** A map $T : V \to W$ between vector spaces (over a field F) is called a *Linear Transformation* or a *Linear Map* if
  1. $T(\boldsymbol{v_1} + \boldsymbol{v_2}) = T(\boldsymbol{v_1}) + T(\boldsymbol{v_2}) \ \forall\, \boldsymbol{v_1}, \boldsymbol{v_2} \in V$;
  2. $T(c\,\boldsymbol{v}) = c\,T(\boldsymbol{v})$ for any $c \in F$ and $\boldsymbol{v} \in V$

  Note: $T(\boldsymbol{0_V}) = \boldsymbol{0_W}$ and $T(\sum_{i=1}^n c_i \boldsymbol{v_i}) = \sum_{i=1}^n c_i T(\boldsymbol{v_i})$

  Examples:
  - ◇ **Scaling:** $T : \mathbb{R} \to \mathbb{R} \mid T(x) = c\,x$
  - ◇ **Dilation:** $T : V \to V \mid T(\boldsymbol{v}) = c\,\boldsymbol{v}$
  - ◇ **Identity:** $I : V \to V \mid I(\boldsymbol{v}) = \boldsymbol{v}$
  - ◇ **Trivial:** $O : V \to W \mid T(\boldsymbol{v}) = \boldsymbol{0_W}$
  - ◇ **Matrices:** $A \in M_{m\times n}(\mathbb{R})$ is a linear map $L_A : \mathbb{R}^n \to \mathbb{R}^m$
  - ◇ **Reflection:** Reflection about some axis in $\mathbb{R}^2$ is $T : \mathbb{R}^2 \to \mathbb{R}^2$
    (i) About x-axis: $T(x, y) = (x, -y)$
    (ii) About y-axis: $T(x, y) = (-x, y)$
  - ◇ **Rotation:** Rotation (anti-clock wise) about origin in $\mathbb{R}^2$ by angle θ is $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ defined as
    $$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
  - ◇ **Projection:** $T : \mathbb{R}^2 \to \mathbb{R}^2$ is projection to
    (i) x-axis: $T(x, y) = T(x, 0)$
    (ii) y-axis: $T(x, y) = T(0, y)$

- ⋄ **Inclusion:** $T : \mathbb{R}^m \to \mathbb{R}^n$ when $m \le n$
- ⋄ **Differentiation:** You already know this :)
- ⋄ **Integration:** You already know this, too :)
- ⋄ **Linear Combination:** Given $\boldsymbol{v_1}, \dots, \boldsymbol{v_n} \in V$, where V is vector space over $\mathbb{R}$, we have

$$T : \mathbb{R}^n \to V \mid T(a_1, \dots, a_n) = \sum_{i=1}^{n} a_i \boldsymbol{v_i}$$

- ⋄ **Quotient:** Given a vector space V and its subspace W, we consider the quotient map

$$Q : V \to V/W, \;\; Q(\boldsymbol{v}) := [\boldsymbol{v}] = \boldsymbol{v} + W$$

- **Null Space and Range:** Let $T : V \to W$ be a linear map, then

$$\begin{aligned} N_{(T)} &:= \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0_W}\} && \text{is Null Space; and} \\ R_{(T)} &:= \{T(\boldsymbol{v}) \in W \mid \boldsymbol{v} \in V\} && \text{is Range.} \end{aligned}$$

Note: $N_{(T)}$ is a subspace of vector space V (convince yourself)

Nullity: $\text{nullity}(T) = \dim_F(N_{(T)})$

Rank: $\text{rank}(T) = \dim_F(R_{(T)})$

- ⋄ **Rank-Nullity Theorem:** Let V be a finite dimensional vector space. If $T : V \to W$ is a linear map, then

$$\text{rank}(T) + \text{nullity}(T) = \dim_F(V)$$

- ⋄ Let $T : V \to W$ be a linear map and suppose that $\dim(V) = \dim(W) < \infty$, then following statement are equivalent:
  1. T is one-to-one;
  2. T is onto.

- **Set of Linear Maps:** For given vector space V and W over the field F, we have

$$\mathcal{L}_{(V,W)} := \{T : V \to W \mid T \text{ is linear map}\}$$

For $S, T \in \mathcal{L}_{(V,W)}$, we define
1. $(S + T)(\boldsymbol{v}) = S(\boldsymbol{v}) + T(\boldsymbol{v})$; and
2. $(cT)(\boldsymbol{v}) = c \cdot T(\boldsymbol{v})$, where $c \in F$.

then, $(S + T), (cT) \in \mathcal{L}_{(V,W)}$

Thus, $\mathcal{L}_{(V,W)}$ is also a *Vector Space* over F.

Note: If $\dim_F(V) = n$ and $\dim_F(W) = m$, then $\dim_F(\mathcal{L}_{(V,W)}) = mn$. *Hint:* Let $\beta = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ and $\gamma = \{\boldsymbol{w_1}, \dots, \boldsymbol{w_m}\}$ be basis for V and W respectively. Then

$$E_{p,q}(\boldsymbol{v_i}) = \begin{Bmatrix} 0 & \text{if } i \neq q \\ \boldsymbol{w_p} & \text{if } i = q \end{Bmatrix} = \delta_{iq} \boldsymbol{w_p}$$

- **Linear Isomorphism:** A linear map $T : V \to W$ is called a *linear isomorphism* if T is *one-to-one* and *onto.* (T is a *bijection*)

We say V and W are *isomorphic* (as vector spaces) if there exists a *linear isomorphism* $T : V \to W$.

- ⋄ Isomorphism $T : V \to W$ guarantees that T^{-1} exists. Therefore, isomorphism is a symmetric relation between vector space V and W.

Ordered Basis & Coordinate Vector:

- **Ordered Basis:** Let V be a finite dimensional vector space. An ordered basis of V is a basis β of V, endowed with a specific order.

For instance, $\{\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}\}$ and $\{\boldsymbol{e_3}, \boldsymbol{e_2}, \boldsymbol{e_1}\}$ are both same as bases of $\mathbb{R}^3$ but different as ordered basis.

- Let $T : V \to W$ be a linear map.
  - ◇ Let $\{\boldsymbol{w_1}, \dots, \boldsymbol{w_n}\} \subset W$ be a linearly independent set and let T be surjective. If $\boldsymbol{v_i} \in V$ such that $T(\boldsymbol{v_i}) = \boldsymbol{w_i}$, then $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ is linearly independent.
  - ◇ Let $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\} \subset V$ be a linearly independent set. If T is an injective map, then $\{T(\boldsymbol{v_1}), \dots, T(\boldsymbol{v_n})\}$ is linearly independent.
  - ◇ Let T be a linear isomorphism. If $\beta = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ is a basis set for V, then $\{T(\boldsymbol{v_1}), \dots, T(\boldsymbol{v_n})\}$ is a basis for W.
  - ◇ Any two vector spaces of the same finite dimension are isomorphic.
    Corollary: Any vector space V of dimension n over $\mathbb{R}$ is isomorphic to $\mathbb{R}^n$.
- Let $T : V \to W$ be a linear map, with Null Space $N_{(T)}$ and Range $R_{(T)}$, then T is a surjective map from V to $R_{(T)}$. Lets define linear map

$$\mathcal{T} : V/N_{(T)} \to R_{(T)} \ \mid \ \forall\, \boldsymbol{v} \in V/N_{(T)}, \mathcal{T}(\boldsymbol{v}) = T(\boldsymbol{v})$$

This $\mathcal{T}$ is well-defined, injective, surjective. Thus, $\mathcal{T} : V/N_{(T)} \to R_{(T)}$ is a *Linear Isomorphism.*

- **Coordinate Vectors:** Let $\beta = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ be an ordered basis of a vector space V. Any $\boldsymbol{v} \in V$ can be expressed uniquely as $\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{v_i}$. Then

$$[\boldsymbol{v}]^\beta = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

is called the *coordinate vector* of $\boldsymbol{v}$ relative to β

Matrix Representation of Linear Transformation:

- **Matrix Representation:** Let $T \in \mathcal{L}_{(V,W)}$, and let $\beta = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ and $\gamma = \{\boldsymbol{w_1}, \dots, \boldsymbol{w_m}\}$ be ordered basis of V and W respectively. We express

$$T(\boldsymbol{v_j}) = \sum_{i=1}^{m} a_{ij} \boldsymbol{w_i}$$

for unique scalars a_{ij}.

Say $\boldsymbol{v} = \displaystyle\sum_{i=1}^{n} x_i \boldsymbol{v_i}$, then

$$\begin{aligned} T(\boldsymbol{v}) &= \sum_{j=1}^{n} x_j T(\boldsymbol{v_j}) \\ &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} \boldsymbol{w_i} \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \boldsymbol{w_i} \end{aligned}$$

Therefore,

$$[\boldsymbol{v}]^\beta = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \Rightarrow \quad [T(\boldsymbol{v})]^\gamma = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Or simply, $[T(\boldsymbol{v})]^\gamma = A[\boldsymbol{v}]^\beta$

The $m \times n$ matrix A associated to $T : V \to W$ is called the *matrix representation* of T with respect to β and γ.

Notation: $A = [T]_\beta^\gamma$. Therefore $[T(\boldsymbol{v})]^\gamma = [T]_\beta^\gamma [\boldsymbol{v}]^\beta$

If $\beta = \gamma$, then A is denoted by $[T]_\beta$

- ⋄ Let V and W to be two vector spaces of same dimensions. Let β and β' be two different ordered basis of V, and γ and γ' be two different ordered basis for W. Let $T \in \mathcal{L}_{(V,W)}$, which have matrix representation as $[T]_\beta^\gamma$ and $[T]_{\beta'}^{\gamma'}$, then
  1. $\det([T]_\beta^\gamma) = \det([T]_{\beta'}^{\gamma'})$
  2. $\operatorname{tr}([T]_\beta^\gamma) = \operatorname{tr}([T]_{\beta'}^{\gamma'})$
- ⋄ Let V and W be vector spaces with $\dim_F V = n$ and $\dim_F W = m$ over some field F, and let β and γ be their ordered basis respectively. then, $\mathcal{L}_{(V,W)}$ is *isomorphic* to $M_{m\times n}(F)$ with *linear isomorphism* defined as

$$\Phi : \mathcal{L}_{(V,W)} \to M_{m\times n}(F), \;\; \Phi(T) = [T]_\beta^\gamma$$

Dual Spaces, Dual Basis & Dual Maps:

- **Dual Spaces:** Let V be a vector space over field F (generally $\mathbb{R}$). Then the dual of V is defined to be $\mathcal{L}_{(V,F)}$, represented as V^*
  - ⋄ $\dim_F(V) = \dim_F(V^*)$
  - ⋄ **Linear Functionals:** Any $f \in \mathcal{L}_{(V,F)}$.
  - ⋄ **Dual Basis:** Let V be a vector space with ordered basis $\beta = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$. Let $\boldsymbol{v} \in V$ as $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v_i}$. Consider $\boldsymbol{v_i^*} : V \to \mathbb{R}$ such that $\sum_{i=1}^n a_i \boldsymbol{v_i} \mapsto a_i$. Then set $\beta^* := \{\boldsymbol{v_1^*}, \ldots, \boldsymbol{v_n^*}\}$ form the basis for V^*. This set β^* is called *Dual Basis.*
  - ⋄ $\boldsymbol{v_i^*}(v_j) = \delta_{ij}$
- **Dual Maps**. Given any $T \in \mathcal{L}_{(V,W)}$, the map T^* defined as

$$T^* : W^* \to V^*, \;\; L \mapsto L \circ T$$

  is called the *dual* of T. (You can check that T^* is a Linear Map ...)
  - ⋄ Let $T \in \mathcal{L}_{(V,W)}$ and $T^* : W^* \to V^*$ be it's *dual.* If β and γ are the ordered basis of V, and W respectively, then

$$([T]_\beta^\gamma)^t = [T^*]_{\gamma^*}^{\beta^*}$$

    where β^* and γ^* are the *dual basis* of V^* and W^* respectively.
  - ⋄ Let U, V, W and X be vector spaces and α, β, γ and δ be the ordered basis for them respectively. Let $R : U \to V,\ S : V \to W$ and $T : W \to X$ be the linear maps, then
    1. $[T \circ S]_\beta^\delta = [T]_\gamma^\delta [S]_\beta^\gamma$
    2. $([T]_\gamma^\delta [S]_\beta^\gamma)[R]_\alpha^\beta = [T]_\gamma^\delta([S]_\beta^\gamma [R]_\alpha^\beta)$
  - ⋄ Let $T \in \mathcal{L}_{(V)}$ and Let β and γ be two ordered basis for V of size n, then the following hold:
    1. $[I_V]_\beta^\gamma [I_V]_\gamma^\beta = [I_V]_\gamma^\gamma = I_n$
    2. $[I_V]_\gamma^\beta [T]_\gamma^\gamma = [T]_\beta^\gamma$
    3. $[T]_\gamma^\gamma [I_V]_\beta^\gamma = [T]_\beta^\gamma$
  - ⋄ $[T]_\beta^\gamma = (\ [T^*]_{\gamma^*}^{\beta^*})^t$

Change of Basis & Invertibility:

- **Change of Basis:** The matrix $[I_V]_\beta^\gamma$ associated to the identity map $I_V : V \to V$ and two ordered bases β, γ of V is called the *change of coordinate/basis* matrix.
- **Inverse Maps:** A linear map $T : V \to W$ is called *invertible* if there exists $S : W \to V$ such that $S \circ T = I_V$ (injectivity) and $T \circ S = I_W$ (surjectivity).
- **Inverse Matrix:** A matrix $A \in M_{m\times n}(\mathbb{R})$ is called invertible if there exists a matrix $B \in M_{n\times m}(\mathbb{R})$ such that

$$AB = I_m, \quad BA = I_n$$

  If a matrix $A \in M_{m\times n}(\mathbb{R})$ is invertible, then $m = n$.

◇ If $S : U \to V$ and $T : V \to W$ are linear invertible maps, then
1. $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$
2. $(T^{-1})^{-1} = T$

◇ Let V and W be finite dimensional vector spaces with ordered basis β and γ respectively. If $T \in \mathcal{L}_{(V,W)}$, then T is invertible iff $[T]_\beta^\gamma$ is invertible. Moreover, we have
$$[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$$

Similarity, Trace & Rank:

- **Similarity:** A matrix $A \in M_n(\mathbb{R})$ is said to be similar to $B \in M_n(\mathbb{R})$ if $\exists\ Q \in M_n(\mathbb{R})$ such that
$$QAQ^{-1} = B$$
This Similarity Relation $(\sim_S)$ is an equivalence relation, since
1. *Reflexive:* $I_n A I_n^{-1} = A$
2. *Symmetric:* If $QAQ^{-1} = B$, then $Q^{-1}BQ = A$
3. *Transitive:* If $QAQ^{-1} = B$ and $PBP^{-1} = C$ then
$$(PQ)A(PQ)^{-1} = PQAQ^{-1}P^{-1} = PBP^{-1} = C$$

◇ Given two similar matrices A and B, Q is not unique, since $(\lambda Q)A(\lambda Q)^{-1} = B$ for any non-zero scalar λ.

◇ Given $T \in \mathcal{L}_{(V)}$, and β and γ are ordered basis of V, then $[T]_\beta$ and $[T]_\gamma$ are similar via the change of basis matrix $(Q = [I_V]_\beta^\gamma)$

- **Trace:** For and square matrix $A \in M_n(F)$, the trace is defined as sum of diagonal elements, *i.e.*,
$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$
Let $T \in \mathcal{L}_{(V)}$ where V has n dimension and let β is ordered basis of V. Then, the trace of T is defined as $\mathrm{tr}([T]_\beta)$

◇ If $A, B, C \in M_n(F)$, then
$$\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$$

◇ Using above property, we get that
$$\mathrm{tr}(QAQ^{-1}) = \mathrm{tr}(A)$$

◇ If β and γ are two ordered basis of V and $T \in \mathcal{L}_{(V)}$, then
$$\mathrm{tr}([T]_\gamma) = \mathrm{tr}([T]_\beta)$$

- **Rank:** For $T \in \mathcal{L}_{(V,W)}$, the *rank* of T is defined as dimension of range T, *i.e.*,
$$\mathrm{rank}(T) := \dim_F(R_{(T)})$$
Let $\beta = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$ be ordered basis of V. Then $T(\beta) = \{T(\boldsymbol{v_1}), \ldots, T(\boldsymbol{v_n})\}$ spans W. By Replacement Theorem, a subset of $T(\beta)$ is a basis for $R_{(T)}$. By rearranging indices, we may assume that $\{T(\boldsymbol{v_1}), \ldots, T(\boldsymbol{v_k})\}$ is basis of $R_{(T)}$.

Choose an ordered basis γ of W and consider $[T]_\beta^\gamma$. The j^{th} column of this matrix is the coordinate vector $[T(\boldsymbol{v_j})]^\gamma$.

◇ The first k columns $\boldsymbol{c_1}, \ldots, \boldsymbol{c_k}$ are linearly independent in F^m

◇ $\mathrm{rank}(T)$ = column rank of $[T]_\beta^\gamma$ = row rank of $[T]_\beta^\gamma$

◇ As $[T]_\beta^\gamma = (\ [T^*]_{\gamma^*}^{\beta^*})^t$, we conclude that
$$\mathrm{rank}(T) = \text{row rank of } [T]_\beta^\gamma = \text{column rank of } [T^*]_{\gamma^*}^{\beta^*} = \mathrm{rank}(T^*)$$

◇ Let $P \in M_m(\mathbb{R}), Q \in M_n(\mathbb{R})$ and $A \in M_{m\times n}(\mathbb{R})$. If P is *injective* and Q is *surjective*, then $\mathrm{rank}(PAQ) = \mathrm{rank}(A)$.

◇ *Corollary:* Let $T \in \mathcal{L}_{(V,W)}$ for finite dimensional V and W. Let β, β' be ordered basis of V and γ, γ' be ordered basis for W. Then
$$\mathrm{rank}([T]_\beta^\gamma) = \mathrm{rank}([T]_{\beta'}^{\gamma'})$$

Determinants:

- **Minors and Co-factors:** For $A \in M_n(F)$, the minor $\tilde{A}_{ij}$ of A associated to the i^{th} row and j^{th} column is the $(n-1) \times (n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column of A. Minor multiplied by the factor $(-1)^{i+j}$, *i.e.*, $(-1)^{i+j}\tilde{A}_{ij}$, is called the co-factor associated with i^{th} row and j^{th} column
- **Determinant:** Let $A \in M_n(F)$, then the determinant is a function $\det : M_n(F) \to F$, which is defined as follows:
  1. For $n = 1$, the determinant of the matrix is the entry itself, *i.e.*, if $A = [a_{11}]$, then $det(A) = a_{11}$
  2. For $n = 2$, the determinant is defined as
  $$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ then } \det(A) = a_{11}a_{22} - a_{12}a_{21}$$
  3. For a general n, the determinant is defined recursively as co-factor expansion along any row or column, which is
  $$\det(A) = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det(\tilde{A}_{ij}) \quad \text{(co-factor expansion along } i^{th} \text{ row)}$$
  We can also write,
  $$\det(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\det(\tilde{A}_{ij}) \quad \text{(co-factor expansion along } j^{th} \text{ column)}$$

We shall think of $A \in M_n(F)$ as an ordered collection of n column vectors, *i.e.*,

$$A = (\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \ \cdots\ \ \boldsymbol{v_n}\ )$$

Key properties of Determinants:

- ⋄ Determinant is a linear function of any column (or row), when the other columns (or rows) are fixed.
$$\begin{aligned}&\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_{j-1}}\ \ \boldsymbol{v_j} + \lambda\,\boldsymbol{w_j}\ \ \boldsymbol{v_{j+1}}\ \cdots\ \boldsymbol{v_n}\ ) = \\ &\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_{j-1}}\ \ \boldsymbol{v_j}\ \ \boldsymbol{v_{j+1}}\ \cdots\ \boldsymbol{v_n}\ ) + \lambda\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_{j-1}}\ \ \boldsymbol{w_j}\ \ \boldsymbol{v_{j+1}}\ \cdots\ \boldsymbol{v_n}\ )\end{aligned}$$
- ⋄ If $A \in M_n(F)$ has one row or column zero, then $\det(A) = 0$
$$\det(\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \cdots\ \boldsymbol{v_n}\ ) = 0 \quad \text{if } \boldsymbol{v_i} = \boldsymbol{0} \text{ for any } i$$
- ⋄ If $A \in M_n(F)$ has two rows or columns identical, then $\det(A) = 0$
$$\det(\ \boldsymbol{v_1}\ \ \boldsymbol{v_2}\ \cdots\ \boldsymbol{v_n}\ ) = 0 \quad \text{if } \boldsymbol{v_i} = \boldsymbol{v_j} \text{ for some } i \neq j$$
- ⋄ If B is obtained from A by interchanging two distinct rows or columns of A, then $\det(B) = -\det(A)$
$$\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_i}\ \cdots\ \boldsymbol{v_j}\ \cdots\ \boldsymbol{v_n}\ ) = -\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_j}\ \cdots\ \boldsymbol{v_i}\ \cdots\ \boldsymbol{v_n}\ )$$
- ⋄ Determinant is unchanged by adding a multiple of another column (resp. row) to a column (resp. row).
$$\det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_i}\ \cdots\ \boldsymbol{v_j}\ \cdots\ \boldsymbol{v_n}\ ) = \det(\ \boldsymbol{v_1}\ \cdots\ \boldsymbol{v_i} + \lambda\,\boldsymbol{v_j}\ \cdots\ \boldsymbol{v_j}\ \cdots\ \boldsymbol{v_n}\ )$$

Important Results:

- ◇ If E is an elementary matrix, then

$$\det(EA) = \det(E)\det(A) = \det(AE)$$

- ◇ A matrix $A \in M_n(F)$ is invertible if and only if $\det(A) \neq 0$
  The restriction of determinant to $GL_n(F)$, the set of invertible matrices, defines a *homomorphism.*
  It follows that if A is invertible, then

$$\det(A^{-1}) = \det(A)^{-1}$$

  Also,

$$\det(CAC^{-1}) = \det(A)$$

- **Determinant of a Linear Map:** Let $T \in \mathcal{L}_{(V)}$ and V be finite dimensional. The determinant of T is defined to be the determinant of $[T]_\beta$ for any choice of ordered basis β.
  If β and γ are two ordered bases, then $[T]_\beta$ and $[T]_\gamma$ are *conjugate* to each other, via the change of basis matrix. As determinant is conjugation invariant, we have

$$\det([T]_\beta) = \det([T]_\gamma)$$

  Thus, the notion of determinant of a linear map $T : V \to V$ is well defined.

Diagonalization:

- **Diagonalizable:** A matrix $A \in M_n(F)$ is called diagonalizable if A is similar to a diagonal matrix, *i.e.*, there exists an invertible matrix $C \in M_n(F)$ such that CAC^{-1} is diagonal matrix.
  A linear map $T : V \to V$ is called diagonalizable if $[T]_\beta$ is diagonal for some basis β.
- If γ is another basis of V, then $[T]_\gamma$, being similar to $[T]_\beta$, is diagonalizable as a matrix. Conversely, suppose $[T]_\gamma$ is diagonalizable for some basis γ, then $Q[T]_\gamma Q^{-1}$ is diagonal for some matrix Q. Use Q to define a basis β such that $[T]_\beta$ is diagonal. Thus, $T : V \to V$ is diagonalizable if $[T]_\gamma$ is diagonalizable for any basis γ.

Eigenvectors, Eigenvalues & Eigenspaces:

- **Eigenvector:** A non-zero vector $\boldsymbol{v} \in V$ is called an eigenvector of $T : V \to V$ if $T(\boldsymbol{v}) = \lambda\boldsymbol{v}$ for some scalar $\lambda \in F$. The scalar λ is called the **eigenvalue** of the eigenvector $\boldsymbol{v}$.
  A non-zero vector $\boldsymbol{v} \in F^n$ is called an eigenvector of $A \in M_n(F)$ if $A\boldsymbol{v} = \lambda\boldsymbol{v}$ for some $\lambda \in F$. The scalar λ is called the **eigenvalue** of the eigenvector $\boldsymbol{v}$.
  - ◇ **Characterization of Eigenvalues:** The following conditions are equivalent:
    1. λ is an eigenvalue for T;
    2. $T - \lambda I_V$ is not invertible;
    3. $\det(T - \lambda I_V) = 0$.
- **Eigenspace:** Let $T : V \to V$ be a linear map. If λ is an eigenvalue, then its eigenspace is defined as

$$E_\lambda := \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \lambda\boldsymbol{v}\}$$

  Note that $E_\lambda \neq \emptyset$ and it is a vector subspace.
  - ◇ If T is diagonalizable, with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ then there exists a basis $\beta = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$ consisting of eigenvectors. Such a basis is called an **eigenbasis**.

Characteristic & Minimal Polynomials:

- **Characteristic Polynomial:** For $A \in M_n(F)$ the polynomial
$$p_A(x) = \det(A - xI_n)$$
is called the characteristic polynomial of A.
For a linear map $T : V \to V$, the polynomial $p_T(x) = \det(T - xI_V)$ is called the characteristic polynomial of T.

- **Minimal Polynomial:** Given $A \in M_n(F)$, consider the set
$$S = \{I_n, A, A^2, \ldots, A^{n^2}\}$$
There are two mutually exclusive and exhaustive cases:
Case 1: Either two elements in S are the same, *i.e.*, $A^k = A^l$ for some $0 \le k < l \le n^2$. Then A satisfies the equation $x^k = x^l$.
Case 2: S is of size $n^2 + 1$. As $M_n(F)$ has dimension n^2, the set S is linearly dependent. There exists $a_i \in F$ such that
$$a_0 I_n + a_1 A + \cdots + a_{n^2} A^{n^2} = 0_n$$
Thus, A satisfies the polynomial equation $a_0 + a_1 x + \cdots + a_{n^2} x^{n^2} = 0$
As any matrix satisfies some polynomial, for each $A \in M_n(F)$, there will be a unique monic polynomial of the smallest degree which A satisfies. This polynomial is called the minimal polynomial of A.
If $A = [T]_\beta$ for some $T \in \mathcal{L}_{(V)}$ and β, then
$$c_0 I_n + c_1 A + \cdots + c_k A^k = 0_n \quad \Leftrightarrow \quad c_0 I_V + c_1 T + \cdots + c_k T^k = 0_V$$
Thus, minimal polynomial of T and $[T]_\beta$ coincide.
  - ◇ **Cayley-Hamilton Theorem:** A square matrix satisfies its characteristic equation, *i.e.*,
  $$p_A(A) = 0_n$$
  - ◇ Similar matrices have the same characteristic and minimal polynomial.

Inner Product:

- **Inner Product Space:** An inner product space is a vector space V equipped with a *inner product* map
$$\langle \cdot , \cdot \rangle : V \times V \to F$$
which satisfies
  1. **Linearity in First Variable:** $\langle \boldsymbol{v_1} + \boldsymbol{v_2}, \boldsymbol{w} \rangle = \langle \boldsymbol{v_1}, \boldsymbol{w} \rangle + \langle \boldsymbol{v_2}, \boldsymbol{w} \rangle$
  2. **Conjugate Symmetry:** $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{w}, \boldsymbol{v} \rangle}$
  3. **Positivity:** $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ if $\boldsymbol{v} \neq \boldsymbol{0}$

  A few conclusions and properties:
  - ◇ $\langle \boldsymbol{0}, \boldsymbol{v} \rangle = \boldsymbol{0} = \langle \boldsymbol{v}, \boldsymbol{0} \rangle$
  - ◇ For $F = \mathbb{R}$, *conjugate symmetry* is *symmetry* (commutativity), *i.e.*, $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$
  - ◇ For $F = \mathbb{C}$, *conjugate symmetry* implies $\langle \boldsymbol{v}, \boldsymbol{v} \rangle$ is real
  - ◇ Note that
  $\langle \boldsymbol{v}, c\boldsymbol{w} \rangle = \overline{\langle c\boldsymbol{w}, \boldsymbol{v} \rangle} = \bar{c}\overline{\langle \boldsymbol{w}, \boldsymbol{v} \rangle} = \bar{c}\langle \boldsymbol{v}, \boldsymbol{w} \rangle$; and
  $\langle \boldsymbol{v}, \boldsymbol{w_1} + \boldsymbol{w_2} \rangle = \overline{\langle \boldsymbol{w_1} + \boldsymbol{w_2}, \boldsymbol{v} \rangle} = \overline{\langle \boldsymbol{w_1}, \boldsymbol{v} \rangle} + \overline{\langle \boldsymbol{w_2}, \boldsymbol{v} \rangle} = \langle \boldsymbol{v}, \boldsymbol{w_1} \rangle + \langle \boldsymbol{v}, \boldsymbol{w_2} \rangle$
  Therefore, $\langle \cdot , \cdot \rangle$ is *linear* in second variable, if $F = \mathbb{R}$, and *conjugate linear* in second variable, if $F = \mathbb{C}$
  - ◇ If $\langle \boldsymbol{v_1}, \boldsymbol{v} \rangle = \langle \boldsymbol{v_2}, \boldsymbol{v} \rangle$ for all $\boldsymbol{v} \in V$, then $\boldsymbol{v_1} = \boldsymbol{v_2}$

◇ **Length:** The length of a vector $\boldsymbol{v} \in (V, \langle \cdot, \cdot \rangle)$ is defined as

$$||\boldsymbol{v}|| := \langle \boldsymbol{v}, \boldsymbol{v} \rangle^{\frac{1}{2}}$$

Some important properties of $|| \cdot ||$:

1. ***Parallelogram Law*:** $||\boldsymbol{v} - \boldsymbol{w}||^2 + ||\boldsymbol{v} + \boldsymbol{w}||^2 = 2||\boldsymbol{v}||^2 + 2||\boldsymbol{w}||^2$
2. ***Cauchy-Schwarz inequality*:** $||\langle \boldsymbol{v}, \boldsymbol{w} \rangle|| \leq ||\boldsymbol{v}|| \cdot ||\boldsymbol{w}||$
3. ***Triangle inequality*:** $||\boldsymbol{v} + \boldsymbol{w}|| \leq ||\boldsymbol{v}|| + ||\boldsymbol{w}||$

◇ Few Standard Inner Products:

* For $V = \mathbb{R}^n$, $\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n$
* For $V = \mathbb{C}^n$, $\langle (z_1, z_2, \ldots, z_n), (w_1, w_2, \ldots, w_n) \rangle := z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n$
* For $V = M_n(\mathbb{R})$, $\langle A, B \rangle := \mathrm{tr}(B^tA)$
* For $V = M_n(\mathbb{C})$, $\langle A, B \rangle := \mathrm{tr}(B^\star A)$, where $B^\star = \bar{B}^t$

Orthogonality:

• **Orthogonality:** Let V be an inner product space. We say that $\boldsymbol{v}, \boldsymbol{w} \in V$ are *orthogonal* to each other if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$.
We may also indicate the same as $\boldsymbol{v} \perp \boldsymbol{w}$.

◇ ***Pythagoras Theorem***: If $\boldsymbol{v} \perp \boldsymbol{w}$, then $||\boldsymbol{v} + \boldsymbol{w}||^2 = ||\boldsymbol{v}||^2 + ||\boldsymbol{w}||^2$

◇ If $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}$ are *mutually orthogonal* to each other, then

$$||\boldsymbol{v_1} + \boldsymbol{v_2} + \cdots + \boldsymbol{v_n}||^2 = ||\boldsymbol{v_1}||^2 + ||\boldsymbol{v_2}||^2 + \cdots + ||\boldsymbol{v_n}||^2$$

◇ **Orthonormal Set:** A collection $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ is called *orthonormal* if

$$\langle v_i, v_j \rangle = \delta_{ij}$$

◇ If V is an *inner product space* of dimension n, then any *orthonormal* set has at most n elements. Moreover, any orthonormal set $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ is a basis.

◇ **Orthonormal Basis:** A basis β of V is called an *orthonormal basis* if it is an orthonormal set and a basis.

◇ If $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ is called *orthonormal basis* of V and $\boldsymbol{v} \in V$ is $\boldsymbol{v} = \sum_{i=1}^n c_i\boldsymbol{v_i}$, then $c_i = \langle \boldsymbol{v}, \boldsymbol{v_i} \rangle$, i.e,

$$\boldsymbol{v} = \sum_{i=1}^{n} \langle \boldsymbol{v}, \boldsymbol{v_i} \rangle \boldsymbol{v_i}$$

◇ Given an orthogonal basis $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$, we may form an orthonormal basis by normalizing. Consider

$$\tilde{\beta} = \left\{ \frac{\boldsymbol{v_1}}{||\boldsymbol{v_1}||}, \frac{\boldsymbol{v_2}}{||\boldsymbol{v_2}||}, \ldots, \frac{\boldsymbol{v_n}}{||\boldsymbol{v_n}||} \right\}$$

Normalized vector $\boldsymbol{v}$ is denoted as $\hat{\boldsymbol{v}}$

◇ **Gram-Schmidt Orthogonalization:** Let $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ be a linearly independent set.
Set $\boldsymbol{w_1} = \boldsymbol{v_1}$, $\boldsymbol{w_2} = \boldsymbol{v_2} - \dfrac{\langle \boldsymbol{v_2}, \boldsymbol{w_1} \rangle}{||\boldsymbol{w_1}||^2}\boldsymbol{w_1}$, $\boldsymbol{w_3} = \boldsymbol{v_3} - \dfrac{\langle \boldsymbol{v_3}, \boldsymbol{w_1} \rangle}{||\boldsymbol{w_1}||^2}\boldsymbol{w_1} - \dfrac{\langle \boldsymbol{v_3}, \boldsymbol{w_2} \rangle}{||\boldsymbol{w_2}||^2}\boldsymbol{w_2}$ and more generally

$$\boldsymbol{w_m} = \boldsymbol{v_m} - \frac{\langle \boldsymbol{v_m}, \boldsymbol{w_1} \rangle}{||\boldsymbol{w_1}||^2}\boldsymbol{w_1} - \cdots - \frac{\langle \boldsymbol{v_m}, \boldsymbol{w_{m-1}} \rangle}{||\boldsymbol{w_{m-1}}||^2}\boldsymbol{w_{m-1}}$$

Then $\boldsymbol{w_i} \neq \boldsymbol{0}$, $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_n}\}$ is an orthogonal set.
And $\mathrm{span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\} = \mathrm{span}\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ for $1 \leq k \leq n$

- ◇ *Corollary:* Any finite dimensional inner product space V admits an orthogonal basis.
- ◇ *Corollary (**Bessel's inequality**)*: If $\{\boldsymbol{e_1}, \dots, \boldsymbol{e_k}\}$ is an orthonormal set, then

$$||\boldsymbol{v}||^2 \geq |\langle \boldsymbol{v}, \boldsymbol{e_1} \rangle|^2 + \cdots + |\langle \boldsymbol{v}, \boldsymbol{e_k} \rangle|^2$$

- **Orthogonal Complement:** Let W be a subspace of an inner product space V . The orthogonal complement $W^\perp$ of W in V is defined as

$$W^\perp := \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0 \text{ for all } \boldsymbol{w} \in W\}$$

  - ◇ Note that $W \cap W^\perp = \{\boldsymbol{0}\}$
  - ◇ The $\dim_F(W^\perp) = n - k$, if $\dim_F(V) = n$ and $\dim_F(W) = k$.

- **Orthogonal Projection:** Let W be a subspace of an inner product space V. Using the decomposition $V = W \oplus W^\perp$, every $\boldsymbol{v} \in V$ can be written uniquely as $\boldsymbol{w} + \boldsymbol{w'}$, where $\boldsymbol{w} \in W$ and $\boldsymbol{w'} \in W'$. The map defined by

$$P : V \to V, \ \ P(\boldsymbol{v}) = \boldsymbol{w}$$

  is called the orthogonal projection of V on W.
  - ◇ P is a linear map, which is a projection.
  - ◇ $N_{(P)} = W^\perp$ and $R_{(P)} = W$

- V is a finite dimensional *inner product space*. We define a linear map:

$$\langle \cdot \,, \boldsymbol{v} \rangle : V \to F, \ \ \boldsymbol{w} \mapsto \langle \boldsymbol{w}, \boldsymbol{v} \rangle$$

  Thus we have a map

$$\Phi_V : V \to V^*, \ \ \boldsymbol{v} \mapsto \langle \cdot \,, \boldsymbol{v} \rangle$$

  Note that $\Phi_V(\boldsymbol{v_1} + \boldsymbol{v_2}) = \Phi(\boldsymbol{v_1}) + \Phi(\boldsymbol{v_2})$ and Φ_V is injective.
  And since $\dim_F(V) = \dim_F(V^*)$, Φ_V is surjective as well.
  - ◇ **Theorem (Finite Dimensional Riesz Representation):** Let V be a finite dimensional inner product space. If $T \in V^*$, then there exists $\boldsymbol{v_0} \in V$ such that $T(\boldsymbol{v}) = \langle \boldsymbol{v}, \boldsymbol{v_0} \rangle$ for $\boldsymbol{v} \in V$.

    Here, $\boldsymbol{v_0} = \overline{T(\boldsymbol{v_1})}\boldsymbol{v_1} + \overline{T(\boldsymbol{v_2})}\boldsymbol{v_2} + \cdots + \overline{T(\boldsymbol{v_n})}\boldsymbol{v_n}$ where $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_n}\}$ is a *orthonormal basis* for V

    Any linear map $T : V \to W$ induces a dual map $T^* : W^* \to V^*$ and for finite dimensional V and W, we have canonical bijections $\Phi_V : V \to V^*$ and $\Phi_W : W \to W^*$

Adjoint & Self-Adjoint:

- **Adjoint:** The composite map

$$(\Phi_V^{-1} \circ T^* \circ \Phi_W) : W \to V$$

  is called the *adjoint* of T, denoted as $T^\star$

$$c\boldsymbol{w} \xmapsto{\Phi_W} \langle \cdot \,, c\boldsymbol{w} \rangle_W = \bar{c} \, \langle \cdot \,, \boldsymbol{w} \rangle_W \xmapsto{T^*} \bar{c} \, \langle T(\cdot), \boldsymbol{w} \rangle_W \xmapsto{\Phi_V^{-1}} c \, \Phi_V^{-1}(\langle T(\cdot), \boldsymbol{w} \rangle_W)$$

  So, $\Phi_V^{-1}(\langle T(\cdot), \boldsymbol{w} \rangle_W) = T^\star(\boldsymbol{w})$. If we apply Φ_V on both side, we'll get

$$\langle T(\cdot), \boldsymbol{w} \rangle_W = \langle \cdot \,, T^\star(\boldsymbol{w}) \rangle_V$$

  So, given a map $T : V \to W$ between *inner product spaces*, a linear map $T^\star : W \to V$ satisfying

$$\langle T(\boldsymbol{v}), \boldsymbol{w} \rangle_W = \langle \boldsymbol{v}, T^\star(\boldsymbol{w}) \rangle_V$$

  for any $\boldsymbol{v} \in V, \boldsymbol{w} \in W$, is called the *adjoint* of T

- ◇ The adjoint of T (*i.e.*, $T^\star$) is *unique*.
- ◇ Let V be a finite dimensional inner product space and β be an *orthonormal* basis of V. If $T : V \to V$, then $[T^\star]_\beta = [T]_\beta^\star$
- ◇ Let two linear maps $S, T : V \to W$ and their adjoints are $S^\star, T^\star$. Then:
  1. $\langle \boldsymbol{v}, (S^\star + T^\star)(\boldsymbol{w})\rangle_V = \langle (S+T)(\boldsymbol{v}), \boldsymbol{w}\rangle_W$
     $\Rightarrow (S+T)^\star = S^\star + T^\star$
  2. $\langle \boldsymbol{v}, \bar{c}T^\star(\boldsymbol{w})\rangle_V = \langle (cT)\boldsymbol{v}), \boldsymbol{w}\rangle_W$
     $\Rightarrow (cT)^\star = \bar{c}T^\star$
  3. $(T^\star)^\star = T$

  Let $T : U \to V$ and $S : V \to W$ be linear maps, then $(ST)^\star = T^\star S^\star$

- **Self-Adjoint:** A linear map $T : V \to V$ is called *self-adjoint* if $T = T^\star$.
  Or equivalently, a linear map $T : V \to V$ is called *self-adjoint* if for any $\boldsymbol{v}, \boldsymbol{w} \in V$

  $$\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$$

  A linear map $T : V \to V$ is called *self-adjoint* if for any *orthonormal* basis β, $[T]_\beta$ is **Hermitian**, *i.e.*, $[T]_\beta^\star = [T]_\beta$
  - ◇ A self-adjoint operator has only *real eigenvalues*. Moreover, the *eigenspaces* are *orthogonal* to each other.
  - ◇ Let V be a finite dimensional real vector space. If $T : V \to V$ is a linear map, then T is a *self-adjoint* iff there exists *orthonormal eigenbasis*.
  - ◇ **Spectral Theorem for Self-Adjoint Operators:** Let V be a finite dimensional complex inner product space. If $T : V \to V$ is a self-adjoint map, then there exists an orthonormal eigenbasis.
  - ◇ **Normal:** A linear map $T : V \to V$ is called normal if $TT^\star = T^\star T$
  - ◇ **Spectral Theorem for Normal Operators:** Let V be a finite dimensional complex inner product space. If $T : V \to V$ is a linear map, then T is normal iff there exists an orthonormal eigenbasis.

Isometry:

- **Isometry:** The distance between two vectors $\boldsymbol{v}, \boldsymbol{w}$ is defined to be $||\boldsymbol{v} - \boldsymbol{w}||$ and a linear map is called an *isometry* if it's distance preserving, *i.e.*,

  $$||T(\boldsymbol{v}) - T(\boldsymbol{w})|| = ||\boldsymbol{v} - \boldsymbol{w}||$$

  Properties:
  - ◇ $\langle T(\boldsymbol{v}), T(\boldsymbol{w})\rangle = \langle \boldsymbol{v}, \boldsymbol{w}\rangle$
  - ◇ $T^\star T = I_V$